PRE-BUDGET CONSULTATION 2026

RETHINKING HEALTHCARE ACCESS THROUGH RESOURCE REALLOCATION

Building Canada's National Virtual Safe Space as a Demonstration of a More Accessible, Continuous, and Fiscally Productive Care Architecture

A Human-Supervised AI Demonstration for Continuous Mental-Health Support, Protected Professional Capacity, and Better Allocation of Public Resources

Toward a Free, Multilingual, 24/7 Foundational Safety Net Before, Between, and Alongside Professional Care

SUBMITTED BY

Dr. Yujia Zhu, D. Div., Ph.D

Solo-Founder and Executive Director, For A Safer Space (FASS) and FASSLING.AI

How can Canada redesign the allocation of human, technological, and financial capacity so that scalable forms of support become more continuously available, scarce professional human attention is protected for the work that genuinely requires human expertise, and fewer people are left without an appropriate and safe place to turn when professional care is unavailable, delayed, or temporarily out of reach?

Artificial intelligence is one instrument within that broader architecture. Its value should therefore not be measured primarily by how many professional tasks it can replace, but by whether it can responsibly create greater availability, continuity, multilingual access, navigation, structure, and scale while allowing therapists, psychologists, counsellors, social workers, physicians, and other helping professionals to devote more of their finite attention to judgment, safety assessment, therapeutic depth, complex decision-making, ethical discernment, intervention, and genuine human connection.

This proposal is therefore not fundamentally about replacing professionals with technology. It is about **redesigning how different forms of capacity are allocated across the healthcare-delivery system**. Computational capacity can be used where structure, repetition, information, navigation, and continuity can be delivered responsibly at scale. Professional human capacity should be protected for the situations in which human presence, expertise, accountability, and relationship are indispensable.

The underlying policy principle is therefore broader than an AI adoption strategy: **Use each form of capacity where it creates the greatest meaningful health value.**

Within that architecture, the role of artificial intelligence is clear: **AI should not replace human care. It should protect human attention for the moments when human care matters most.**

Dedication

“Whatever living creatures there are, with not a one left out … may all beings be happy.”
— *Karaṇīya Mettā Sutta* (Sutta Nipāta 1.8)

“无论世间有何众生，一个也不遗漏……愿一切众生安乐。”
——《慈经》（*Karaṇīya Mettā Sutta*，《经集》1.8）

This proposal did not emerge from a sudden idea, nor did it begin with artificial intelligence.

It began with listening.

It began more than a decade ago, through years of observing the Canadian healthcare system from the front line, advocating for people struggling to access care, building social innovations, delivering services, experimenting with new models, and, most importantly, listening to people across a wide range of socioeconomic circumstances describe what healthcare access actually looked like from where they stood.

This work has always been personal to me because behind every policy discussion about access is a human being who, at some point, quietly reaches the realization: **“I need help.”**

But recognizing that need does not mean knowing where to go next.

Some people finally gather the courage to reach out, only to encounter a waitlist. Some can afford one appointment but not the many attempts it may take to find the right professional. Some struggle to explain deeply personal experiences in an unfamiliar language. Some live too far from appropriate services. Some complete a short-term program while the circumstances of their lives continue long after the administrative endpoint of that program.

And many experience the difficult space between **“I need help”** and **“I have found the right person to help me.”**

For many years, I believed one of the central problems was relatively straightforward: there were not enough accessible mental-health professionals, and professional care was too often financially out of reach.

That belief led me to advocate for broader public access to professional mental-health services. The underlying principle has never changed: **No person who genuinely requires appropriate professional care should be prevented from receiving it primarily because they cannot afford it.**

But years of frontline experience gradually taught me that affordability, although fundamental, is only one dimension of accessibility.

People also wait.

They search.

They struggle to find the right clinical, cultural, linguistic, or relational fit.

They move between disconnected services.

They need support between appointments.

They complete programs while their lives continue.

They encounter geography, disability, limited operating hours, fragmented navigation, repeated intake processes, and the emotional burden of starting over after an unsuccessful attempt to find help.

Eventually, I came to understand something that became foundational to this proposal:

Formal coverage does not automatically mean practical accessibility.

A healthcare service can technically exist and still be unreachable.

A person can technically be covered and still have nowhere appropriate to turn at the moment support is needed.

A referral can be clinically appropriate while someone remains on a waitlist for weeks or months.

A program can successfully complete its mandate while the person's underlying needs continue.

That realization changed the question I was asking.

Instead of asking only:

"How do we put more money into healthcare?"

I began asking:

"How do we redesign the way existing healthcare resources are allocated, connected, and delivered so that every public dollar, every professional hour, and every appropriate technological capability can create more meaningful health outcomes?"

That question eventually became the foundation of this nearly 200-page proposal.

My first practical attempt to address one piece of the problem was **For A Safer Space (FASS)**.

FASS was created as a free human bridge for people who might otherwise have nowhere to turn while waiting, searching, navigating the system, or simply trying to understand what kind of support they needed.

Its purpose was never to replace professional care. It was to help address the difficult interval between recognizing a need and reaching the right professional.

FASS showed me how meaningful an accessible human bridge can be.

But it also taught me an equally important lesson:

Human beings can offer extraordinary compassion, judgment, wisdom, care, accountability, and presence. But human availability is finite.

A counsellor can speak with only one person at a time.

A therapist has only so many hours in a day.

A supervisor can responsibly oversee only so much work.

Another language requires another form of human capacity.

Another hour of service requires another hour of labour.

And when an organization removes the financial barrier for the person seeking support, the underlying cost does not disappear.

Someone else still has to carry it.

That realization eventually led to **FASSLING™** and to another question:

If technology can responsibly create greater availability, could we use it not to replace human care, but to protect human care for the moments when human beings are truly irreplaceable?

FASSLING™ was never conceived as a replacement for psychotherapy, psychologists, counsellors, social workers, physicians, or other healthcare professionals.

It explored something different.

Could certain appropriately bounded forms of emotional support, guided reflection, non-clinical coaching, psychoeducation, navigation, multilingual interaction, structured exercises, and continuity become more continuously available without requiring one additional professional to become personally available for every additional appropriate interaction?

Over time, I realized that **FASS and FASSLING™ were revealing something much larger than the strengths and limitations of two individual projects**.

Together, they became real-world cases through which a broader question about healthcare financing and delivery could be examined.

FASS explored the value—and the limits—of a free human-delivered bridge.

FASSLING™ explored whether technology could expand appropriate availability beyond some of those human-capacity limits.

And together they led me toward a deeper conclusion:

The problem is not only how much healthcare capacity exists. It is also how different forms of capacity are allocated.

If public budgets are finite, professional human attention is finite, and people still continue to fall through gaps despite substantial public expenditure, then the question cannot only be how much more capacity we need.

We must also ask how the capacity we already have is being used.

Which functions genuinely require professional expertise?

Which primarily require availability, information, structure, repetition, navigation, or continuity?

Where is scarce professional human attention currently being spent?

Where could technology responsibly create greater availability?

Where does the individual fall out of the system?

Where do services fail to connect?

And how should public funding be structured so that the right form of capacity reaches the right need?

These are not merely technology questions.

They are healthcare budgeting questions.

This is why the proposal uses the specific cases of **FASS, FASSLING™, and mental-health service delivery to demonstrate a possible budgeting-reallocation model**.

The approximate 50/50 funding hypothesis proposed in the document is not intended to become a permanent formula, nor does it suggest that half of Canada's healthcare budget should be directed toward artificial intelligence.

It is an experiment.

One portion of a defined or comparable funding envelope could purchase greater **availability, continuity, multilingual reach, accessibility, navigation, infrastructure, computation, and scale**.

Another portion could protect **professional judgment, therapeutic depth, complex care, safety, intervention, supervision, ethical accountability, and genuine human connection**.

The real question is not whether AI is cheaper than people.

It is whether every healthcare function should continue to be financed through the same form of capacity simply because, historically, human labour was the only mechanism available.

The proposed 50/50 allocation may turn out to be wrong.

One population may require substantially more professional capacity. Another low-risk, highly structured use case may be able to rely more heavily on scalable infrastructure. Different languages, disabilities, cultures, risk levels, professional requirements, and service environments may require completely different allocations.

That is precisely why the model should be tested rather than assumed.

The number itself is not the most important idea.

The more important question is:

Can we separate, where appropriate, the cost of availability from the cost of expertise, and allocate each form of capacity to the functions where it creates the greatest meaningful health value?

This leads to another conclusion that has become central to my thinking:

Professional human attention is itself a scarce public resource.

A therapist's hour can only be spent once.

If a highly trained professional is spending that hour performing a standardized or highly repeatable function that another form of capacity could responsibly support, that same hour cannot simultaneously be spent with someone whose circumstances require complex judgment, trauma-informed care, risk assessment, ethical discernment, therapeutic depth, intervention, or genuine human presence.

The cost of professional time is therefore not only financial.

It also includes the **opportunity cost of human attention**.

That is why I do not believe the right policy question is:

"How many therapists can artificial intelligence replace?"

That framing is too narrow and misses the larger opportunity.

The more consequential question is:

How can Canada redesign the allocation of human, technological, and financial capacity so that scalable forms of support become more continuously available, scarce professional human attention is protected for the work that genuinely requires human expertise, and fewer people are left without an appropriate and safe place to turn when professional care is unavailable, delayed, or temporarily out of reach?

This proposal is therefore not fundamentally about replacing professionals with technology.

It is about **redesigning how different forms of capacity are allocated across the healthcare-delivery system**.

Use scalable capacity where scalability is appropriate.

Protect professional human capacity where human expertise is indispensable.

Connect the two rather than forcing one to replace the other.

Technology should not make human beings more disposable.

It should make human attention more precious.

This is also why I do not see patients, healthcare professionals, taxpayers, government, and social innovators as groups whose interests must necessarily conflict.

They can be part of the same solution.

Patients need greater access, affordability, continuity, dignity, and appropriate care.

Healthcare professionals need their expertise to be protected and their limited attention increasingly concentrated on work that genuinely requires human capability.

Government and taxpayers need to know not simply how much is being spent, but what meaningful access, health outcomes, continuity, and public value that expenditure is creating.

And social entrepreneurs should be able to identify neglected needs, listen, experiment, build prototypes, and demonstrate what may be possible without being expected to personally carry essential public infrastructure forever.

That lesson has also become deeply important to me.

A social entrepreneur can enter a gap before institutions are ready to follow.

They can listen when a problem has not yet become a formal indicator.

They can experiment when no established pathway exists.

They can contribute labour, expertise, credibility, financial resources, and pieces of their own lives to demonstrate another possibility.

But innovation and permanent infrastructure are not the same responsibility.

Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.

If rigorous evidence eventually demonstrates that a particular capability serves a persistent and legitimate public need, responsibility should become proportionate to the importance of that capability.

Government does not necessarily need to operate every component directly. It can regulate, fund, procure, contract, coordinate, certify, evaluate, establish standards, or use other institutional arrangements.

But the public capability itself should not remain indefinitely dependent on one founder, one nonprofit, one donor, one vendor, one AI model, or one technology cycle.

Government should protect the promise, not the platform.

And this brings me to the larger purpose of the entire proposal.

The ultimate goal is not simply to create a better AI-enabled mental-health service.

Mental health is the demonstration. FASS and FASSLING.AI™ are real-world examples of social experimentation that have achieved substantial results despite severe constraints on human and social capital resources and the absence of funding/grants/donation.

The 50/50 allocation is a testable budgeting hypothesis.

The larger ambition is to explore how the healthcare-delivery system itself might be redesigned.

The question I have ultimately been trying to answer is remarkably simple:

How can we make healthcare truly universally accessible?

To me, universal healthcare should mean more than universal coverage.

Coverage is foundational, but it does not automatically guarantee meaningful access.

A person can be covered and still wait.

A person can be covered and still be unable to find the right professional.

A person can be covered and still encounter language, geography, disability, scheduling, cultural, or navigational barriers.

A person can be covered and still fall into the spaces between services.

I therefore believe the next evolution of universal healthcare must move from **universal coverage toward universal practical accessibility**.

That means a healthcare system in which people can actually reach the right form of care at the right level and, as far as responsibly and realistically possible, at the right time.

It means a system that is more holistic, continuous, person-centred, and deliberately designed around the realities of how people actually seek and receive care.

It means reducing unnecessary out-of-pocket barriers where appropriate while also refusing to assume that every improvement must simply be financed by continuously adding new funding on top of an unchanged delivery architecture.

It means asking not only:

“How much are we spending?”

but:

“What is every dollar actually purchasing, and could those resources be organized more intelligently?”

It means asking where professional expertise creates the greatest value.

It means asking where technology can responsibly create appropriate abundance.

It means asking where people disappear between services.

And it means asking how those pieces can be connected so that the person experiences something closer to one coherent healthcare continuum rather than a collection of disconnected programs.

The deepest aspiration of this proposal is therefore a **healthcare-delivery system that is designed, as far as responsibly and realistically possible, to leave no one behind**.

That does not mean promising the impossible.

No healthcare system can eliminate every waitlist, every scarcity, every provider mismatch, or every barrier.

Professional capacity will always remain finite.

Public resources will remain finite.

Technology will remain imperfect.

But scarcity does not always have to become abandonment.

A program may end.

The person's life does not.

A professional may temporarily be unavailable.

That does not necessarily mean the person must have absolutely nowhere appropriate to turn.

And a healthcare system can acknowledge finite resources while still asking whether those resources could be allocated, connected, and delivered more intelligently.

That is the larger principle I hope this proposal can test.

If one day this budgeting model is further researched, independently evaluated, refined, and successfully implemented, I hope Canada can move toward a broader understanding of universal healthcare: not simply **"everyone is covered,"** but **"everyone can meaningfully reach appropriate care."**

If such a model proves viable, its significance may extend beyond Canada.

The lesson for other countries would not be:

"Use the same AI."

Nor would it be:

"Adopt the same 50/50 funding ratio."

The potentially transferable lesson would be a way of thinking:

Identify what form of capacity each healthcare function actually requires, fund that capacity deliberately, protect scarce human expertise where it is indispensable, use scalable infrastructure where it can responsibly create abundance, connect the pieces into a continuum, and measure success by meaningful human outcomes.

That is the system-level possibility I believe is worth exploring.

This nearly 200-page proposal is the first time I have brought together more than a decade of frontline observation, listening, advocacy, social innovation, service delivery, experimentation, creation, and system-level thinking into one comprehensive model.

It has now been formally submitted to the Government of Canada.

At this point, the next stage should belong to evidence and public institutions.

If the 50/50 hypothesis is wrong, test another allocation.

If certain functions should remain entirely human-led, preserve them.

If technology introduces hidden costs, unacceptable risks, or inequities, acknowledge them.

If the architecture does not work, redesign it or stop.

But if rigorous evaluation demonstrates that a different allocation of human, technological, and financial capacity can expand practical access, strengthen continuity, protect professional expertise, preserve safety and dignity, and create greater meaningful health value from finite resources, then the next question becomes institutional:

Is government prepared to carry the responsibility required to make a demonstrated public capability durable?

I have taken this particular question as far as I reasonably can as a social innovator.

I listened.

I advocated.

I built.

I tested.

I learned from what worked and what did not.

I translated those experiences into a system-level hypothesis.

And now I have formally placed that hypothesis before government.

Social innovation can reveal the possibility. Evidence can test it. Professional expertise can safeguard it. Government must decide whether the demonstrated public value warrants implementation and institutional durability.

So, in that sense, my chapter with this particular Canadian healthcare question can temporarily close.

There are many other social questions, places, and systems in the world that I remain curious about and hope to continue exploring.

But if I ever look back and this proposal has contributed in some small way to meaningful healthcare reform, I hope what remains is not the particular AI model, the original funding ratio, or even the organizations that helped generate the idea.

I hope what remains is the principle behind all of it:

Universal healthcare should not merely mean that everyone is covered. It should mean that the system is designed so that people can actually reach the care they need.

And a truly human-centered healthcare system should strive toward an even simpler promise:

No one should be left behind simply because the architecture of the system failed to connect them to the right care.

Technology will change.

Providers will change.

Funding models will change.

Governments will change.

But that principle can endure.

And if Canada can eventually make it real—not simply as an aspiration, but as reliable, accountable, human-centered public infrastructure—then perhaps a person experiencing one of the most vulnerable moments of their life will not need to understand the complicated funding streams, professional scopes, technologies, regulations, or institutional architecture behind the system.

They will only need to know something much simpler:

My country has my back.

Contents

This submission presents the policy case, proposed service architecture, safeguards, funding hypothesis, evaluation framework, and requested federal action.

Executive Summary

Canada has made substantial investments in healthcare, mental-health services, employee assistance programs, community-based supports, healthcare modernization, digital infrastructure, and the mental-health workforce. These investments are important and should continue. Yet a consequential challenge remains within the architecture of access itself: a person can live within a publicly funded healthcare system and still experience periods in which the care or support they need is not practically reachable.

This proposal begins from that gap.

It is the result of more than a decade of frontline observation, advocacy, service delivery, social innovation, experimentation, and sustained listening to people across different socioeconomic circumstances. Over time, I repeatedly encountered the difference between the healthcare system as it exists institutionally and the healthcare system as people actually experience it. A service may exist without being readily accessible. A person may have formal coverage and still wait. A referral may be appropriate while the next appointment remains weeks or months away. A qualified professional may exist somewhere in the system while remaining inaccessible because of language, geography, disability, operating hours, provider mismatch, specialization, cost, cultural fit, or uncertainty about where to begin. A short-term program may fulfill its administrative mandate while the circumstances that brought the person into care continue.

These experiences gradually changed the question I was asking. The issue was no longer simply:

How do we put more money into healthcare?

It became:

How do we redesign the way existing and future healthcare resources are allocated, connected, and delivered so that every public dollar, every professional hour, and every appropriate form of technological capacity can create greater meaningful health value?

That is the larger policy question behind this proposal.

The underlying proposition is not that additional public investment is unnecessary. Genuine shortages of professionals, services, infrastructure, specialized care, or other forms of capacity may require additional resources, and this proposal should not be interpreted as an argument against such investment. Rather, it introduces a complementary budgeting question: **Before assuming that every improvement in access requires a proportionate increase in expenditure, can Canada examine more carefully what different forms of public funding are purchasing, which functions genuinely require scarce professional expertise, which primarily require availability or continuity, and whether those forms of capacity can be organized more deliberately?**

Mental health is the first and most fully developed demonstration environment for that broader healthcare resource-allocation question.

It is particularly useful because the difference between formal coverage and practical accessibility is visible throughout the mental-health continuum. People often recognize that they need support long

before the right professional becomes available. They may wait weeks or months for care, spend substantial time and money trying to find a provider who is the right clinical, cultural, linguistic, or relational fit, lose access when a time-limited program ends, need support outside conventional service hours, encounter geographic or disability-related barriers, struggle to obtain help in a preferred language, or simply remain uncertain about what kind of service they need and where to begin.

In many cases, the problem is not an absence of willingness to seek help. It is the absence of an **appropriate, accessible, and continuous pathway through which the person can move from recognizing a need to reaching the right level of care**. This distinction is central to the proposal because it suggests that universal healthcare should ultimately be evaluated not only by whether people are formally covered, but by whether they can practically reach, navigate, and remain connected to appropriate care.

FASS and FASSLING™ as Real-World Demonstration Cases

My first practical attempt to address one part of this problem was **For A Safer Space (FASS)**.

FASS's unlimited free mental health counseling, psychotherapy, emotional support, and coaching sessions globally provided by clinical interns were developed as a free human bridge for people who might otherwise have nowhere to turn while waiting, searching, navigating, or trying to understand what kind of help they needed. Its purpose was never to replace professional care. It was to reduce the difficult gap between **"I need help"** and **"I have found the right professional."**

FASS demonstrated the value of accessible human support. It also revealed a structural limitation that goodwill and free access alone could not overcome: **human availability is finite**.

A counselor can speak with only one person at a time. A therapist has only so many hours in a day. A supervisor can responsibly oversee only so much work. Additional language coverage requires additional human capacity. Longer service hours require additional staffing. Greater demand creates additional requirements for recruitment, training, coordination, administration, quality assurance, supervision, and funding. Removing the fee charged to the individual does not eliminate the underlying cost of human-delivered support. It changes who carries that cost.

That realization led to **FASSLING.AI**, which explored another part of the same access problem.

FASSLING™ was not designed to replace psychotherapy, counseling, psychological care, social work, medicine, or professional judgment. It explored whether appropriately bounded forms of emotional support, guided reflection, non-clinical coaching, life-skills assistance, psychoeducation, navigation, multilingual interaction, structured exercises, and continuity could become more continuously available without requiring another professional to become personally available for every additional appropriate interaction. It's offered 24/7, multilingual support, via both audio and text, as a virtual safe space for help-seekers around the globe.

Together, FASS and FASSLING™ became more than two complementary social innovations. They became **real-world cases through which a larger healthcare-budgeting question could be examined**.

FASS demonstrated the value and limits of a free human bridge.

FASSLING™ explored whether technological capacity could extend appropriate availability beyond some of those human-capacity limits.

Together, they led to a deeper conclusion:

The healthcare-access problem is not only about how much capacity exists. It is also about how different forms of capacity are allocated, connected, and used.

From an Access Problem to a Resource-Allocation Problem

Canada's mental-health access challenge is therefore not only a question of whether enough professionals exist. It is also a question of **how scarce professional human attention is allocated across different categories of need**.

Highly trained professionals may spend meaningful portions of their time on structured intake, standardized guidance, routine service navigation, basic psychoeducation, documentation, repetitive exercises, routine follow-up, or other structured functions. These activities may be valuable and necessary, but they coexist with responsibilities that depend far more heavily on distinctly human expertise: complex professional judgment, therapeutic relationships, trauma-informed care, ethical discernment, risk assessment, ambiguity, specialized intervention, and sustained accountable human presence.

Where professional supply is finite, the allocation of professional time becomes more than a workforce-management issue. It becomes a matter of **access, service quality, public-resource stewardship, and healthcare-system design**.

A therapist's hour can only be spent once. A psychologist, counsellor, social worker, physician, or other professional who spends that hour performing a highly standardized function cannot simultaneously spend the same hour with a person whose circumstances require complex assessment, trauma-informed judgment, therapeutic depth, ethical reasoning, risk management, intervention, or sustained human presence.

The cost of professional time is therefore not only financial. It also includes the **opportunity cost of human attention**.

For this reason, the scarce resource being protected in this proposal is not only money. It is also **professional human attention**.

Artificial intelligence creates an opportunity to reconsider this architecture without diminishing the importance of helping professionals. The relevant question is not how many therapists, psychologists, counsellors, social workers, physicians, or other professionals technology can replace. That framing is too narrow and reduces a system-design problem to a labour-substitution question.

The more consequential question is:

How can Canada redesign the allocation of human, technological, and financial capacity so that scalable forms of support become more continuously available, scarce professional human attention is protected for the work that genuinely requires human expertise, and fewer people are

left without an appropriate and safe place to turn when professional care is unavailable, delayed, or temporarily out of reach?

Artificial intelligence is one instrument within that broader architecture. Its value should therefore not be measured primarily by how many professional tasks it can replace, but by whether it can responsibly create greater availability, continuity, multilingual access, navigation, structure, and scale while allowing professionals to devote more of their finite attention to judgment, safety assessment, therapeutic depth, complex decision-making, ethical discernment, intervention, and genuine human connection.

The underlying policy principle is therefore broader than an AI-adoption strategy:

Use each form of capacity where it creates the greatest meaningful health value.

Within that architecture, the role of artificial intelligence is clear:

AI should not replace human care. It should protect human attention for the moments when human care matters most.

The National Virtual Safe Space Demonstration

I therefore respectfully recommend that the Government of Canada establish a **Human-Supervised National Virtual Safe Space Demonstration Initiative**: a controlled, staged, time-limited, independently evaluated demonstration of free-at-the-point-of-appropriate-use, multilingual, trauma-informed, 24/7 AI-assisted emotional support, guided reflection, non-clinical coaching, holistic life-skills support, routine psychoeducation, and service navigation.

The initiative would operate within clearly defined professional boundaries, privacy requirements, accessibility standards, safety protocols, qualified human governance, escalation pathways, and reliable connections to professional, community, crisis, or emergency services.

The proposed National Virtual Safe Space would **not** replace psychotherapy, counselling, regulated clinical care, crisis services, emergency services, diagnosis, prescribing, or licensed professional judgment. Nor should it be presented as a lower-cost substitute for services that properly require qualified professional expertise.

Its purpose would be to test whether Canada can create a missing layer around existing services: **an always-available foundational safety net before, between, and alongside professional care**.

A waiting list should not automatically become a support vacuum. A referral should not necessarily mean that all appropriate support disappears while the person waits. An unsuccessful provider match should not require a return to complete isolation. The completion of a short-term program should not imply that the person's underlying circumstances have ended.

The program may end; the person's life does not.

The importance of a **National Virtual Safe Space** lies precisely in its ability to address a form of healthcare scarcity that traditional service expansion alone cannot fully eliminate. Even in a well-funded professional system, human capacity will remain finite, conventional operating hours will remain limited,

geographic and linguistic barriers will persist, and people will continue to experience moments when the right professional is not immediately available. A national foundational layer could reduce the likelihood that those temporary gaps become complete abandonment.

Its national significance is therefore larger than the existence of another digital service. It creates a practical environment in which Canada can test whether continuity can become a more dependable public capability—less dependent on geography, office hours, the duration of individual programs, or an individual's ability to pay for another supportive interaction.

The National Virtual Safe Space is therefore also the **concrete mechanism through which the broader resource-allocation hypothesis of this proposal can be tested**.

A Budgeting-Reallocation Demonstration

The proposed initiative is not fundamentally an AI project. It is a **healthcare budgeting and resource-allocation demonstration conducted through a concrete national mental-health service architecture**.

Historically, greater supportive availability generally required greater human labour. More users required more professional or para-professional hours. Longer operating hours required additional staffing. Broader language coverage required additional human capacity. Expanded geographic reach often required parallel organizational and workforce expansion.

In practical terms, **availability and expertise were frequently purchased through the same scarce input: human labour**.

Artificial intelligence changes that economic relationship for some suitable and carefully bounded functions.

Once appropriate infrastructure, governance, privacy protections, cybersecurity, safety systems, accessibility features, multilingual performance, qualified human supervision, professional boundaries, and independent evaluation are established, some forms of structured support may potentially become available to additional people without requiring professional labour to increase in direct proportion to every additional appropriate interaction.

This does not mean that AI makes support costless. Safe public infrastructure requires substantial and continuing expenditure on computational resources, model access, secure hosting, cybersecurity, privacy architecture, accessibility, multilingual quality, safety systems, monitoring, governance, human supervision, professional involvement, and independent evaluation.

The relevant distinction is therefore not between **"expensive humans"** and **"cheap AI."**

It is between **different categories of public expenditure that purchase different forms of capacity**.

The demonstration should therefore test whether a defined existing or comparable public funding envelope can be organized around two complementary forms of capacity.

One is **scalable digital infrastructure**, which purchases availability, continuity, multilingual reach, accessibility, repeatability, navigation, and scale.

The other is **protected professional human capacity and governance**, which purchases judgment, expertise, therapeutic depth, complex-care capability, safety, intervention, relationships, supervision, ethical accountability, and genuine human presence.

The premise is not that one form of capacity can replace the other. It is that they create **different forms of public value and may therefore be financed and deployed more deliberately**.

The proposed **National Virtual Safe Space** is the concrete mechanism through which this broader resource-allocation principle can be tested. Its significance lies not only in creating a continuously available layer of mental-health support, but in providing government with a bounded and measurable environment in which two different forms of public capacity can be financed side by side and evaluated as one integrated system.

The National Virtual Safe Space can test whether public funding directed toward scalable digital infrastructure can purchase greater availability, continuity, multilingual reach, accessibility, navigation, and structured support, while a protected share of the same demonstration envelope preserves professional human capacity for judgment, therapeutic depth, complex care, safety, intervention, supervision, and accountability.

In this sense, the National Virtual Safe Space is more than a service proposal. **It is the practical demonstration case for a possible new approach to healthcare budgeting**, allowing Canada to examine whether different forms of capacity can be funded according to the functions they are best suited to perform while still operating as one connected continuum of care.

The 50/50 Funding Hypothesis

For demonstration purposes, I propose testing an approximate **50/50 funding hypothesis**.

Approximately half of the demonstration envelope could support AI-enabled public digital infrastructure, including model access, inference and computational costs, secure hosting, privacy architecture, cybersecurity, multilingual capability, accessibility systems, trauma-informed design, automated safety and boundary mechanisms, maintenance, monitoring, testing, and independent technical evaluation.

The remaining portion could support protected human professional capacity and governance, including therapists, counsellors, psychologists, social workers, supervisors, intervention and escalation specialists, clinical advisers, service navigators, safety reviewers, quality-assurance functions, professional-boundary oversight, culturally responsive review, referral pathways, and institutional governance.

The proposed ratio is **not a permanent funding formula, ideological commitment, or prediction of the eventual optimal allocation**.

It is a testable budgeting hypothesis.

One population may ultimately require substantially more professional human capacity. Another lower-risk and highly structured use case may be able to rely more heavily on scalable infrastructure. Different

languages, disabilities, cultures, clinical complexities, risk levels, supervision requirements, and service environments may require different allocations.

If 50/50 is wrong, government should test another allocation.

The number itself is less important than the method.

The underlying question is:

Can Canada separate, where appropriate, the cost of availability from the cost of expertise, finance each form of capacity according to the functions it is best suited to perform, connect them into one coherent continuum, and evaluate whether the resulting architecture produces greater meaningful public value?

The deeper fiscal proposition is therefore neither simply to spend more nor simply to spend less.

It is to determine whether Canada can generate **greater practical access, stronger continuity, better protection of professional capacity, and greater meaningful health value from finite resources by allocating each form of capacity more deliberately**.

The 50/50 demonstration should therefore evaluate not merely whether each half of the budget performs as expected, but **how the two forms of capacity interact**. Greater digital availability has limited public value if it creates disproportionate supervision, escalation, remediation, or governance burdens. Likewise, preserving substantial professional capacity does not necessarily improve access if scarce professional hours remain concentrated on functions that other forms of capacity could safely and responsibly support.

The appropriate long-term allocation, if any, should therefore be determined by evidence rather than assumption.

A Four-Part Public-Value Proposition

If designed and evaluated responsibly, this architecture could create value for several constituencies simultaneously.

For the public, it could provide immediate, free-at-the-point-of-appropriate-use, multilingual, text- and audio-enabled access to an appropriate foundational layer while people wait for care, move between appointments, search for the right provider, navigate uncertainty, or adjust after a time-limited program ends. The value would not lie in pretending that AI is equivalent to therapy. It would lie in reducing the periods during which people have no appropriate support at all.

For helping professionals, the model could protect and elevate the work that most requires human capability. If appropriate structured, repetitive, informational, and navigational functions can be responsibly supported through scalable infrastructure, professionals may be able to devote a greater proportion of their finite attention to therapeutic relationships, complex care, trauma, professional judgment, risk assessment, ethical ambiguity, intervention, supervision, quality assurance, and genuine human connection.

The relevant workforce principle is therefore **job evolution, not job elimination**.

For government and taxpayers, the demonstration would test whether a comparable public funding envelope can serve more people, reduce avoidable gaps in continuity, protect the value of professional hours, and produce greater public value through more deliberate allocation of human and computational capacity.

The fiscal question is not:

"How cheaply can AI conduct a conversation?"

It is:

"Can the complete human-AI architecture produce greater practical access, continuity, safety, equity, protected professional capacity, and meaningful public value from a comparable funding envelope?"

Fiscal productivity should therefore mean **greater meaningful public value per dollar**, not simply lower spending.

For social entrepreneurs, a mature public model could reduce the expectation that individual founders or small nonprofit organizations should personally carry indefinitely the financial, operational, technological, ethical, regulatory, professional, and legal burden associated with maintaining a public-interest service at scale.

Social entrepreneurs can identify unmet needs, listen, experiment, build prototypes, challenge assumptions, and demonstrate new possibilities.

They should not be required to become permanent private substitutes for public infrastructure.

Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.

These constituencies should not automatically be treated as having competing interests. Better system design may allow their interests to reinforce one another. Greater scalable availability may protect professional attention. Better allocation of professional attention may strengthen complex human care. Greater continuity may reduce the number of people who disappear into the gaps between formal services. Better resource stewardship may make broader access more sustainable.

The objective is therefore not simply to spend more, spend less, automate more, or employ fewer people.

It is to **design the system better**.

From Universal Coverage to Universal Practical Accessibility

The longer-term policy vision can be expressed through two complementary aspirations.

The **Care Guarantee** states:

No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.

The **Continuity Guarantee** states:

No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.

The two guarantees address different forms of scarcity.

The Care Guarantee addresses the affordability of professional expertise.

The Continuity Guarantee addresses the availability gaps created by finite professional capacity.

Neither replaces the other.

Expanding AI-enabled continuity cannot substitute for improving access to professional care, while increasing professional funding alone cannot eliminate every temporal, geographic, linguistic, relational, accessibility, or navigational gap.

A more complete system should therefore pursue both.

Professional care provides depth. AI-enabled infrastructure can provide appropriate continuity and scale. Qualified human supervision provides safety and accountability. Government provides standards, equitable access, financing architecture, institutional durability, governance, evaluation, and public responsibility.

Together, these capacities could form a more complete continuum of support.

But the deeper significance of these guarantees extends beyond mental health.

They point toward a more complete understanding of **universal healthcare itself**.

Universal healthcare should not mean only that everyone is formally covered.

It should increasingly mean that people can **practically reach the appropriate care they need**.

A person can be formally covered and still wait.

A person can be formally covered and still struggle to find the right professional.

A person can be formally covered and still encounter language, geography, disability, cultural, scheduling, or navigational barriers.

A person can be formally covered and still fall into the spaces between services.

The next evolution of universal healthcare should therefore move, as far as responsibly possible, from **universal coverage toward universal practical accessibility**.

A Healthcare System Designed to Leave Fewer People Behind

The long-term objective is therefore **not unlimited AI psychotherapy**.

It is not maximum automation.

It is not indiscriminate cost reduction.

And it is not the assumption that every healthcare problem can be solved without additional public investment.

The larger objective is to test whether healthcare delivery can be designed around a more deliberate relationship among **funding, infrastructure, professional expertise, technological capacity, continuity, and human need**.

The aspiration is a system in which people can actually reach appropriate care; in which financial hardship does not become the principal barrier to necessary professional expertise; in which language, geography, disability, fragmented navigation, provider mismatch, or the end of one finite program do not unnecessarily leave people without an appropriate pathway; in which highly trained professionals can increasingly devote their scarce attention to the work that genuinely requires them; and in which public resources are evaluated not only by how much is spent, but by how much practical access, continuity, safety, equity, professional capacity, and meaningful health value they create.

No healthcare system can eliminate every waitlist, every scarcity, every mismatch, or every barrier.

Professional capacity will remain finite.

Public resources will remain finite.

Technology will remain imperfect.

But **scarcity does not always have to become abandonment**.

A program may end.

The person's life does not.

A professional may temporarily be unavailable.

That does not necessarily mean the person must have absolutely nowhere appropriate to turn.

A universal healthcare system should therefore not merely cover everyone.

It should be designed, as far as responsibly and realistically possible, so that fewer people are left behind because the architecture of the system failed to connect them to the right care.

Mental Health Is the Demonstration; Healthcare Redesign Is the Larger Question

This proposal should therefore be understood at several levels.

FASS and FASSLING™ are the concrete real-world cases.

The National Virtual Safe Space is the practical demonstration architecture.

The approximately 50/50 allocation is the initial testable budgeting hypothesis.

Mental health is the demonstration environment.

But the deeper subject is **healthcare-delivery architecture**.

The proposal does not assume that the same technology, funding ratio, or division of labour should automatically be applied elsewhere in healthcare. Different healthcare domains involve different clinical risks, professional scopes, regulatory requirements, technologies, workflows, and populations.

If the demonstration succeeds, the transferable contribution is therefore not a universal 50/50 formula or a mandate to deploy AI throughout healthcare.

It is a method of asking better questions.

Which functions genuinely require scarce professional expertise?

Which primarily require availability, infrastructure, navigation, coordination, monitoring, or continuity?

Which costs are currently bundled together because historical service models required them to be?

Where does professional attention create the greatest human value?

Where does the system lose people?

Where can technology responsibly create appropriate abundance?

Where should any released professional capacity be reinvested?

And how can the different components be connected so that a person experiences something closer to one coherent healthcare continuum rather than a collection of disconnected programs?

These are not merely questions about artificial intelligence.

They are questions about how healthcare systems are financed and designed.

The ultimate policy question is therefore:

How can Canada redesign the allocation, connection, and delivery of finite healthcare resources so that every public dollar, every professional hour, and every appropriate technological capability is

used where it can create the greatest meaningful health value—and so that fewer people are left behind by the architecture of the system itself?

The Government Decision

This proposal therefore asks the Government of Canada to do more than finance another digital service or technology platform.

It asks government to test whether responsible technological capacity, protected professional human expertise, qualified human supervision, rigorous governance, public accountability, and deliberate budgeting reallocation can be combined into a more accessible healthcare-delivery architecture.

The demonstration should be designed so that government can reach either conclusion with confidence.

If the evidence shows that the model improves accessibility, continuity, professional-capacity use, safety, user dignity, workforce sustainability, equity, connection to appropriate human care, and fiscal productivity, Canada would have a basis for considering further development.

If the evidence shows unacceptable safety risks, inadequate public trust, poor value, weak outcomes, excessive supervision burdens, inequitable performance, or governance problems that cannot be satisfactorily addressed, the model should not advance in its proposed form.

That neutrality is essential.

The purpose of a demonstration is not to prove that a predetermined solution is correct.

It is to determine whether a promising new architecture deserves a place in the Canadian public system.

Social innovation can reveal the possibility.

Evidence can test it.

Professional expertise can safeguard it.

Government must determine whether the demonstrated public value warrants implementation and institutional durability.

If the evidence ultimately supports the model, Canada would have an opportunity not merely to modernize mental-health service delivery, but to demonstrate a broader evolution in universal healthcare: **professional care when professional care is needed, an appropriate place to turn when that care is not immediately available, human expertise protected for the moments when human beings matter most, and public resources organized so that each form of capacity is used where it can create the greatest meaningful health value.**

In its simplest form, the aspiration behind this proposal is therefore larger than artificial intelligence, larger than mental health, and larger than any one platform or organization:

Universal healthcare should not merely mean that everyone is covered. It should mean that the system is designed so that people can actually reach the care they need.

And the ultimate ambition is equally simple:

A healthcare-delivery system that, as far as responsibly and realistically possible, leaves no one behind.

Proposal at a Glance

The proposed **Human-Supervised Virtual Safe Space Demonstration Initiative** would test whether Canada can combine scalable AI-enabled availability with protected professional human capacity to create a safer, more continuous, accessible, and fiscally productive mental-health support architecture. The proposal is grounded in a simple resource-allocation insight: Canada's challenge is not only how many qualified professionals are available, but also how scarce professional human attention is deployed across different categories of need.

The initiative would not replace psychotherapy, counselling, regulated clinical care, crisis services, emergency services, or professional judgment with artificial intelligence. Rather, it would test whether AI can responsibly support appropriately bounded functions that are structured, repetitive, informational, navigational, or continuity-oriented, thereby allowing qualified professionals to devote a greater share of their time to the work that most depends on human expertise: complex judgment, therapeutic relationships, trauma-informed care, risk assessment, ethical discernment, intervention, supervision, and genuine human connection.

The demonstration would therefore evaluate a new division of labour between **computational capacity and professional human capacity**. AI-enabled infrastructure would be used to create greater availability, continuity, multilingual access, accessibility, repeatability, and scale. Human professionals would remain responsible for judgment, safety, governance, escalation, intervention, quality assurance, ethical oversight, culturally responsive review, and complex care. The purpose is not to diminish the role of professionals, but to protect it.

Government Decision at a Glance

Element	Proposed Approach
Policy Challenge	Canadians can experience periods in which no appropriate support is readily available before, between, or after professional encounters because of affordability, wait times, provider mismatch, limited operating hours, language barriers, geography, disability-related barriers, fragmented navigation, or the finite duration of existing programs.
Central Policy Insight	The challenge is not only the supply of mental-health professionals. It is also how scarce professional human attention is allocated. Some structured, repetitive, informational, navigational, and text-based functions may increasingly be suitable for responsible AI assistance, allowing professional attention to be concentrated on needs that require human expertise.
Immediate Recommendation	Establish a time-limited, staged, independently evaluated Human-Supervised Virtual Safe Space Demonstration Initiative, with evidence generation—not immediate national deployment or maximum user volume—as its first objective.
Public-Service Model	Provide free-at-the-point-of-appropriate-use, multilingual, trauma-informed, 24/7 AI-assisted emotional support, guided reflection, non-clinical coaching, life-skills support, routine psychoeducation, and service navigation within clearly defined professional and safety boundaries.

Element	Proposed Approach
Access Model	Test an appropriate foundational layer that can be available 24 hours a day, 365 days a year, through multilingual text and audio-enabled interaction, without a per-interaction out-of-pocket charge, subject to privacy, safety, technical-capacity, accessibility, and service-boundary requirements.
Role of AI	Provide availability, continuity, structured support, multilingual reach, repeatability, accessibility, and scale.
Role of Human Supervision	Provide governance, safety review, escalation, direct intervention when appropriate, service navigation, quality assurance, professional-boundary oversight, cultural review, ethical accountability, and referral.
Role of Professional Care	Preserve licensed and specialized human care for psychotherapy, assessment, treatment where authorized, therapeutic relationships, complex trauma, risk evaluation, professional judgment, and other functions requiring regulated or specialized expertise.
Funding Hypothesis	Test an approximate 50/50 allocation within a defined existing or comparable public funding envelope: approximately 50% for AI-enabled public digital infrastructure and approximately 50% for human professional capacity and governance. The ratio is a *demonstration hypothesis*, not a permanent formula or predetermined optimal allocation.
Economic Principle	*Separate the cost of availability from the cost of expertise*. Use computational capacity to create appropriate abundance in scalable functions while protecting scarce professional human attention for complex human needs.
Fiscal Objective	Determine whether a complete human-AI system can produce *greater access, continuity, safety, and public value from a comparable funding envelope*, rather than evaluating AI merely on the unit cost of an individual interaction.
Workforce Principle	*Job evolution, not job elimination.* The relevant measure is not how many professional tasks AI can replace, but how much professional attention can be made available for people whose needs genuinely require human expertise.
Safety Standard	AI must not misrepresent itself as a licensed professional, independently diagnose or prescribe without legal authority, replace crisis or emergency services, discourage appropriate professional care, commercialize sensitive disclosures, or become a mechanism for individualized surveillance.
Privacy Standard	Treat privacy as core infrastructure through data minimization, meaningful and understandable consent, cybersecurity, appropriate retention limits, role-based access, transparent rules for human review, strict limits on secondary use, and independent auditing.
Accessibility Standard	Build multilingual, disability-inclusive, culturally responsive, geographically inclusive, low-bandwidth, text- and audio-enabled access into the service from the outset, while maintaining appropriate human and non-AI alternatives for people who cannot or do not wish to use AI.
Service Architecture	**Layer 1:** Always-available AI-enabled foundational support. **Layer 2:** Qualified human supervision, intervention, safety, navigation, and governance. **Layer 3:** Licensed and specialized professional care.

Element	Proposed Approach
Evaluation Framework	Test nine domains: accessibility, continuity, professional capacity, workforce sustainability, safety, fiscal productivity, user trust and dignity, frontline learning, and legal and governance clarity.
Care Guarantee	No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.
Continuity Guarantee	No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.
Social-Innovation Principle	Social entrepreneurs can identify unmet needs, experiment, and demonstrate possibilities, but should not be expected to bear indefinitely the financial, operational, technological, ethical, regulatory, and legal burdens associated with essential public infrastructure.
Government Role	Define public rights, standards, professional boundaries, privacy requirements, accessibility obligations, human-supervision requirements, escalation pathways, evaluation criteria, accountability mechanisms, and appropriate long-term funding arrangements.
Technology-Neutrality Principle	Government should protect the promise, not the platform. The public capability should not depend indefinitely on one founder, nonprofit organization, vendor, AI model, or technology company.
Long-Term Objective	If the evidence supports the model, establish a durable, technology-neutral public capability for free, continuous, multilingual, trauma-informed, accessible, human-supervised support connected to appropriate professional care.
Public Promise	Professional care when professional care is needed; somewhere appropriate and safe to turn when it is not immediately available.
National Opportunity	Position Canada as a leader in human-centered, publicly accountable institutional design for AI-enabled social infrastructure, demonstrating how technological innovation can expand social protection while preserving professional human care.

The Proposal Architecture at a Glance

The proposed architecture begins with a recognition that mental-health access is shaped by more than the number of available appointments. Cost, waitlists, provider mismatch, operating hours, language, geography, disability, navigation challenges, and the finite duration of programs can all create periods during which a person has no appropriate place to turn. These gaps become particularly important because **professional human attention is inherently finite**.

1. The Public-Policy Problem

Canada can continue expanding professional capacity, and there are compelling reasons to improve access to appropriate professional care. However, no realistic system can make every qualified professional continuously available to every person at every hour. The policy challenge is therefore

twofold: Canada must improve access to professional expertise while also reducing the periods in which people are left entirely unsupported because that expertise is temporarily unavailable.

At the same time, frontline experience suggests that supportive-service systems contain a range of activities with very different requirements. Some depend heavily on professional expertise and human relationship. Others are more structured, standardized, informational, repetitive, navigational, or protocol-driven. Historically, human workers performed almost all of these functions because there was no technologically viable alternative. AI creates an opportunity to examine whether that assumption should continue to govern public-service architecture.

2. The Resource-Allocation Opportunity

The proposed demonstration would test whether appropriately bounded structured functions can be responsibly automated or augmented without undermining professional standards. If so, the principal benefit should not be understood simply as reducing labour costs. The more important benefit would be the **release of scarce professional attention**.

Professional time that is no longer required for suitable standardized or repetitive functions could potentially be redirected toward the situations in which human capability produces the greatest value: complex judgment, therapeutic relationships, trauma-informed care, difficult ethical decisions, risk assessment, intervention, supervision, ambiguity, and human connection.

The governing principle is therefore: **AI for availability and scale. Humans for judgment, safety, complexity, and connection.**

3. The 50/50 Demonstration Hypothesis

The initiative would test an approximate division of a defined existing or comparable funding envelope between two complementary categories of capacity.

- The first category—approximately half of the demonstration envelope—would purchase **AI-enabled public digital infrastructure**. This could include continuous availability, multilingual capability, text and audio-enabled access, secure computation, model and inference access, accessibility systems, privacy protections, cybersecurity, structured support functions, safety infrastructure, testing, and technical evaluation.
- The second category—approximately half—would purchase **professional human capacity and governance**. This could include professional judgment, therapeutic depth, complex-care capacity, safety review, escalation, direct intervention, supervision, service navigation, ethical oversight, quality assurance, cultural review, and institutional accountability.

The distinction is important. The proposed experiment is not based on the premise that AI is "cheap" and humans are "expensive." It asks whether public systems can **separate the cost of availability from the cost of expertise**, purchasing each through the form of capacity best suited to provide it.

The 50/50 allocation is intentionally provisional. Government should be free to conclude that the optimal ratio is different—or that the model should not proceed at all—based on evidence generated through the demonstration.

4. A Three-Layer Continuum of Support

The proposed service model would operate through three connected layers rather than presenting AI as a standalone substitute for care.

Layer 1: Always-Available Virtual Safe Space. This layer would provide appropriate first-line emotional support, guided reflection, non-clinical coaching, life-skills assistance, routine psychoeducation, structured exercises, service navigation, multilingual interaction, and text- and audio-enabled access. Its principal contribution would be **availability and continuity**.

Layer 2: Human-Supervised Support. Qualified humans would provide governance, safety review, professional-boundary oversight, escalation, intervention where appropriate, culturally responsive review, quality assurance, service navigation, and connections to professional, community, crisis, or emergency services. Its principal contribution would be **accountability and safety**.

Layer 3: Licensed and Specialized Professional Care. Qualified professionals would continue to provide assessment, psychotherapy, authorized treatment, therapeutic relationships, complex risk evaluation, trauma care, specialized intervention, and other services requiring regulated or specialized human expertise. Its principal contribution would be **depth and professional judgment**.

The architecture can therefore be summarized as follows:

- **AI creates availability.**
- **Humans create accountability.**
- **Professionals provide depth.**
- **Government provides continuity, standards, and public responsibility.**

5. The Intended Public Experience

From the user's perspective, the system should feel less like a collection of disconnected programs and more like a continuum. A person who recognizes a need could first access an immediate foundational layer rather than having to wait until a conventional service becomes available. That interaction could provide emotional support, structured reflection, basic guidance, or navigation.

Where a person's needs exceed the appropriate boundaries of the AI-enabled layer, qualified human involvement could be introduced. Where professional care is required, the individual could be connected or referred to the appropriate professional, community, crisis, or emergency pathway.

Importantly, access to the foundational layer would not necessarily disappear once professional care begins. It could remain available, where appropriate, while the individual waits for an appointment, between appointments, during the search for a better provider fit, or after a finite episode of structured service concludes.

The objective is continuity rather than substitution. **The program may end; the person's life does not.**

6. Evidence Before Scale

The demonstration should not assume that the proposed architecture will succeed. Its first purpose should be to generate credible evidence.

Evaluation should examine nine domains: **accessibility, continuity, professional capacity, workforce sustainability, safety, fiscal productivity, user trust and dignity, frontline learning, and legal and governance clarity**.

The central question should be whether the model creates meaningful improvements without introducing unacceptable risks. A credible demonstration must be designed so that government can legitimately conclude either that the evidence supports expansion or that it does not.

This principle—**evidence before national scale**—is essential to the initiative's legitimacy.

7. Two Complementary National Aspirations

The long-term public vision rests on two different forms of protection:

The **Care Guarantee** addresses financial scarcity: **No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.**

The **Continuity Guarantee** addresses availability scarcity: **No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.**

Neither guarantee replaces the other. Expanding digital continuity cannot substitute for making professional care more accessible. Equally, expanding professional funding does not eliminate every temporal, geographic, linguistic, relational, or navigational gap in availability. A more complete system should pursue both.

8. From Demonstration to Durable Public Capability

If independent evaluation supports the model, the long-term objective should not be to institutionalize one specific application, founder, nonprofit, commercial vendor, or AI model. Technology will change. Providers will change. Models will improve, merge, or disappear. Public policy should therefore be designed around the **capability that Canadians require**, not around the current technology used to deliver it.

That capability would be a free-at-the-point-of-appropriate-use, continuously available, multilingual, accessible, trauma-informed, privacy-preserving, human-supervised foundational safety net with reliable connections to appropriate professional care. In this sense, **government should protect the promise, not the platform**.

9. Four Dimensions of Public Value

A successful model could generate public value across four constituencies simultaneously.

- For Canadians, it could mean greater access, continuity, multilingual availability, and fewer periods of complete absence of support.
- For helping professionals, it could mean protection of scarce human attention, greater capacity for complex care, and a professional role increasingly concentrated on the aspects of helping work in which human beings remain indispensable.
- For government and taxpayers, it could mean stronger fiscal productivity, better use of professional hours, improved system learning, and the possibility of producing greater access and continuity from a comparable funding envelope.
- For social entrepreneurs, it could mean greater freedom to identify needs and demonstrate innovations without becoming indefinitely responsible for carrying the legal, financial, technological, ethical, and operational burden of essential social infrastructure.

Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.

10. The Canadian Public Promise

At its most mature, the proposed architecture would support a simple but consequential public promise: **Professional care when professional care is needed—and somewhere appropriate and safe to turn while it is not immediately available.**

This does not mean government can guarantee the continuous availability of every professional service. It means that finite professional capacity need not translate into complete absence of support.

If Canada can establish that capability safely, equitably, and sustainably, the resulting public experience could ultimately be expressed in more human terms: **“My country has my back.”**

11. An International Leadership Opportunity

The strongest international opportunity is not for Canada to claim that it possesses the most advanced AI. Such claims would be transient in a rapidly changing technological environment. The more durable opportunity is for Canada to demonstrate **how a democratic public system can govern AI-enabled social infrastructure responsibly**.

A successful Canadian model could show that technological progress need not require abandoning human care; that productivity need not mean dehumanization; that innovation need not require deregulation; and that continuity can be expanded without pretending that professional capacity is infinite.

The international lesson should not be that Canada replaced therapists with artificial intelligence. It should be that Canada found a way to preserve professional human care while using responsible

technology and strategic public investment to ensure that people were no longer left entirely unsupported when that care was temporarily unavailable.

The Proposal in One Sentence

Canada should test whether strategically reallocating finite public resources between scalable AI-enabled availability and protected professional human expertise can create a free, continuous, multilingual, accessible, human-supervised mental-health safety net—while strengthening, rather than replacing, professional care.

1. The Case for Action

The Policy Problem: Access Is More Than the Number of Professionals

Canada Faces an Access-Architecture Problem

A person may recognize that something is wrong long before they understand what kind of help they need or where they should begin. They may be uncertain whether the appropriate next step is a psychologist, psychotherapist, counsellor, social worker, physician, peer-support service, community organization, coach, crisis resource, or simply a psychologically safer place to speak, reflect, and determine what kind of support would be appropriate.

Even when a person is ready and willing to seek help, the existence of services does not necessarily translate into meaningful access. The appropriate provider may not be available when the need arises. A person may struggle to find someone who speaks their preferred language, understands their cultural context, has the necessary specialization, provides accessible services, or feels like the right relational fit. They may already have a therapist yet need appropriate support between appointments. They may complete a short-term counselling or structured program while the circumstances that brought them to the service continue. They may receive a valid referral yet still wait weeks or months before the next service becomes available.

These experiences demonstrate that mental-health access is not a single doorway. It is a **continuum shaped by affordability, timing, provider availability, provider fit, continuity, navigation, geography, language, disability, culture, trust, and a person's willingness and emotional capacity to begin again after an unsuccessful attempt to obtain help**.

That distinction is important because a service can be formally available while remaining practically inaccessible. A system can increase the number of appointments while still leaving people unsupported between them. A program can satisfy its administrative mandate while providing no continuity after the intervention ends. A practitioner can be highly qualified while not being the right clinical, cultural, linguistic, or relational fit for a particular individual. A person can possess a referral to an appropriate service and still have nowhere appropriate to turn tonight.

Canada's policy challenge is therefore broader than the absolute number of mental-health professionals. It also concerns the architecture through which professional capacity is accessed, allocated, connected, and surrounded by appropriate support at the moments when people actually need it.

If the problem is understood only as a shortage of practitioners, the natural policy response is to increase the supply of professional labour. In many areas, expanding professional capacity may indeed be necessary. But workforce expansion alone cannot resolve every temporal, geographic, linguistic, accessibility, relational, and navigational gap in care. Nor can professional human capacity realistically become unlimited.

Canada should therefore pursue two objectives at the same time: (a) expand access to appropriate professional care where additional capacity is required, and (b) examine whether the broader system can

make better use of the professional capacity already available while maintaining appropriate continuity when that capacity is temporarily unavailable.

The policy question is not simply how many professionals Canada has. It is also **how effectively professional expertise is connected to the people who genuinely require it—and what surrounds that expertise before, between, and after professional encounters**.

Universal Coverage Is Not the Same as Universal Availability

In 2021, I initiated a public petition calling on the Government of Ontario to expand OHIP coverage to include mental-health services delivered by mental-health professionals. The principle underlying that advocacy remains central to this proposal:

A person's access to appropriate professional mental-health care should not depend primarily on their ability to pay.

Financial barriers can delay the decision to seek care, interrupt an established course of support, limit a person's ability to try another provider after a poor match, and create significant inequality in the process of finding an appropriate professional. Expanding public access to necessary professional mental-health services should therefore remain an important policy objective.

Frontline experience, however, has also made clear that **affordability is only one dimension of accessibility**.

Even in a hypothetical system in which every appropriate professional service were fully publicly covered, significant access problems would remain. What happens while someone is waiting for an appointment? What happens at night or on a weekend? What happens when the first provider is not the right fit? What happens when the individual requires support in a language that is not readily available? What happens in rural, remote, or underserved communities? What happens when disability-related accessibility needs are not adequately accommodated? What happens when a short-term program ends while the person's circumstances continue? And what happens when an individual recognizes a need for support but does not yet understand what type of professional or service would be appropriate?

Public coverage can substantially reduce financial scarcity. It cannot eliminate availability scarcity, because professional human attention remains finite. Canada should therefore distinguish between two complementary public-policy aspirations.

The Care Guarantee

No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.

The Care Guarantee concerns access to professional expertise. It recognizes that when assessment, treatment, therapeutic care, complex judgment, or specialized intervention is required, a person's financial circumstances should not be the principal reason that appropriate care remains beyond reach.

The Continuity Guarantee

No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.

The Continuity Guarantee addresses a different problem. It recognizes that even a well-resourced professional system will contain periods during which the appropriate professional is not immediately available. During those periods, people should not necessarily move from receiving professional care to receiving nothing at all.

The two guarantees are mutually reinforcing rather than competing. The first asks: **Can I obtain the professional care I need?** The second asks: **Do I have somewhere appropriate and safe to turn right now?**

A mature mental-health support architecture should aspire to answer **yes** to both. This distinction is central to the proposal. The long-term objective is not to substitute continuous AI interaction for appropriate professional care. It is to build a system in which access to human expertise is strengthened while an appropriate foundational layer of support remains available around it.

Professional capacity will remain finite. The safety net does not have to disappear when professional capacity is temporarily unavailable.

The Scarce Resource Is Not Only Money—It Is Professional Human Attention

Therapists, counsellors, psychologists, social workers, physicians, and other helping professionals possess capabilities that cannot be reduced to scripts, standardized information, or procedural exchanges. They exercise judgment when circumstances are ambiguous. They establish therapeutic and helping relationships. They interpret context, silence, power, trauma, culture, behaviour, and relational dynamics. They assess risk, navigate ethical complexity, recognize when a person does not fit a predefined pathway, and determine when a situation requires a different form of intervention.

Most importantly, they provide something that technology cannot fully reproduce: **accountable human presence**. Professional attention is valuable precisely because it is finite. A therapist's hour can only be spent once. A counsellor cannot use the same hour both for a routine structured interaction and for a person experiencing complex trauma. A social worker cannot simultaneously devote professional attention to repetitive service navigation and to a difficult systems intervention involving significant judgment and advocacy. A supervisor cannot allocate the same professional capacity both to standardized administrative functions and to complex case consultation or risk oversight.

This creates an opportunity cost that deserves greater attention in mental-health policy.

Many structured supportive-service environments contain a combination of activities. Some require substantial professional expertise; others may involve structured intake, routine navigation, standardized psychoeducation, basic goal-setting, repeatable exercises, documentation, routine follow-up, information delivery, or suitable lower-complexity supportive interactions. These functions can be useful

and important. The relevant question is not whether they have value but **whether every instance of every such function must continue to consume scarce professional human attention**.

Historically, the answer was largely predetermined by technological constraint. Human beings had to perform almost every supportive interaction because there was no sufficiently capable alternative. Contemporary artificial intelligence changes that constraint for some appropriately bounded functions. Structured, repetitive, informational, protocol-guided, navigational, and text-based activities can increasingly be assisted by computational systems, subject to appropriate safeguards, professional boundaries, privacy protections, accessibility standards, and human oversight.

This development should not be interpreted as an argument for automating professional care. It is an argument for examining more carefully which functions truly require professional expertise, which can be responsibly augmented, and how the resulting capacity should be reinvested in human care.

The central workforce question is therefore not: Do professionals still matter? They unquestionably do. The more consequential question is: **Is Canada consistently protecting scarce professional human attention for the work that most requires judgment, safety, complexity, therapeutic relationships, ethical discernment, intervention, and genuine human presence?**

If professional shortages are real, then the allocation of professional attention becomes more—not less—important. Canada may simultaneously need more professionals and a better architecture for using the professionals it already has.

This is the basis for the proposal's central workforce principle: **job evolution, not job elimination**. The success of a responsible AI-enabled model should not be measured primarily by how many professional tasks are removed from human workers. It should be measured by whether professionals gain greater capacity to spend their time on people and situations that genuinely require human expertise.

In this sense, the proposal seeks to elevate rather than diminish professional work. **Human professional attention has become too valuable to use inefficiently.**

Why This Belongs in a Pre-Budget Consultation

Although this proposal concerns mental-health accessibility, its implications extend considerably beyond the delivery of mental-health services. It is equally a proposal about public-sector productivity, affordability, workforce transformation, digital public infrastructure, responsible artificial intelligence, service modernization, privacy and cybersecurity, accessibility, legal and regulatory clarity, social innovation, and the allocation of finite public resources. For that reason, the proposal belongs squarely within a federal pre-budget discussion.

The relevant fiscal question is not simply: **How much more should government spend?** It is also: **How can government obtain greater access, continuity, safety, equity, and public value from the resources it is already committing and may commit in the future?**

Historically, the economics of personalized supportive services were closely linked to human labour. Serving more people generally required more professional hours. Extending operating hours required additional staffing. Increasing language coverage required additional human capacity. Expanding geographic reach frequently required parallel organizational and workforce expansion. In other words, greater availability generally required a corresponding increase in labour.

Artificial intelligence changes that economic relationship for some suitable and carefully bounded functions. For the first time, governments can meaningfully test whether two categories of expenditure that historically had to be purchased together can be partially separated:

- **the cost of making appropriate support available;** and
- **the cost of providing professional expertise, judgment, therapeutic depth, and human care.**

This distinction is economically important.

Continuous availability can increasingly be purchased in part through computational resources: model access, token or inference capacity, secure infrastructure, multilingual capability, accessibility systems, cybersecurity, safety mechanisms, testing, maintenance, and responsible human oversight. Professional expertise, by contrast, remains fundamentally dependent on scarce human capacity.

The public-finance opportunity is therefore not to replace one category of expenditure with the other. It is to determine whether government can purchase availability and expertise differently, while integrating them into a safer and more effective whole.

Could part of a defined public funding envelope support scalable digital availability, while another part protects professional human capacity for complex care, judgment, therapeutic relationships, safety, intervention, supervision, and accountability? Could such an architecture serve more people without compromising professional standards? Could it reduce periods of complete absence of support while increasing the proportion of professional time devoted to needs that genuinely require professionals? And could it do so within a comparable overall funding envelope?

These are empirical questions. They should be tested rather than assumed. That is why this proposal should not be characterized as a conventional cost-reduction initiative. Its purpose is not to determine how cheaply artificial intelligence can generate an interaction but to examine whether a complete human-AI system can create **greater access, greater continuity, greater professional capacity, and greater public value from finite resources**.

The central fiscal proposition is therefore resource reallocation, not simply resource expansion. It begins not from the premise that human beings have become less valuable, but from the opposite conclusion: **professional human attention has become too valuable to allocate casually.**

A pre-budget consultation is particularly appropriate for this question because a proposal of this significance should not proceed directly from concept to nationwide implementation. Responsible public innovation requires institutional capacity to co-design, test, evaluate, revise, compare alternatives, identify failure modes, establish professional boundaries, assess workforce effects, develop privacy and governance standards, determine the appropriate role of human oversight, and measure whether the resulting public value warrants longer-term investment.

The immediate request is therefore deliberately limited. It is not that Canada commit now to a permanent national AI-enabled mental-health system. Rather, it is that government recognize the significance of the resource-allocation question and create the fiscal and institutional space necessary to test it responsibly through a controlled, time-limited, independently evaluated demonstration.

The first objective should be evidence generation, not maximum user volume or premature national scale. A credible demonstration should be designed so that government can reach either conclusion. If the evidence indicates that the model improves access, continuity, safety, professional-capacity utilization, workforce sustainability, user dignity, and fiscal productivity, government would have a basis for considering further development. If the evidence does not justify expansion, the model should not advance in its proposed form. That neutrality strengthens, rather than weakens, the case for testing.

If successful, however, the implications could be significant. Canada could move toward a system in which professional mental-health care remains deeply human while an appropriate foundational safety net remains continuously available around it—before appointments, between them, while people search for appropriate care, and after finite programs end.

Such a public capability cannot be built responsibly through short-term technological enthusiasm or institutional improvisation. It requires long-horizon thinking, rigorous evaluation, professional partnership, privacy and safety governance, fiscal discipline, and clarity about public responsibility.

The underlying principle is straightforward: **When technology changes the scarcity assumptions on which public systems were designed, government should examine whether those systems can be reorganized to serve people better.**

This proposal asks Canada to begin that examination—not by assuming that artificial intelligence is the answer, and not by deploying an untested model nationally, but through a disciplined demonstration designed to determine whether a more continuous, equitable, professionally protective, and fiscally sustainable architecture of support is possible.

The objective is ultimately national in significance: **To determine whether Canada can build a mental-health support architecture in which appropriate professional care is protected, scarce human expertise is used where it creates the greatest value, public resources are allocated more intelligently, and no Canadian is left with absolutely nowhere appropriate and safe to turn simply because the formal system is temporarily unavailable.**

2. From Frontline Experience to Public Policy

From Frontline Experience to a New Public-Infrastructure Question

Frontline Experience Reveals What Aggregate Data Cannot

Governments possess substantial knowledge about mental-health systems. Administrative data, population surveys, academic research, program evaluations, performance indicators, economic analysis, professional expertise, and public-service experience are all indispensable to responsible policymaking. They reveal patterns at a scale that no individual practitioner, organization, or service provider can observe alone.

Frontline experience contributes a different form of knowledge. It reveals how policies and programs are actually encountered by people in moments of need.

A program may appear accessible in administrative terms while remaining difficult to use in practice. An increase in funded appointments may expand nominal capacity without resolving provider mismatch, language barriers, geographic constraints, limited operating hours, or discontinuity between services. A short-term intervention may appropriately address an immediate concern yet leave unanswered what happens after the final funded session. A highly trained professional may spend part of the day delivering standardized or repetitive functions while another individual whose circumstances require complex judgment, relational depth, or sustained human presence remains on a waitlist.

These observations should not be interpreted automatically as failures of a particular professional, organization, contractor, or program. In many cases, they reflect the rational operation of systems designed around finite resources and finite human capacity. The more important question is whether the architecture itself should evolve as technological possibilities change.

Aggregate data can help government understand **where resources are being allocated, how services are being used, and what outcomes are being reported**. Frontline experience can help illuminate an additional question: **whether scarce professional human attention is consistently being used where it creates the greatest public value**.

That distinction is central to this proposal. The argument that follows is not derived from the assumption that existing services are ineffective. It arises from the possibility that many valuable services were designed under technological constraints that are now beginning to change.

My Original Assumption Changed

Before developing For A Safer Space (FASS), and later FASSLING™, I began with an assumption that many people understandably share: that the principal cause of inadequate mental-health access was simply an insufficient number of professionals. Frontline experience complicated that assumption.

I repeatedly encountered people whose difficulties could not be explained solely by the absolute supply of practitioners. Some described spending considerable amounts of money trying different therapists before finding an appropriate clinical or relational fit—or abandoning the search before finding one. Others waited weeks or months for care. Some required support outside conventional operating hours. Some struggled to obtain services in a preferred language or within an appropriate cultural context. Others were not yet ready for formal psychotherapy or did not understand what kind of assistance they required.

Some already had a therapist and nevertheless experienced periods between appointments when additional appropriate support would have been valuable. Others were navigating unclear distinctions among psychotherapy, counselling, social work, coaching, peer support, community services, medical care, crisis services, and other forms of assistance.

These experiences led me to reframe the problem. The relevant question was not only: Are there enough professionals? It was also: **Can people reach the right form of support at the right time—and what happens to them while they are trying?**

Mental-health accessibility is therefore not simply a workforce-supply issue. It is also an **access-architecture problem** involving how people enter the system, how they navigate between different forms of support, how scarce professional attention is allocated, and what remains available when professional care is delayed, interrupted, unsuitable, or temporarily unavailable. That realization became the starting point for the subsequent development of FASS and FASSLING™.

FASS: Testing the Human Bridge

My first practical response to this problem was not technological. It was human. I founded For A Safer Space (FASS) around a simple but ambitious principle: **A person seeking emotional or mental-health support should not be left entirely without support simply because they cannot afford another session, have reached the limit of a short-term program, are waiting for formal care, or have not yet found an appropriate helping professional**.

FASS was established as what I believe was the world's first organization designed to provide unlimited free mental-health counselling, psychotherapy, emotional support, and coaching sessions to help seekers globally. Because this is a potentially verifiable superlative, the claim should remain carefully qualified unless independent evidence is available to substantiate a stronger formulation.

The distinctive feature of the model was continuity. Rather than organizing access around a predetermined number of sessions, FASS was designed so that eligible help seekers could return when appropriate support was needed without first having to calculate whether they could afford another interaction or whether they had exhausted a fixed allocation of sessions.

Through a model involving clinical interns and trainees from universities in different parts of the world, appropriate supervision and encouragement to use licensed or specialized professional services where necessary, FASS sought to address one of the most difficult periods in the mental-health journey: the space between "I need help" and "I have found the right professional care."

That interval can be lengthy and uncertain. A person may contact multiple providers before finding someone with whom they feel sufficiently safe. They may encounter cultural or linguistic mismatch. A particular therapeutic modality may not suit their needs. The provider may lack availability. The person may spend time on a waitlist. They may already be receiving professional care while still experiencing difficult periods between appointments. Or they may not yet know whether their circumstances call for psychotherapy, counselling, social work, coaching, psychiatric care, community support, or another form of assistance.

FASS was designed so that people did not have to navigate that period entirely alone. Importantly, the FASS model was never premised on asking people to choose between FASS and appropriate professional care. Help seekers were encouraged to use FASS alongside an appropriate therapist or other licensed helping professional whenever such care was available and suitable to their needs. Its role was complementary rather than substitutive.

The objective was not to draw people away from formal mental-health services. It was to provide an accessible bridge while they entered those services, navigated them, received care within them, or searched for a more suitable professional fit. Within appropriate service boundaries, a person could begin speaking rather than remaining entirely unsupported. They could organize their thoughts, gain clarity, receive emotional support, obtain basic navigation assistance, prepare themselves to engage more effectively with professional care, and be encouraged toward licensed or specialized services where their circumstances required them. If one professional relationship did not work out, they could retain a supportive place to return while continuing the search.

The broader lesson was significant: **Support should not exist only inside isolated professional encounters.** Human need does not organize itself according to appointment calendars, program limits, or service contracts. Need may arise before the first appointment, between appointments, during a provider transition, after an unsuccessful match, after a short-term intervention ends, or at a moment when the appropriate professional is simply unavailable.

FASS therefore provided an early practical test of an accessible human bridge before, between, and alongside professional care.

What FASS Revealed About the Limits of Human-Only Delivery

The FASS experience also revealed an unavoidable constraint: **Human capacity is finite.** A counsellor has only so many working hours in a day. A clinical intern can support only a finite number of people. A supervisor can safely oversee only a finite amount of work. Expanding into another language requires additional linguistic capability. Extending coverage across time zones requires new scheduling capacity. Continuous 24-hour availability requires shifts. Every meaningful increase in demand requires some combination of recruitment, screening, training, supervision, coordination, administration, infrastructure, and quality assurance.

This remains true even when the person receiving support pays nothing. Removing the price charged to the user does not eliminate the cost of producing the service. It changes who absorbs that cost. That distinction became fundamental to the evolution of the model. Free human-delivered support can reduce financial barriers for users, but it does not make human labour abundant. A system can eliminate a

per-session fee without eliminating the scarcity of counsellors, supervisors, operating hours, languages, administrative capacity, or professional attention.

FASS therefore revealed two important propositions at the same time:

- **First, people can benefit from an accessible bridge before, between, and alongside professional care**.
- **Second, a human-only bridge cannot realistically provide unlimited, continuously available support at population scale without continuously expanding human resources**.

This did not weaken the original mission. It clarified the next design question: **How can the safety-net principle—free access, continuity, emotional support, appropriate human connection, and encouragement toward professional care—be preserved while reducing its dependence on an ever-expanding supply of human availability?** That question ultimately led to FASSLING™.

FASSLING™: Exploring Whether the Safety Net Can Become Continuously Available

FASSLING™ emerged from a practical realization: **The need for support can be continuous even when human availability cannot be.**

The underlying public-interest objective remained substantially the same. The delivery architecture changed.

FASSLING™ was developed to explore whether people could have an immediately available, non-judgmental place to turn for appropriate non-medical emotional support, guided reflection, non-clinical coaching, navigation, and holistic life-skills support regardless of conventional appointment schedules, geography, language, or the immediate availability of a human helping professional.

Artificial intelligence introduces a fundamentally different form of service capacity. An appropriately designed AI-enabled system can operate across time zones, provide text-based and audio-enabled interaction, function in multiple languages, respond without conventional appointment scheduling, and support additional interactions without requiring a corresponding additional professional to become personally available for every use. That distinction is economically and operationally significant.

FASSLING™ was not designed to replace counselling, psychotherapy, psychology, psychiatry, social work, crisis services, emergency services, or other regulated or specialized professional care. Its intended role is **non-medical, non-diagnostic, non-therapy emotional support, guided reflection, life coaching, navigation, and life-skills assistance**, operating within clear boundaries and complementing rather than competing with professional services.

The proposition is therefore not that AI is sufficient for every person or every need. It is much narrower: **The temporary absence of immediately available professional care does not necessarily have to mean the complete absence of appropriate support.**

A person waiting for a therapist may still benefit from a place to reflect and organize their thoughts. Someone between appointments may still need an appropriate supportive interaction. A person whose first provider was not the right fit need not necessarily return to total isolation while searching again. Someone who requires support outside ordinary service hours can have a place to begin. A person uncertain about what they are experiencing can potentially obtain basic structured reflection and navigation before approaching the next appropriate service. In this progression, FASS demonstrated the need for the safety net; FASSLING™ explored whether technology could make an appropriate part of that safety net continuously available.

The policy significance does not depend on whether FASSLING™ itself becomes a permanent public platform. Its significance lies in the institutional question the experience raises: whether the underlying capability can be governed, evaluated, and, if justified by evidence, developed as durable public infrastructure.

Beyond Emotional Support: Building Holistic Capacity

A further lesson emerged from the development of this work: Psychological safety and well-being cannot be supported through emotional conversation alone. People also rely on practical and internal capacities that shape how they navigate ordinary life. These can include communication, boundaries, self-awareness, emotional regulation, decision-making, stress management, relationships, resilience, goal-setting, conflict, purpose, and other everyday life skills.

For this reason, I developed a broader ecosystem of FASSLING™ programs oriented toward holistic essential life skills. The intention was to allow people to engage with structured material addressing particular capacities while also having access to emotional or coaching-oriented support when thoughts, frustrations, uncertainties, or difficult feelings arose during that process.

The broader concept is therefore not simply a conversational AI interface. It is an integrated support environment based on a wider public-interest proposition: **Support should be available not only after a person reaches a crisis threshold, but also throughout the ordinary processes of living, learning, adapting, reflecting, and growing.**

This observation reinforces the importance of differentiated service design. Not every human difficulty constitutes a clinical condition. Not every supportive interaction requires psychotherapy. Not every life challenge requires regulated professional expertise. Conversely, some circumstances unquestionably do require licensed, specialized, or emergency intervention.

A mature support architecture should be capable of distinguishing among those levels of need rather than collapsing them into a single service category. The purpose of a foundational support layer is therefore not to blur the boundaries between emotional support, coaching, counselling, psychotherapy, crisis intervention, and medical care. It is to make those boundaries more explicit and to help ensure that each form of capacity is used appropriately.

Experience Inside a Government-Funded Program Revealed a Broader Public-Finance Question

My perspective is also informed by previous frontline work as a Government Services Counsellor and Integrated Health Solution Specialist within a government-funded program operated by Morneau Shepell and LifeWorks, now TELUS Health. The precise employer, program, title, funding relationship, and corporate-successor wording should be verified before formal submission so that this firsthand experience is described with complete institutional accuracy.

That experience provided exposure to an important feature of the broader mental-health landscape: Governments already commit significant resources to mental-health and supportive-service delivery through a range of mechanisms and institutional arrangements. These can include structured counselling services, digital mental-health programs, Employee Assistance Programs, public-sector supports, community-based services, contracted delivery models, and other forms of funded intervention. Depending on the jurisdiction and program, public resources may flow through governments, healthcare institutions, nonprofit organizations, community providers, or private-sector contractors.

The relevant policy question is therefore not simply: Why is government not funding mental health? Governments already invest in mental-health services and supportive infrastructure. The more consequential question is: **Can those resources be structured in ways that generate still greater access, continuity, safety, professional quality, and public value?**

Structured counselling programs, i-CBT interventions, Employee Assistance Programs, and similar services can provide meaningful benefits. They can offer people a point of entry, support an immediate concern, provide coping strategies, structure reflection, facilitate navigation or referral, and deliver appropriate short-term assistance.

The argument of this proposal is not that these programs lack value. It is that their architecture necessarily reflects a fundamental economic constraint: **Professional human time is finite.**

Human-delivered programs therefore manage scarcity through mechanisms such as defined scopes, eligibility criteria, session limits, standardized pathways, operating hours, structured protocols, waitlists, and referral once the boundaries of the funded service have been reached. Those constraints are not inherently evidence of poor program design. They are often rational responses to the economics and capacity limitations of human-delivered services.

The deeper policy question is: **What happens when the person's need does not end when the program ends?** Grief does not necessarily resolve within a predetermined number of sessions. Trauma may require extended processing. Anxiety can recur. Loneliness may persist. Relationships evolve. Family circumstances change. Employment pressures return. Someone may resolve one difficulty and encounter a different one months later.

The program may end. The person's life does not.

This creates a continuity problem that cannot necessarily be solved simply by funding additional finite episodes of human-delivered service. It raises the possibility that Canada requires a complementary layer of support with a different economic architecture.

A Frontline Observation That Deserves Testing: Not Every Supportive Function Requires the Same Level of Professional Labour

One of the most consequential observations informing this proposal emerged from the internal structure of supportive-service delivery itself. Many structured mental-health and supportive programs contain activities that may be standardized, procedural, repetitive, protocol-guided, informational, or text-based. Depending on the program and the needs of the individual, examples can include structured intake questions, routine psychoeducation, guided reflection, standardized CBT-informed exercises, basic goal-setting activities, journaling prompts, coping-skills practice, routine follow-up, progress tracking, emotional check-ins, service navigation, resource information, structured homework, and text-based supportive communication.

These activities can be useful. Some may require professional involvement depending on context, risk, service scope, and the individual's circumstances. The relevant proposition therefore is not that they should automatically be removed from human professionals. Rather, the demonstration should test whether **some appropriately bounded instances of these functions can be responsibly AI-assisted or automated without reducing safety, service quality, user dignity, or access to professional care**.

Historically, human workers had to deliver much of this work because no system could sustain responsive natural-language interaction, explain information dynamically, provide adaptive prompts, support multiple languages at scale, or remain continuously available on demand. That technological constraint is changing. Contemporary AI systems can increasingly perform certain structured, repetitive, informational, navigational, and language-based functions through computational capacity. This creates the possibility of examining the internal composition of supportive services more precisely.

Therefore, the appropriate policy question is not: Can the entire service be automated? It is: **Which functions require professional human expertise from the outset; which can be responsibly augmented by AI; which may be suitable for automation within clear safeguards; and when must responsibility return immediately to a qualified human professional?**

That is a more credible framework for human-AI service design because it begins with functional differentiation rather than replacement.

The Opportunity Cost of Professional Human Attention

This observation leads to a concept that is central to the economic case for the demonstration: **the opportunity cost of professional human attention**.

A therapist, counsellor, psychologist, social worker, physician, or other qualified helping professional represents years of education, supervision, accumulated judgment, professional development, ethical responsibility, and relational experience. Their attention, therefore, is not an interchangeable input. It is a scarce form of human capital.

A professional hour can only be spent once. If a professional spends part of that hour facilitating a highly standardized activity that an appropriately governed system could safely support, the relevant cost is not only the salary or contracted rate associated with that time. There may also be an opportunity cost: **Another person whose circumstances genuinely require professional judgment, trauma expertise, relational depth, risk assessment, ethical reasoning, or direct intervention may remain waiting**.

This distinction changes how workforce scarcity should be understood. A shortage of mental-health professionals may indeed mean that Canada needs more qualified practitioners, and this proposal should not be interpreted as an alternative to workforce expansion where additional professionals are required.

However, a second problem may coexist with the first: Existing professional capacity may be used for a mixture of functions that require very different levels of expertise. These two challenges call for different policy responses. The first may require **workforce expansion**. The second may require **service redesign**. Responsible artificial intelligence makes the second possibility increasingly practical.

The objective is not to reduce the value of professional work. It is to recognize that professional human attention is valuable precisely because certain aspects of human care cannot be replicated through computation. Accordingly, the central workforce metric for the demonstration should not be: How many professional tasks did AI replace? It should be: **How much additional professional attention became available for people whose needs genuinely required human expertise?**

That is the meaning of **job evolution, not job elimination**.

From Professional Labour to a Different Form of Capacity

Artificial intelligence also changes the economics of service availability. Under conventional human-delivered service models, greater availability generally requires greater labour capacity. More users require additional staff time. More interactions require additional paid hours. Longer operating hours require additional shifts. Additional languages require additional linguistic capacity. Wider geographic coverage usually requires corresponding organizational expansion.

AI-enabled infrastructure does not eliminate costs. Instead, it introduces a different category of cost. Appropriate structured functions can potentially be supported through expenditures on model access, token or inference usage, computational capacity, secure hosting, multilingual infrastructure, text and audio-enabled interaction, privacy protections, cybersecurity, accessibility, automated safety mechanisms, testing, maintenance, monitoring, evaluation, and qualified human oversight.

Government would still pay for capacity. But it would be purchasing **a different form of capacity with a different scaling relationship**.

The significance is therefore not merely that an individual AI interaction might cost less than an individual professional interaction. Such comparisons can be misleading because responsible public deployment requires governance, privacy, cybersecurity, human supervision, safety systems, evaluation, and infrastructure.

The more important possibility is that AI may weaken the historical relationship between **one additional appropriate supportive interaction** and **one additional equivalent increment of professional labour**. That creates the economic foundation for a different public-service architecture.

Government may now be able to distinguish more deliberately between **the cost of making an appropriate foundational support layer available** and **the cost of providing professional expertise, therapeutic depth, judgment, safety, and human care.**

This is the distinction between the **cost of availability** and the **cost of expertise**. The two remain connected. They should not be treated as substitutes. But they no longer necessarily need to be purchased through the same delivery mechanism. That possibility provides the intellectual basis for the 50/50 demonstration hypothesis developed later in this proposal.

The Policy Journey: From Listening to Infrastructure

The development of this proposal can therefore be understood as a sequence of increasingly institutional questions.

- **First, public-coverage advocacy identified the affordability problem.** I began from the principle that appropriate professional mental-health care should not be denied primarily because an individual is unable to pay.
- **Second, FASS tested free human support as an accessible bridge.** It explored the value of providing continuity before, between, and alongside professional care rather than limiting support to isolated formal encounters.
- **Third, FASS revealed the capacity limits of human-only delivery.** Free access could remove a financial barrier for the person using the service, but it could not eliminate the underlying scarcity of human labour, supervision, languages, operating hours, and professional attention.
- **Fourth, FASSLING™ explored whether technology could extend availability.** Artificial intelligence created a potential pathway for making appropriate non-medical emotional support, guided reflection, coaching, navigation, and life-skills assistance available across time zones and languages without requiring an additional professional interaction for every additional use.
- **Fifth, frontline experience within government-funded service delivery raised a broader public-finance question.** Governments already invest in mental-health and supportive services, but human-delivered programs necessarily operate within the economics of finite professional time.
- **Sixth, frontline observation suggested that not every supportive function necessarily requires the same level of professional labour.** Structured, repetitive, informational, navigational, and text-based functions may increasingly be candidates for carefully governed AI assistance, while complex judgment and professional care remain human responsibilities.
- **Seventh, AI creates the possibility of distinguishing availability from expertise.** Computational infrastructure may be able to create greater continuity and scale while scarce professional attention is protected for complex care, relationships, safety, judgment, intervention, and oversight.

The next question is no longer primarily entrepreneurial or technological. It is institutional: **Can government combine AI-enabled availability, qualified human supervision, professional care, strong privacy and safety safeguards, public accountability, and sustainable financing into durable social infrastructure?**

This progression reflects how responsible social innovation should develop. It should begin by listening to people experiencing the problem. It should remain close to frontline reality. It should test assumptions rather than defend them. It should identify where an existing approach succeeds and acknowledge where that approach encounters structural limits. It should redesign in response to those limits. It should generate evidence. And when an innovation appears capable of addressing a persistent public need, government should determine whether the underlying capability merits rigorous public evaluation.

The innovation did not begin with AI. It began with listening. Artificial intelligence became relevant only after frontline experience revealed a problem that goodwill, additional staffing, and conventional service structures alone could not fully resolve: How to maintain an appropriate foundational layer of support without assuming an unlimited supply of professional human attention.

FASS and FASSLING™ should therefore be understood not as the proposed national solution, but as part of the evidentiary and experiential pathway that produced the public-policy question now before government. The next stage should not depend on one founder, one nonprofit organization, one platform, one vendor, or one AI model. It should ask whether the underlying public capability deserves to be tested under professional governance, independent evaluation, fiscal discipline, and democratic accountability. That is the transition from **social innovation** to **public-infrastructure policy**—and it is the question to which the remainder of this proposal responds.

3. The Proposed Initiative

The Proposed Demonstration: Building a Human-Supervised Continuity Layer Around Professional Care

Recommendation for Proposed Human-Supervised Virtual Safe Space Demonstration Initiative

I respectfully recommend that the Government of Canada establish a **time-limited, staged, independently evaluated Human-Supervised Virtual Safe Space Demonstration Initiative** to test a new model of publicly supported mental-health and emotional-support infrastructure. The demonstration would examine whether Canada can provide an appropriate first layer of **free-at-the-point-of-appropriate-use, multilingual, trauma-informed, 24/7 emotional support, guided reflection, non-clinical coaching, life-skills support, routine psychoeducation, and service navigation**, accessible through text and audio-enabled interaction, while preserving clear and enforceable boundaries around psychotherapy, regulated clinical care, crisis intervention, emergency services, diagnosis, prescribing, and other functions requiring professional or legal authority.

The proposed model would combine two complementary forms of capacity. The first would be **AI-enabled public digital infrastructure**, designed to create availability, continuity, multilingual reach, accessibility, structured support, and scale. The second would be **qualified human professional capacity and governance**, responsible for judgment, supervision, safety, complex needs, escalation, intervention, quality assurance, ethical oversight, cultural review, professional boundaries, and accountability.

The premise is not that one form of capacity can replace the other. It is that they purchase different forms of public value and may therefore be used more deliberately together.

The demonstration should also test whether a defined existing or comparable public funding envelope can be allocated differently between these two categories of capacity, rather than assuming that each additional appropriate supportive interaction must require an equivalent additional increment of professional human labour.

For demonstration purposes, I propose that government evaluate an approximate **50/50 resource-allocation hypothesis**: approximately half of the relevant demonstration envelope directed toward scalable AI-enabled infrastructure and continuous availability, and approximately half directed toward human professional capacity, supervision, intervention, complex care, safety, and governance.

This ratio should not be interpreted as a proposed permanent national formula, an ideological commitment, or an assumption about the eventual optimal allocation. It is deliberately presented as an **empirical starting point**. The appropriate long-term allocation—if the model proves worthy of further development—should be determined by evidence concerning safety, utilization, professional standards, user needs, privacy, multilingual performance, accessibility, supervision requirements, escalation rates, workforce effects, costs, user trust, and measurable public outcomes.

The central fiscal question is therefore not whether AI can produce an individual conversation at lower unit cost than a professional. Such a comparison would be incomplete because responsible public deployment requires substantial investment in secure infrastructure, privacy, cybersecurity, accessibility, testing, safety systems, human supervision, evaluation, and governance.

The more meaningful question is: **Can a complete human-AI system generate substantially greater access, continuity, safety, and public value from a comparable funding envelope while protecting appropriate professional care?**

The demonstration should therefore examine whether this mixed architecture can produce greater access, stronger continuity, more effective use of scarce professional attention, broader linguistic and geographic reach, meaningful support outside conventional operating hours, responsible workforce evolution, and greater fiscal productivity **without compromising safety, professional standards, privacy, trust, dignity, or the availability of appropriate human care**.

Its first objective should be **evidence generation**. It should not begin with an assumption of universal implementation. It should not optimize primarily for user volume. It should not treat technological deployment as an objective in itself. It should be structured so that government can legitimately conclude either that the evidence supports further development or that the risks, costs, outcomes, or governance challenges do not justify expansion.

The immediate request is therefore modest in institutional terms even though the long-term vision is ambitious: **Create the fiscal, regulatory, professional, and evaluative space to test the model responsibly**.

The public proposition being tested is more enduring: **No person should be left with absolutely nowhere appropriate and safe to turn simply because professional care is not immediately available, affordable, geographically or linguistically accessible, or because the right professional match has not yet been found.**

The demonstration would be temporary. The public capability it is designed to evaluate could, if the evidence warrants it, become durable.

Purpose: Building the Missing Layer Around Professional Care

The purpose of the initiative would be to create and evaluate an appropriate **continuity layer around professional care**. This layer would be particularly relevant for people who recognize that they need help but do not yet know where to begin; who are waiting for professional mental-health services; searching for an appropriate therapist or helping professional; attempting to find a better clinical, cultural, linguistic, or relational fit; moving between programs or providers; or experiencing gaps between counselling, psychotherapy, or other professional appointments.

It could also support people whose time-limited counselling or structured intervention has ended while their underlying circumstances continue; people facing affordability, geographic, linguistic, cultural, disability-related, or operating-hours barriers; and people seeking appropriate non-clinical emotional support, guided reflection, coaching, life-skills assistance, or basic service navigation.

The proposed layer would not assume that AI is the appropriate final destination for the individual. For many people, appropriate professional care should remain the destination. For others, community services, peer support, social services, coaching, primary care, or another form of assistance may be more suitable. In higher-risk circumstances, crisis or emergency services may be necessary. The value of a virtual safe space lies precisely in not requiring the person to know the final destination before they are permitted to begin receiving appropriate support.

A person should not have to fall into a complete support vacuum while searching for the right destination. This principle reflects a basic reality of help-seeking. People do not move neatly through mental-health systems according to institutional boundaries. They wait. They search. They encounter provider mismatch. They complete programs while their lives continue. They require support between appointments. Their circumstances change. Needs recur. New challenges emerge. Sometimes they know exactly what kind of professional assistance they require; sometimes they simply know that they are struggling and need somewhere appropriate to begin.

The proposed initiative would test whether continuity can be created around those transitions. A person could access the foundational layer before entering professional care. They could use it while waiting for an appointment, where appropriate. They could return between professional encounters. They could access it while searching for another provider after an unsuccessful match. They could continue to

use appropriate non-clinical functions after a time-limited program ends, and they could return when a later life difficulty creates a new need for support or navigation.

This continuity should not be confused with indefinite AI therapy. It is better understood as **persistent access to an appropriate foundational support layer within clearly defined boundaries**. Professional services would necessarily remain finite.

The foundational safety net should not have to disappear simply because one episode of professional service has ended.

The Long-Term Vision: From Episodic Access to a Continuity Guarantee

The larger significance of the proposed demonstration is that it would test the operational and economic foundation for a new form of public assurance: not unlimited professional care at every moment, but **continuous access to an appropriate first layer of support when professional care is not immediately available**.

Historically, such a promise has been difficult to contemplate because the economics of supportive interaction were tightly tied to human labour. If every additional conversation required another therapist, counsellor, social worker, coach, navigator, or other human provider to become personally available, then unlimited continuous access would imply an unrealistic requirement for unlimited professional hours.

Artificial intelligence changes that scarcity assumption for certain carefully bounded functions. It does not make psychotherapy unlimited. It does not make clinical care unnecessary. It does not eliminate the importance of professional relationships or human judgment. However, it may make an **appropriate foundational layer of support substantially more continuously available** than would be possible through one-to-one human labour alone.

The long-term objective is therefore not unlimited AI psychotherapy. It is more precise: **continuous access, within appropriate service, safety, privacy, and professional boundaries, to a free foundational mental-health and emotional-support safety net.**

If the demonstration establishes sufficient evidence of safety, usefulness, equity, trust, accessibility, and fiscal sustainability, Canadians could eventually have access to an appropriate first-line support capability 24 hours a day, 365 days a year, without paying out of pocket for each interaction, without exhausting an arbitrary number of foundational supportive interactions, and with multilingual text and audio-enabled options designed around accessibility and privacy.

Such a capability could remain available while a person is waiting for care, between appointments, searching for an appropriate provider, navigating a transition between services, or after a structured program ends. Its value would lie not in replacing specialized human care, but in reducing the periods during which people are left with no appropriate support at all.

In human terms, the desired public expectation is simple: “My country has my back.” This should not be understood as a promise that government can make every therapist, psychologist, counsellor,

social worker, physician, or other professional continuously available to every person. No public system can eliminate the scarcity of human time. It means something more achievable: **Professional scarcity should not automatically become complete abandonment**. This is why the proposal distinguishes between two complementary aspirations.

The Care Guarantee

No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.

The Care Guarantee addresses the affordability of professional expertise.

The Continuity Guarantee

No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.

The Continuity Guarantee addresses the gaps created by finite professional availability.

The two guarantees serve different purposes and should not be conflated. The Care Guarantee protects access to professional depth. The Continuity Guarantee protects against periods of complete absence of appropriate support.

Professional care provides depth. AI-enabled infrastructure can provide availability and scale. Qualified human supervision provides accountability and safety. Government provides standards, equitable access, institutional durability, and public responsibility. The proposed demonstration should determine whether these elements can be integrated safely, effectively, equitably, and sustainably enough to make the Continuity Guarantee a credible long-term public objective.

A Virtual Safe Space as Public Infrastructure

A virtual safe space should not be evaluated merely as software, a chatbot, or another digital mental-health product. If the underlying capability proves safe and valuable, it should be considered as a potential form of **social and digital public infrastructure**.

Public infrastructure is valuable partly because certain capabilities become dependable rather than episodic. Libraries, parks, and community centres serve different purposes from healthcare, and the analogy should not be taken literally. Their relevance lies in a more limited principle: public institutions sometimes create value by ensuring that useful social capacity exists before the precise moment of need can be predicted.

A human-supervised virtual safe space could serve a comparable continuity function in the digital environment. It could provide an appropriate place where a person can begin, reflect, organize thoughts, receive bounded emotional support, practice life skills, engage in non-clinical coaching, understand

available options, navigate services, and connect with qualified human care when human care becomes necessary. Its basic promise would be: **There is somewhere appropriate I can turn right now.**

That promise has value even when the virtual safe space is not—and should not be—the person's ultimate destination. A bridge derives its value partly from helping people reach somewhere else. The infrastructure should therefore be designed **around professional care rather than in competition with it**. It should help people enter appropriate services, remain supported while waiting, maintain continuity between professional encounters, and navigate transitions when a particular service or provider is no longer available.

The ultimate public asset would not be one specific chatbot, AI model, nonprofit organization, commercial provider, or technology company. It would be the underlying capability: **free-at-the-point-of-appropriate-use, continuously available, multilingual, trauma-informed, accessible, privacy-preserving, human-supervised emotional-support and navigation infrastructure connected to appropriate professional care.**

Technology will change. Vendors will change. AI models will change. Organizations and procurement arrangements will change. If the capability itself proves to have durable public value, government policy should be designed so that the **public promise can endure even when the technology delivering it changes**.

This leads to an important governance principle: **Government should protect the promise, not the platform.** The infrastructure would therefore be technological in its means, but social in its purpose: To create dependable continuity within responsible boundaries and accountable human governance.

Responsible Automation and the Protection of Human Attention

The demonstration should also test one of the central propositions emerging from frontline experience: **different functions within mental-health and supportive-service programs require different levels of human expertise**.

Structured counselling pathways, i-CBT interventions, coaching programs, Employee Assistance Programs, psychoeducational services, intake processes, and other supportive models may contain combinations of structured questioning, routine psychoeducation, guided reflection, goal-setting, journaling prompts, coping-skills education, behavioural exercises, routine check-ins, progress tracking, service navigation, resource information, structured homework, and text-based supportive communication.

These functions may be valuable. Some may require professional involvement depending on context, risk, program design, and individual need. The proposal does not assume that they should simply be transferred from humans to AI.

The demonstration should instead examine them at the functional level.

Modern AI systems can increasingly support some activities that are structured, repetitive, informational, protocol-guided, navigational, or text-based. This development creates an opportunity to

ask whether every component of a structured support program must continue to consume the same level of scarce professional attention merely because human labour was historically the only available delivery mechanism.

The correct policy question is not: Can we automate the entire service? It is: **Which functions require professional human expertise from the outset; which can be safely augmented by technology; which may be suitable for bounded automation; and when must responsibility shift immediately to a qualified human?**

That distinction has significant resource implications. A professional hour devoted to a highly standardized function cannot simultaneously be used for complex trauma, difficult relational circumstances, professional judgment, therapeutic relationship-building, risk assessment, ethical ambiguity, direct intervention, or another situation requiring human expertise and presence. Professional attention therefore carries an opportunity cost.

The demonstration should test whether responsible automation can reduce the amount of professional attention consumed by suitable standardized functions and whether that released capacity is actually redirected toward people whose needs require higher levels of human expertise.

The governing principles should be:

- **Automate what is structured and demonstrably appropriate to automate.**
- **Augment what benefits from technology without displacing necessary professional judgment.**
- **Escalate when complexity, risk, professional responsibility, or human relationship requires it.**
- **Protect human attention for what is irreducibly human.**

This is not a professional-replacement agenda. It is a **professional-capacity protection strategy**. The most meaningful workforce metric should therefore not be how many professional tasks the technology performs. It should be **how much professional attention becomes available for people whose circumstances genuinely require human expertise**.

Guiding Principles

The demonstration should be governed by a clear set of principles established before public deployment and refined through co-design, professional consultation, lived-experience participation, legal review, and independent evaluation.

Complementarity, Not Replacement

Artificial intelligence should strengthen the architecture around professional care rather than become a substitute for it. The demonstration should maintain explicit boundaries distinguishing foundational emotional support, non-clinical coaching, navigation, and psychoeducation from psychotherapy, regulated clinical care, specialized treatment, crisis intervention, and emergency services.

Where professional care is required, the system should support access to that care rather than discourage, delay, or compete with it.

Continuity as a Public Objective

The service should be designed around a straightforward principle: a person's access to appropriate support should not collapse to zero merely because they are waiting, between appointments, changing providers, navigating a poor professional fit, completing a short-term program, or temporarily unable to obtain the right professional care.

The objective is not uninterrupted clinical treatment. It is **continuity of an appropriate foundational safety net**.

No Per-Interaction Financial Barrier to the Foundational Layer

The demonstration should test whether appropriate foundational support can be made available without requiring a person to pay out of pocket each time they need to use it. If continuity is the policy objective, affordability must be incorporated into the architecture from the outset rather than treated as an optional add-on.

Human Accountability

The scalability of AI must not dilute accountability. Qualified humans should remain responsible for governance, professional boundaries, safety protocols, escalation standards, quality assurance, ethical oversight, direct intervention where appropriate, and the determination of how the service interacts with professional, community, crisis, and emergency systems. The greater the technological scale, the stronger the human accountability framework must become.

Responsible Automation

Functions should not be automated merely because automation is technically possible. The relevant question is whether AI assistance demonstrably improves access, continuity, or efficiency **without compromising safety, dignity, professional standards, user autonomy, or outcomes**. Automation should therefore be selective, evidence-based, reversible, auditable, and governed by clearly defined use boundaries.

Protection of Scarce Professional Attention

Therapists, psychologists, counsellors, social workers, physicians, and other helping professionals should be deployed where their education, judgment, ethics, relational capabilities, and accountable human presence create the greatest value. The demonstration should specifically measure whether responsible AI assistance enables professionals to devote more time to complex care, therapeutic relationships, safety, supervision, intervention, and professional judgment.

Trauma-Informed Design

Trauma-informed design should be treated as a foundational requirement rather than an optional content feature. The relevant question is not simply whether an AI response is technically plausible. It is also: **How might this interaction affect a person who has experienced coercion, abuse, rejection, discrimination, marginalization, fear, trauma, or institutional betrayal?**

Design and evaluation should therefore consider dignity, autonomy, transparency, choice, trust, non-judgment, avoidance of unnecessary re-traumatization, and the potential psychological consequences of automated interaction.

Privacy by Design

Privacy should be treated as core infrastructure. The demonstration should incorporate data minimization, meaningful and understandable consent, strong cybersecurity, appropriate retention limits, role-based access, transparent rules governing human review, strict secondary-use limitations, and independent privacy auditing. Sensitive emotional disclosures should not be collected simply because technology makes collection possible.

No Individualized Surveillance

Private emotional conversations must not become individualized government intelligence. They should not be repurposed for advertising, unrelated commercial profiling, political targeting, eligibility enforcement unrelated to the service, or other forms of individualized surveillance. Where frontline learning is part of evaluation, it should rely on appropriately de-identified, aggregated, purpose-limited information subject to rigorous governance.

The governing principle should be: **The purpose is not watching people at scale. It is understanding systemic barriers at scale.**

Accessibility as Infrastructure

Accessibility should be built into the service from the beginning rather than retrofitted after deployment. The initiative should consider multilingual interaction, text, audio-enabled access, screen-reader compatibility, simplified-language options, neurodivergent accessibility, low-bandwidth functionality, rural and remote access, disability-inclusive interface design, and different levels of digital literacy.

Equally important, people who cannot or do not wish to use AI should retain appropriate human and non-AI pathways. A public system should not transform digital inclusion into digital coercion.

The underlying public message should be simple: **You belong here.**

Cultural Responsiveness

Translation alone is not cultural responsiveness. The demonstration should assess whether interactions appropriately account for cultural context, communication norms, family structures, trust,

stigma, understandings of distress, and different relationships to institutions and professional care. Particular attention should be paid to the possibility that standardized AI systems may reproduce cultural assumptions that do not fit the people being served.

Professional Partnership

Relevant helping professions should participate meaningfully in the design and governance of the initiative. Therapists, psychologists, counsellors, social workers, physicians, clinical supervisors, and other qualified professionals should help shape professional boundaries, escalation protocols, training requirements, quality standards, safety mechanisms, and evaluation criteria. AI should be **professionally governed, not merely professionally tolerated**.

Lived-Experience Participation

People who have personally navigated mental-health and support systems should have a substantive role in co-design and evaluation. Participation should include, where feasible, people with experience of waitlists, affordability barriers, provider mismatch, short-term programs, Employee Assistance Programs, language or cultural barriers, rural and remote access challenges, disability-related barriers, community and nonprofit services, and AI-enabled support.

A public virtual safe space should not be designed only *for* people. It should be designed *with* them.

Resource Stewardship

The initiative should evaluate the use of both financial resources and professional human attention. Its objective should not be to minimize expenditure at any cost. It should determine whether public resources are being allocated to the functions where they generate the greatest public value. The governing fiscal principle is **better architecture, not indiscriminate austerity**.

Workforce Evolution, Not Workforce Elimination

The initiative should deliberately examine how professional roles may evolve as routine functions become AI-assisted. Potential roles may include AI-support supervision, escalation and intervention, complex-care practice, safety review, quality assurance, trauma-informed AI evaluation, clinical advising, culturally responsive AI review, professional governance, and human-machine service design. The objective should be to increase the proportion of professional time devoted to high-value human functions, not simply to reduce professional headcount.

Job evolution, not job elimination.

Evidence Before Scale

National expansion should follow evidence rather than precede it. The demonstration should evaluate safety, usefulness, trust, equity, accessibility, privacy, continuity, professional sustainability, workforce effects, governance, and fiscal value before any decision concerning broader deployment. The first phase should optimize for **learning quality**, not maximum user volume.

Technology Neutrality and Public Control of the Standard

Government should define the public capability and its governing standards: scope, professional boundaries, privacy requirements, accessibility obligations, safety requirements, human-supervision expectations, escalation pathways, evaluation criteria, and accountability. No single AI model, technology company, nonprofit organization, contractor, or founder should define or permanently control the national architecture. Canada should invest in a **public capability**, not structural dependence on a particular technology provider.

Durability Beyond Individual Organizations

If the capability proves valuable, its continued existence should not depend indefinitely on one founder, nonprofit organization, corporation, donor, AI provider, or short-term funding cycle. Social innovators can identify unmet needs and demonstrate possibilities. They should not be expected to carry essential public infrastructure indefinitely. A nationally significant safety net should not contain a single organizational point of failure.

What This Proposal Is—and Is Not

Clarity about scope is essential because the proposal could otherwise be misunderstood as an argument for replacing professional mental-health care with artificial intelligence. It is not.

This proposal does **not** recommend replacing psychotherapy, counselling, psychology, social work, medicine, or other professional services with AI. It does not propose unlimited AI psychotherapy. It does not argue that professional care is unnecessary or that every emotional interaction should be automated. It does not authorize AI to diagnose, prescribe, or independently perform regulated clinical functions beyond applicable legal authority. It does not replace crisis or emergency systems. It does not imply that existing i-CBT programs, Employee Assistance Programs, counselling programs, community services, or publicly funded interventions lack value. And it does not propose automating human work merely because automation is technically possible.

Nor is the proposal a mechanism for commercializing emotional disclosures, weakening privacy, reducing human accountability, creating individualized government surveillance, or measuring success primarily through reductions in professional employment or public spending. Its proposition is narrower and more disciplined:

- **Use AI where availability, structured support, multilingual reach, repeatability, and scale can create public value responsibly.**
- **Use qualified humans where judgment, complexity, ethics, professional responsibility, safety, intervention, relationships, and genuine human presence matter most.**
- **Do not consume scarce professional attention on functions that can be responsibly supported through other forms of capacity merely because those functions historically required a person.**
- **Reallocate resources so that computational capacity creates appropriate abundance while professional human attention is protected for complex human needs.**

The objective is not unlimited technology. It is **continuity without abandonment**.

The long-term public promise, if evidence eventually demonstrates that the model is safe, effective, equitable, trusted, and economically sustainable, can be stated simply:

- **Professional care when professional care is needed.**
- **An appropriate foundational safety net when that care is not immediately available.**
- **No per-interaction out-of-pocket barrier to that foundational layer.**
- **Multilingual and accessible entry points.**
- **Qualified human oversight and clear professional boundaries.**
- **Somewhere appropriate and safe to turn when the next difficult moment comes.**

The proposed **Human-Supervised Virtual Safe Space Demonstration Initiative** would give Canada an opportunity to determine whether responsible AI, protected professional human capacity, strong safeguards, and public accountability can make that vision both credible and sustainable. The demonstration should not presume that the answer is yes. It should give Canada a responsible way to find out.

4. A Three-Layer Continuum of Support

The proposed model is built around a straightforward principle: **Not every supportive function requires the same level of human expertise, but every level requires appropriate boundaries, accountability, and a clear pathway to human care.**

The purpose of the three-layer architecture is therefore not to separate artificial intelligence from human services, nor to create parallel systems that compete with one another. It is to assign different functions to the form of capacity best suited to perform them, while maintaining continuity across the entire support journey.

The model seeks to create abundance where technology can safely create abundance; accountable human supervision where judgment and governance are required; and protected professional capacity

where regulated expertise, therapeutic depth, complexity, and genuine human presence are indispensable. This architecture is intended to make continuity possible without confusing continuity with clinical care. It also gives practical effect to the resource-allocation principle at the centre of this proposal. Structured, repetitive, informational, navigational, and highly text- or audio-mediated functions should not automatically consume scarce professional attention when they can be responsibly supported through appropriately governed AI infrastructure. At the same time, AI should not be assigned functions that require regulated expertise, diagnosis, treatment, complex risk assessment, therapeutic relationships, specialized intervention, or professional accountability. The three-layer model provides a practical mechanism for drawing and governing those distinctions.

The architecture can be summarized simply:

- **Layer 1 creates availability.**
- **Layer 2 creates accountability.**
- **Layer 3 provides professional depth.**
- **Government provides continuity, standards, and public responsibility.**

Layer 1: The Always-Available Virtual Safe Space

The first layer would provide **free-at-the-point-of-appropriate-use, immediate, multilingual, 24/7 access to clearly bounded AI-assisted support**. Its purpose is to ensure that a person does not always have to begin from zero, remain entirely unsupported while waiting, or incur an out-of-pocket charge each time they need an appropriate first place to turn.

Within clearly defined service boundaries, Layer 1 could provide basic emotional support, guided reflection, non-clinical coaching, holistic life-skills support, routine psychoeducational information, structured exercises, journaling prompts, coping-skills practice, goal clarification, basic problem organization, service navigation, resource information, and other appropriate forms of first-line support. Access could be provided through text and audio-enabled interaction and designed to support multilingual use, accessibility needs, different levels of digital literacy, and use outside conventional service hours.

The layer could also provide continuity during common gaps in the service journey: while a person is waiting for professional care, between professional appointments, while searching for a more appropriate provider, during transitions between programs, or after a short-term intervention has concluded.

Layer 1 is particularly suited to functions that are **structured, repetitive, informational, protocol-guided, reflection-based, navigational, or primarily text- or audio-mediated**. These are the functions for which AI may be able to create substantial public value by increasing availability without requiring an additional therapist, counsellor, social worker, navigator, coach, or other professional to become personally available for every additional interaction.

In this sense, Layer 1 purchases something conventional one-to-one service systems have historically struggled to provide at population scale: **continuous availability**. That availability must remain carefully bounded. Layer 1 would not independently diagnose. It would not prescribe treatment. It

would not provide unbounded psychotherapy. It would not represent itself as a licensed professional. It would not determine that professional care is unnecessary when circumstances indicate otherwise. It would not replace crisis, emergency, medical, or specialized professional services. Its role would be more precise: to provide an appropriate first layer of support, help people organize and understand their needs, maintain continuity during periods of delay or transition, and support movement toward human care when human care is required.

The foundational promise of Layer One is therefore simple: **There is somewhere appropriate I can turn right now**. For the public, this means that an appropriate first-line safety net need not disappear simply because an office is closed, a short-term counselling allocation has ended, the next appointment is weeks away, the first provider was not the right fit, a person is still determining what kind of help they need, or private care is temporarily unaffordable.

Layer 1 is therefore the technological foundation of the proposed **Continuity Guarantee**. It creates the possibility of maintaining access to an appropriate first layer of support 24 hours a day, 365 days a year, without a per-interaction out-of-pocket charge, through multilingual text and audio-enabled interaction, before, between, and alongside professional care. This does not make professional care unlimited, but it makes the surrounding safety net potentially continuous.

Layer Two: Human-Supervised Support and Accountability

The second layer provides the **human governance necessary to make the first layer safe, accountable, and worthy of public trust**. Artificial intelligence may create scale. It should not create unaccountable scale.

Qualified humans would therefore oversee the operation of the AI-enabled layer, review system performance, monitor professional boundaries, evaluate defined safety concerns, manage escalation, support navigation, review difficult or ambiguous interactions where appropriate, assess cultural and linguistic performance, conduct quality assurance, and ensure that automation remains suitable for the functions assigned to it.

Depending on the eventual service design, this layer could include qualified counsellors, mental-health professionals, clinical advisers, supervisors, social workers, appropriately supervised trainees, non-clinical coaches operating within defined scopes, service navigators, quality-assurance professionals, AI-safety reviewers, cultural reviewers, escalation specialists, and other appropriately trained personnel. Their function would not be to manually reproduce every interaction that an AI system could responsibly provide. That would undermine the logic of the model. Their role is to perform the distinctly human functions that make scaled support accountable.

Layer 2 would therefore continuously ask questions such as: Is the system remaining within its intended scope? Are users being told clearly what the system can and cannot do? Are appropriate professional referrals being made? Are defined safety signals being handled correctly? Are necessary escalations being missed? Are unnecessary escalations creating burdens or discouraging use? Are cultural, linguistic, accessibility, or bias-related problems emerging? Are some groups receiving poorer-quality interactions? Are automated functions generating confusion, inappropriate reassurance,

overdependence, or other unintended effects? At what point should an interaction cease to remain automated and move into qualified human hands?

The significance of Layer 2 extends beyond supervision of individual interactions. It would also provide supervision of the infrastructure itself. This distinction matters for responsible AI governance. In a system capable of serving large numbers of people, safety cannot depend solely on correcting problematic individual outputs after they occur. Government and professional partners need visibility into patterns, failure modes, population effects, escalation behaviour, accessibility performance, and the continuing appropriateness of automated functions.

Layer 2 also creates an important workforce opportunity. As AI systems assume some bounded structured and repetitive functions, helping professionals may increasingly contribute through higher-value roles such as AI-support supervision, complex-case review, escalation, trauma-informed evaluation, quality assurance, clinical advising, professional-boundary governance, cultural review, safety oversight, and human-machine service design.

The objective is therefore not to remove professionals from the support system. It is to **use their expertise differently and more deliberately**. Layer 2 is the accountability bridge between technological scale and professional human care.

Layer 3: Licensed and Specialized Professional Care

The third layer contains the functions for which **regulated expertise, professional judgment, therapeutic depth, specialized knowledge, relational care, and direct human accountability are indispensable**. Depending on jurisdiction, professional scope, and individual need, this layer may include psychologists, psychotherapists, physicians, psychiatrists, social workers, counsellors, trauma specialists, and other regulated or appropriately credentialed professionals.

This is the layer where professional human expertise should be protected most carefully. Its functions may include assessment, diagnosis where legally authorized, treatment planning, psychotherapy, complex risk evaluation, trauma treatment, specialized clinical intervention, therapeutic relationship-building, ethical decision-making, management of ambiguity, high-risk care, and other circumstances in which professional judgment and genuine human presence are central to the service itself.

Layer Three remains fundamentally human. AI does not eliminate the need for this layer. **AI should help protect professional capacity for it.** That distinction is central to the proposal.

If Canada faces shortages of therapists, psychologists, counsellors, social workers, physicians, and other helping professionals, then the policy response should not be to consume their limited hours unnecessarily on activities that can be responsibly supported by other forms of capacity. A professional who spends less time manually delivering standardized psychoeducation, repetitive exercises, routine text prompts, basic navigation, or other appropriately automatable functions may be able to spend more time with people whose circumstances require trauma expertise, therapeutic relationships, complex judgment, risk assessment, intervention, or human presence.

The policy objective is therefore not fewer professionals. It is **more professional attention available for genuinely professional work.** Layer 3 is where the system preserves depth.

- Layer One creates continuity.
- Layer Two creates accountability.
- Layer Three provides professional expertise and human care.
- Together, the three layers create a continuum that no single layer could responsibly provide alone.

Illustrative User Journey

The value of the model becomes clearer when viewed from the perspective of a person seeking support.

1. A person recognizes a need.

A person may be experiencing grief, stress, loneliness, relationship difficulty, anxiety, uncertainty, a life transition, or another emotional or practical concern. They may know that they need support without yet knowing whether the appropriate destination is psychotherapy, counselling, coaching, medical care, peer support, community services, social services, or another form of assistance.

2. The person receives immediate access to the virtual safe space.

Rather than waiting until a conventional appointment becomes available, the individual can begin through text or audio-enabled interaction. They do not need to pay for every appropriate foundational interaction, wait for business hours, or correctly identify a professional category before entering the support system. This creates an accessible starting point rather than requiring the individual to navigate the entire service landscape before receiving any assistance.

3. The person receives appropriately bounded support.

Depending on the situation, Layer 1 may offer guided reflection, emotional support, structured exercises, non-clinical coaching, psychoeducation, basic goal clarification, life-skills assistance, or service navigation. The system should remain transparent throughout about its identity, scope, limitations, and the circumstances in which professional care may be more appropriate.

4. Safety and boundary systems remain active.

1Automated safeguards should identify predefined situations requiring additional caution, different response handling, specific safety information, restricted functionality, escalation, or human review. These safeguards should support—not replace—professional governance.

5. Human involvement occurs where it adds value or becomes necessary.

An appropriately trained human may become involved when the situation becomes more complex, the system reaches a defined boundary, a safety concern arises, the user requests human involvement where such access is available, or professional review is otherwise indicated. The purpose of escalation is not to demonstrate that the AI system has failed. Escalation is an intended part of the design. A responsible system should know when continued automation is no longer appropriate.

6. The individual is connected to an appropriate next level of support.

Depending on the person's circumstances, the next step may involve a therapist, psychologist, physician, social worker, counsellor, community organization, social service, peer-support service, crisis resource, or emergency service. The system should assist navigation without presenting itself as a substitute for the destination.

7. The foundational safety net remains available while the person waits.

This is one of the most important features of the architecture. Referral should not necessarily mean the immediate termination of all support.If a professional appointment is several days or weeks away, the foundational layer may remain available within its appropriate scope so that the individual does not return to complete absence of support during the waiting period.

8. Professional care is provided when professional care is needed.

Layer 3 becomes central whenever professional judgment, assessment, treatment, therapeutic depth, specialized intervention, or regulated expertise is required.

9. The foundational layer may remain available between appointments.

The person should not be forced into a false choice between AI-enabled support and professional care. Where clinically and operationally appropriate, the two can function complementarily. The professional remains responsible for professional care. The foundational layer remains a supplementary continuity resource.

10. The person can return when a later need emerges.

The end of one episode of professional or structured care does not imply that the individual will never need support again. Circumstances evolve. Stressors recur. Relationships change. New difficulties arise. A durable foundational layer would allow the person to return without requiring every new difficult moment to begin with another search for an entry point. The result is not perpetual treatment. It is continuity of access.

The professional may change. The funded program may conclude. The person's needs may evolve. The safety net does not automatically disappear. That is the practical meaning of the Continuity Guarantee.

The Three-Layer Division of Labour

The proposed system relies on an intentional division of responsibility rather than a binary choice between humans and machines.

Layer 1: AI Provides Availability and Scale

AI should be used where its comparative advantages include continuous availability, multilingual interaction, structured guidance, routine psychoeducation, basic navigation, reflection, repeatable exercises, goal tracking, and scalable low-marginal-cost interaction. Its purpose is to reduce avoidable gaps in foundational support.

Layer 2: Humans Provide Accountability and Escalation

Qualified humans should provide supervision, quality assurance, safety review, professional-boundary oversight, ethical governance, culturally responsive review, intervention where appropriate, and escalation. Their purpose is to ensure that technological scale does not come at the expense of accountability, dignity, or professional standards.

Layer 3: Professionals Provide Expertise and Depth

Qualified and regulated professionals should provide assessment, judgment, therapeutic relationships, specialized treatment, complex risk evaluation, trauma care, ethical decision-making, and genuine human care. Their professional attention should be protected for the circumstances in which it creates the greatest value. The model can therefore be expressed through four complementary responsibilities:

- AI provides availability.
- Humans provide accountability.
- Professionals provide depth.
- Government provides continuity, standards, equitable access, and public responsibility.

This division of labour is not intended to create rigid walls. The value of the model lies in the ability to move safely between layers as a person’s needs change.

A Five-Level Safety Architecture

Accessibility and safety should not be treated as competing investments. They are part of the same public responsibility. The more widely available an AI-enabled support service becomes, the stronger and more explicit its safety architecture must become. The demonstration should therefore incorporate at least five reinforcing levels of protection.

Level 1 — Universal but Bounded Access

Users should receive immediate, multilingual, accessible supportive interaction within a clearly defined scope. The system should explain in understandable language that it is AI, what functions it can provide, what it cannot provide, when professional care may be more appropriate, and how other forms of assistance can be accessed.

Universal availability should never be confused with unlimited scope. The service may be continuously available while remaining tightly constrained in function.

Level 2 — Automated Safety and Boundary Systems

Technical mechanisms should identify predefined circumstances that require a different response, additional caution, specific safety messaging, restricted functionality, or escalation. Automated tools may also help detect when an interaction is moving beyond the intended scope of foundational support. Such mechanisms should themselves be evaluated continuously. Safety systems can fail both by missing circumstances that require escalation and by escalating excessively. Both types of error matter.

Level 3 — Qualified Human Oversight

Qualified professionals and trained reviewers should oversee system performance, boundary adherence, safety patterns, escalation behaviour, cultural and linguistic issues, quality, accessibility, and defined circumstances requiring human attention. Human oversight should be focused and governed. It should not become indiscriminate monitoring of private conversations. The principle is **accountable safety, not surveillance**.

Level 4 — Human Intervention and Connection to Care

Where appropriate, the system should facilitate timely transition to qualified human support, licensed professional care, community resources, social services, crisis services, or emergency services. Escalation is not an operational failure. It is one of the functions that makes the system responsible. A safe AI-enabled system should recognize when it has reached the limits of its role.

Level 5 — Independent Evaluation

External evaluation should assess the full human-AI system rather than focusing only on model performance. Evaluation should examine safety, privacy, cybersecurity, bias, cultural responsiveness, accessibility, multilingual quality, user trust, user dignity, user outcomes, professional workload, workforce effects, escalation quality, missed escalation, unnecessary escalation, continuity, utilization, fiscal productivity, use of professional attention, and unintended consequences.

Independent evaluators should also answer a critical workforce question: **Does responsible automation actually release professional capacity for higher-value human work, or does it merely create new forms of workload, oversight burden, or risk?** Evidence should determine redesign.

Safety, Scale, and Accountability Must Grow Together

The central governance principle of the model is straightforward: **The more scalable the system becomes, the more deliberate human accountability must become.**

AI can make supportive interaction available at a scale that conventional one-to-one service delivery cannot easily reproduce. That capability creates public value, but it also magnifies the consequences of errors. A system capable of serving thousands or potentially much larger populations cannot rely on the assumption that most individual interactions will simply be reasonable enough. Public infrastructure requires systematic governance.

At minimum, this means clear professional boundaries, privacy and cybersecurity protections, transparent escalation rules, appropriate human review, continuous quality assurance, lived-experience participation, independent auditing, accessible complaint and redress mechanisms, and public accountability. Scale should therefore never become an independent policy objective.

The desired relationship is:

more availability → stronger safeguards

more automation → more deliberate oversight

more scale → greater accountability

This is one of the principal distinctions between responsible public infrastructure and an ordinary consumer-facing chatbot.

A consumer product can often rely largely on individual terms of service and user choice. A publicly supported mental-health and emotional-support capability would require a higher standard because government would be explicitly inviting people to rely on it as part of a social safety net. Public trust must therefore be earned through governance, not assumed because the technology is useful.

The Public Promise Created by the Three-Layer Model

The importance of the proposed architecture ultimately lies not in its internal technical complexity, but in the experience it could create for the public. A person should not need to understand the distinction between computational infrastructure, AI inference costs, professional supervision, clinical pathways, procurement arrangements, or government funding structures in order to benefit from the system. Their experience should be simpler. They should be able to know:

- There is somewhere appropriate I can turn right now.
- I will not be charged every time I need an appropriate foundational supportive interaction.
- I can communicate through text or audio-enabled interaction.
- I can seek support in a language I understand, where the system can safely support that language.

- I can be connected to qualified human and professional care when my circumstances require it.
- I can continue to have an appropriate place to turn while I am waiting.
- If one provider is not the right fit, I do not have to navigate the next search in complete isolation.
- If a short-term program ends, the entire support architecture does not disappear with it.
- And human professionals remain protected for the moments when human expertise, responsibility, and connection matter most.

At its most human level, that public experience can be expressed in a simple phrase: *"My country has my back."* The phrase should not be understood as a promise that government can eliminate all suffering, remove every waitlist, or make every professional continuously available. It represents a more grounded public commitment: Finite professional capacity should not require people to fall into complete absence of support.

This is what the three-layer continuum is intended to make operationally and economically possible.

- Layer 1 creates scalable availability.
- Layer 2 creates accountable human governance.
- Layer 3 protects professional depth.

Government can bind the three together through standards, funding, evaluation, privacy protections, workforce policy, professional partnership, and durable public responsibility. The objective is therefore not to place artificial intelligence at the centre of Canada's mental-health system. It is to place *continuity* at the centre.

AI is one instrument for creating that continuity. Human beings remain the source of professional judgment, accountable intervention, therapeutic depth, ethical responsibility, and genuine connection. The proposed three-layer architecture brings those capabilities together in a form that can be tested, measured, governed, corrected, and—only if the evidence supports it—developed into durable public infrastructure.

5. Human-AI Symbiosis and the Better Allocation of Professional Attention

Moving Beyond the False Choice Between Humans and Machines

This proposal rejects the assumption that Canada must choose between human-delivered services and artificial intelligence. That is the wrong policy frame. The more useful question is functional:

Which activities can be responsibly supported by scalable technology, and which should remain protected for qualified human professionals because their value depends on judgment, relationship, ethics, complexity, accountability, or genuine human presence?

This distinction matters because AI and professional human labour have different comparative advantages. Artificial intelligence may be particularly useful where the principal requirements are continuous availability, multilingual interaction, structured reflection, repeatability, routine psychoeducation, guided exercises, non-clinical coaching, basic service navigation, goal clarification, appropriate low-complexity emotional support, and continuity outside conventional service hours.

These functions are not necessarily trivial. Many are valuable precisely because they help people begin, organize their thoughts, develop practical skills, obtain information, remain engaged with a support pathway, or navigate toward the next appropriate service. The policy question is whether every instance of these functions requires a highly trained professional to personally generate every interaction.

Human professionals, by contrast, remain indispensable where the service depends on distinctly human capabilities. These include professional judgment, therapeutic and helping relationships, complex or ambiguous circumstances, trauma-informed professional care, ethical reasoning, individualized interpretation, risk assessment, escalation, culturally nuanced decision-making, management of uncertainty, complex family or relational dynamics, professional accountability, and genuine human presence.

The appropriate division is therefore not “humans or AI.” It is a deliberate allocation of different forms of capacity.

- AI for availability, structure, continuity, multilingual reach, and scale.
- Humans for judgment, complexity, relationships, safety, ethics, and accountability.

This distinction is also central to the fiscal case for the proposed demonstration. If every additional supportive interaction must always be purchased through another increment of professional human labour, then continuous and broadly accessible support will remain constrained by staffing, scheduling, geography, language availability, and cost.

If, however, government can responsibly use computational capacity for functions in which availability, repetition, structure, and scale are the principal requirements, professional human attention can be protected for the circumstances in which it creates substantially greater value.

The proposition is therefore not that professional work has become less important. It is that professional attention has become too valuable to allocate indiscriminately.

Professional Human Attention Is a Scarce Public Resource

Mental-health workforce discussions typically focus on two forms of scarcity: funding and the number of available professionals. Both are important. This proposal introduces a third resource that deserves equally serious consideration: **professional human attention**.

A therapist's hour can only be used once. A counsellor cannot simultaneously facilitate a highly standardized interaction and sit with a person experiencing complex trauma. A social worker cannot use the same hour for repetitive service navigation and for a difficult systems intervention requiring advocacy, judgment, and coordination. A supervisor cannot devote the same professional attention both to routine procedural work and to the careful oversight of complex cases.

This creates an opportunity cost. When highly trained professionals spend substantial portions of their working time on structured, repetitive, protocol-guided, informational, or primarily text-based functions that could potentially be supported through appropriately governed technology, the cost is not limited to salary or contracted service expenditure. There is also a less visible cost: The professional is unavailable during that same period to someone whose needs require their full expertise.

Many structured mental-health and supportive-service environments contain activities such as routine psychoeducation, standardized exercises, guided reflection, goal tracking, journaling prompts, basic navigation, information delivery, routine check-ins, follow-up activities, and text-based supportive communication. These activities can be useful and, in some circumstances, may still require human involvement. The demonstration should not assume otherwise. It should test the question at the level of specific functions and risk profiles.

Historically, trained people performed nearly all such work because no viable alternative could sustain responsive natural-language interaction, explain concepts dynamically, adapt structured prompts, operate across languages, or remain continuously available on demand. That technological constraint is changing. Contemporary AI systems can increasingly support some appropriately bounded, repetitive, informational, navigational, and conversational functions through computational resources rather than requiring a professional to personally produce every response.

This creates a legitimate policy question: If professional expertise is scarce, should government continue using that expertise for every function that historically required human labour when some of those functions may now be responsibly augmented or automated?

The answer should not be ideological. It should be empirical. Where responsible AI assistance improves access without compromising safety, quality, dignity, professional boundaries, or outcomes, government should test whether professional time can be redirected toward higher-complexity human needs. Where a function depends on judgment, therapeutic depth, risk assessment, ambiguity, professional responsibility, or relationship, human involvement should remain central. The workforce principle is therefore analogous to the broader resource-allocation model: **Allocate professional human attention according to comparative advantage.**

Job Evolution, Not Job Elimination

The proposed initiative is not designed to eliminate therapists, counsellors, psychologists, social workers, physicians, or other helping professionals. It is designed to **increase the proportion of their working time devoted to the functions for which professional human expertise is most valuable**.

If AI-enabled infrastructure can safely absorb or assist some appropriate structured, repetitive, informational, and navigational work, professionals may have greater capacity for deep listening,

therapeutic relationship-building, complex cases, trauma-informed care, professional judgment, ethical decision-making, risk assessment, ambiguous circumstances, human intervention, difficult navigation, supervision, quality assurance, escalation, and situations in which human presence itself forms part of the intervention. The desired workforce outcome is therefore not simply lower labour intensity. It is higher-value professional practice.

A system in which professionals spend less time reproducing standardized functions could potentially allow them to spend more time with people who have been waiting for care, whose needs do not fit standardized pathways, who have struggled to find an appropriate provider, whose circumstances involve significant complexity or risk, or whose outcomes depend on trust, interpretation, professional judgment, and relationship.

This suggests a broader workforce transition from task substitution toward role evolution. New or expanded roles could include AI-support supervision, escalation practice, AI-safety review, trauma-informed AI evaluation, clinical AI advising, quality assurance, professional-boundary review, culturally responsive AI evaluation, human-machine service design, digital mental-health governance, AI-assisted navigation, complex-care specialization, and workforce-transition training.

Some of these roles already exist in limited forms. Others may emerge only if AI-enabled support becomes more deeply integrated into public systems. Government should therefore avoid evaluating the initiative primarily through a labour-displacement lens.

The better question is: **Does this architecture allow qualified professionals to spend more of their limited time doing work that genuinely requires qualified professionals?** That is the meaning of *job evolution, not job elimination.*

Better Use of Existing Professionals Can Complement Workforce Expansion

This distinction has important implications for how Canada responds to workforce shortages. A shortage of mental-health professionals may genuinely mean that Canada needs more qualified professionals. In many settings, expanding training pipelines, improving retention, supporting rural and remote practice, and strengthening access to regulated care may remain necessary.

This proposal does not argue otherwise. However, workforce scarcity can also be intensified when existing professionals devote substantial portions of their time to functions that do not consistently require their highest level of expertise. These are different problems. The first calls for **workforce expansion**. The second calls for **service redesign**. Canada should be prepared to pursue both.

Increasing the number of professionals without examining how their time is used could reproduce the same allocation patterns at a larger scale. Conversely, redesigning services without addressing genuine shortages would simply redistribute an insufficient workforce. A stronger strategy is therefore complementary: Increase professional supply where evidence demonstrates need while redesigning workflows so that professional expertise is concentrated where it creates the greatest value.

Responsible AI assistance could potentially contribute to this second objective by weakening the historical assumption that every additional appropriate supportive interaction must require a

corresponding increase in professional labour. This is why the proposed workforce transition should be understood not merely as a technology strategy, but as a professional-capacity expansion strategy. Its core performance measure should not be: How many professional tasks did AI replace? It should be: How much additional professional attention became available for people whose circumstances genuinely required professional expertise?

This metric would provide government with a far more meaningful indication of whether the initiative is strengthening or weakening the mental-health workforce.

Professionals Must Be Partners in the Transition

A human-supervised public system will not be credible if professionals experience it as technology designed elsewhere and imposed on them after the fact. Therapists, counsellors, psychologists, social workers, physicians, clinical supervisors, and other relevant helping professionals should therefore participate meaningfully in the design of the demonstration from its earliest stages.

Their knowledge is essential because automation decisions in emotionally sensitive services are not purely technical. Helping professionals understand dimensions of care that may be difficult to capture through software development alone: therapeutic alliance, professional boundaries, trauma, relational dynamics, risk, ethical ambiguity, cultural context, professional liability, clinical uncertainty, and the difference between a response that is technically coherent and one that is genuinely appropriate for a particular human being in context.

Their participation should inform service scope, system design, escalation protocols, professional-boundary definitions, quality assurance, safety standards, governance, training requirements, workforce transition, evaluation criteria, and decisions about which functions should or should not be automated.

The question: “Can AI technically perform this function?” is not sufficient. A responsible public system must also ask: “Should AI perform this function?” and “If it does, under what boundaries, safeguards, professional supervision, accountability structures, and escalation rules?” Qualified professionals are indispensable to answering those questions.

The governing principle should therefore be: **AI should be professionally governed, not merely professionally tolerated.**

Professional associations, regulatory bodies, educational institutions, practitioners, supervisors, and relevant workforce representatives may also have an important role in defining competencies for emerging hybrid functions such as AI supervision, safety review, professional-boundary oversight, and human-AI service delivery. The transition should be something the workforce helps shape, not something it is simply expected to absorb.

Workforce Transition Must Be Funded, Not Assumed

If government expects professionals to assume new responsibilities in an AI-enabled service architecture, it should not assume that those competencies will emerge automatically. Workforce transition should therefore be treated as a substantive and funded component of the demonstration.

Professionals participating in the model may require education and practical training in AI capabilities and limitations, human-AI service design, privacy and data governance, algorithmic bias, trauma-informed digital practice, safety-signal review, escalation procedures, quality assurance, professional liability, scope-of-practice boundaries, multilingual and cultural considerations, accessibility, and appropriate collaboration with automated systems.

Supervisors may require additional competencies for evaluating AI-mediated interactions. Service navigators may need to understand when technology-assisted support is appropriate and when it is not. Clinical advisers may need frameworks for reviewing system behaviour without becoming responsible for impossible volumes of individualized oversight. Organizations may need clear protocols governing documentation, accountability, escalation, complaints, and redress.

Government should therefore evaluate whether new professional guidance, continuing-education requirements, supervision models, competencies, credentialing mechanisms, or standards of practice are required. This investment serves two purposes. First, it improves safety and service quality. Second, it makes the workforce transition more credible. Professionals are more likely to participate constructively in system redesign when they can see a clear and respected role for their expertise within the emerging architecture.

The public should not be asked to trust a human-supervised AI-enabled mental-health system if the humans expected to govern it have not been adequately prepared and resourced to do so.

People With Lived Experience Must Also Shape the System

Frontline expertise does not belong only to professionals. People who have navigated mental-health and supportive-service systems possess forms of knowledge that may not be visible in administrative datasets, professional standards, program evaluations, or technology development.

This includes people who have waited for therapy, struggled with the cost of private care, experienced unsuccessful provider matches, repeatedly retold difficult personal histories during intake, used Employee Assistance Programs or short-term counselling, completed structured interventions while their underlying needs continued, relied on nonprofit services, sought support in languages other than English or French, experienced cultural mismatch, encountered disability-related barriers, lived in rural or remote communities, required assistance outside conventional operating hours, or used AI-enabled emotional-support or coaching tools. These experiences should inform the design and evaluation of the demonstration.

Lived-experience participants can help identify whether consent is understandable, whether privacy expectations feel realistic, whether escalation procedures feel supportive or punitive, whether language access is genuinely usable, whether navigation is confusing, whether the distinction between AI

support and professional care is sufficiently clear, whether interfaces create unintended barriers, and whether the service actually feels psychologically safe to the people for whom it is intended. Without these perspectives, Canada risks reproducing digitally many of the access barriers already present in conventional systems.

A platform may technically support multiple languages while failing to communicate appropriately across cultural contexts. It may satisfy formal accessibility standards while remaining difficult for some people with disabilities to use. It may offer immediate access while creating uncertainty about whether the user is interacting with a support tool, a professional, or a healthcare service. These are not merely usability concerns. They affect trust, dignity, safety, and equitable access. The system should therefore be designed not only *for* people with lived experience. It should be designed *with them*.

Professional Expertise and Lived Experience Should Inform One Another

Professionals and people with lived experience should not be treated as isolated consultation groups whose recommendations are combined only after major design decisions have already been made. Their perspectives should interact throughout the demonstration.

Professionals may recognize clinical, ethical, safety, or boundary risks that users do not immediately identify. Users may identify barriers or institutional behaviours that professionals have come to regard as normal. Technology developers may identify capabilities and constraints that affect what can be implemented safely. Privacy and cybersecurity experts may identify data practices that should not be permitted. Accessibility specialists may detect exclusions that are invisible to other participants. Cultural and language experts may identify assumptions embedded in model behaviour. Legal experts may clarify accountability and regulatory boundaries. Researchers can help convert these concerns into measurable evaluation questions.

Responsible governance therefore requires a multidisciplinary co-design structure. That structure should meaningfully include helping professionals, people with lived experience, social innovators, AI developers, privacy and cybersecurity experts, accessibility specialists, cultural and language experts, legal experts, researchers, economists, and public-sector decision-makers.

The purpose is not to require unanimous agreement on every aspect of the model. It is to prevent any one professional group, technology provider, institution, commercial actor, or founder from unilaterally defining what a psychologically safe digital public service should be. A public capability that may eventually serve a highly diverse population requires governance that reflects that diversity of expertise.

Human-AI Symbiosis as a Public-Service Model

The broader workforce vision of this proposal can be described as **human-AI symbiosis**: an institutional model in which different forms of capacity perform different functions while remaining connected through clear governance and accountability.

- Artificial intelligence contributes availability, scale, multilingual reach, structured support, repeatability, and continuity.
- Helping professionals contribute judgment, complexity management, therapeutic and helping relationships, ethics, professional accountability, safety, intervention, and genuine human presence.
- People with lived experience contribute reality testing, context, trust signals, insight into barriers, and direct knowledge of whether the system actually works for the people it is intended to serve.
- Government contributes standards, funding, privacy and safety requirements, regulatory clarity, workforce transition, independent evaluation, equitable access, accountability, and institutional continuity over time.

No single one of these actors should carry the entire burden.

The objective is not to build a mental-health system in which AI dominates human services. Nor is it to preserve existing workflows simply because they are familiar. The objective is to construct an architecture in which each form of capacity is used where it produces the greatest public value.

The governing principles can therefore be summarized as follows:

- AI should create abundance where abundance is safe and appropriate.
- Professionals should provide depth where human expertise and accountability are indispensable.
- People with lived experience should keep the system grounded in the realities of those it serves.
- Government should ensure that the resulting capability remains equitable, accountable, professionally governed, and durable.

This is not human replacement. It is a redesign of how scarce professional human attention is protected and deployed. If successful, the result could help Canada move toward the broader objective of this proposal: **continuous access to an appropriate foundational safety net while preserving professional human care for the moments when human beings matter most**.

The measure of success should therefore never be how much human involvement Canada can remove from mental-health support. It should be how intelligently Canada can use technology to protect and extend the value of human involvement where it is most needed.

6. The Funding and Resource-Allocation Model

The Economic Architecture: Reallocating Scarcity to Create Abundance

The Core Economic Principle

This proposal rests on a central economic proposition: **Canada should distinguish between the cost of availability and the cost of expertise.**

Historically, mental-health and emotional-support systems had little practical choice but to purchase both through the same scarce input: human labour. If a program sought to serve more people, it generally required additional professional or support-worker hours. If it sought to extend operating hours, it required additional staffing. Twenty-four-hour availability required shifts. Additional languages required additional linguistic capacity. Expanding geographic reach often required corresponding organizational and workforce growth. As demand increased, labour requirements generally increased with it.

In practical terms, availability and expertise were economically bundled together because human beings were the primary mechanism through which both could be delivered. Artificial intelligence changes that relationship for some appropriate and carefully bounded functions. It does not eliminate the need for professional mental-health care. It does not make public infrastructure costless. It does not imply that every supportive function can or should be automated. Nor does it eliminate the need for qualified human oversight. What it creates is a new possibility: Government may increasingly be able to **purchase availability differently from expertise**.

Appropriately governed computational infrastructure can potentially provide continuous access, multilingual interaction, structured support, routine information, guided reflection, navigation, and other suitable functions at a scale that does not require an equivalent additional professional interaction every time another person uses the service.

Professional human capacity can then be concentrated where it creates the greatest value: complex judgment, therapeutic relationships, trauma-informed care, risk assessment, ethical ambiguity, intervention, supervision, and genuine human presence. The emerging economic principle is therefore:

- Use computational capacity to create abundance where availability, structure, repetition, and scale are appropriate.
- Protect scarce professional human attention for the circumstances in which human expertise is indispensable.

This distinction becomes particularly important when supportive services are examined at the functional level. Many programs contain combinations of structured questioning, psychoeducation, guided reflection, standardized exercises, goal-setting activities, journaling prompts, routine check-ins, basic navigation, resource information, progress tracking, and other repeatable interactions.

These functions may be useful, and some will continue to require human involvement depending on context and risk. The proposal does not assume otherwise. The demonstration should instead test

which functions can be responsibly AI-assisted, which may be suitable for bounded automation, and which should remain in professional hands from the outset.

Historically, professionals delivered nearly all of these activities because there was no viable alternative. That constraint is changing. The economic implication is significant. If therapists, counsellors, psychologists, social workers, physicians, and other helping professionals are scarce, then their time should itself be treated as a scarce public resource.

A professional hour spent facilitating a highly standardized function that could be safely supported through another form of capacity cannot simultaneously be spent with a person experiencing complex trauma, significant risk, difficult relational circumstances, ethical ambiguity, or another situation requiring professional expertise. The cost of that hour is therefore not merely the salary, fee, or reimbursement attached to it.

There is also an opportunity cost of professional human attention. That opportunity cost is central to this proposal. Resource reallocation is not based on an assumption that professional labour has become less valuable. It is based on the opposite conclusion: **Human professional attention has become too valuable to use inefficiently.**

Why This Matters for Existing and Future Public Spending

This proposal does not begin from the assumption that governments are indifferent to mental health or that existing public expenditures lack value. Governments already devote resources to a range of mental-health, counselling, digital-health, community-support, workplace-support, and contracted service-delivery arrangements.

The specific scale, composition, jurisdictional distribution, and funding mechanisms of those expenditures should be independently documented in the final submission. The relevant policy point, however, is broader: Canada is already making choices about how public resources purchase mental-health and supportive capacity. The more consequential fiscal question is therefore not simply: Should government spend more on mental health? It is also: **How should existing and future funding be structured to produce the greatest possible combination of access, continuity, safety, equity, professional quality, and public value?**

Under conventional service architecture, a significant share of expenditure ultimately purchases units of human time. That structure was rational when virtually every meaningful supportive interaction required a person to deliver it.

The technological environment is changing. If some structured components can now be responsibly supported through scalable digital infrastructure, government has an opportunity to examine the internal architecture of funded services rather than treating every supportive function as economically identical. The relevant questions become more precise:

- Which functions genuinely require professional human expertise?
- Which functions can be responsibly augmented by AI?

- Which structured or repetitive functions may be suitable for automation under clear safeguards?
- Which interactions can begin within a bounded digital support layer and transition to human involvement when complexity or risk increases?

And, critically:

- What happens to the professional capacity released when an appropriate function no longer requires a professional to deliver every interaction manually?

This final question is essential. Resource reallocation has little public value if professional time is nominally "saved" but does not translate into greater access to professional expertise where that expertise is needed. The demonstration should therefore test not only whether AI can perform selected functions appropriately, but whether service redesign actually results in more professional attention being available for complex human needs.

An Approximate 50/50 Demonstration Hypothesis

Within a defined demonstration funding envelope, government should test an approximate 50/50 allocation between AI-enabled public digital infrastructure and human professional capacity and governance. The ratio is intentionally simple. Its purpose is not to prescribe a permanent national funding formula. Nor does it assume that half of every mental-health budget should ultimately be spent on technology. The 50/50 allocation is a demonstration hypothesis: a disciplined starting point from which government can test how two distinct forms of capacity perform when funded together within a comparable envelope.

The underlying logic is straightforward: **One portion of the envelope purchases availability, continuity, accessibility, and scale, while the other purchases judgment, depth, safety, intervention, professional accountability, and human connection.**

Illustrative 50/50 Resource-Allocation Model

Approximate Allocation	Illustrative Uses	Primary Public Value Purchased
≈50% — AI-Enabled Public Digital Infrastructure	Model access; token and inference costs; computational capacity; secure hosting; technical infrastructure; multilingual capability; text and audio-enabled interaction; accessibility; privacy protections; cybersecurity; trauma-informed design; automated safety and boundary systems; monitoring; maintenance; testing; research; and independent technical evaluation.	**Availability, continuity, accessibility, multilingual reach, repeatability, and scale.**
≈50% — Human Professional Capacity and Governance	Therapists; counsellors; psychologists; social workers; other qualified professionals; supervisors; intervention and escalation specialists; clinical advisers; safety and quality reviewers; culturally responsive reviewers; service navigators; legal and ethical advisers; referral functions; AI supervision; professional-boundary oversight; and governance.	**Judgment, professional depth, relationships, complexity management, safety, intervention, accountability, and genuine human connection.**

The model can be expressed even more simply:

≈50% → create scalable availability

≈50% → protect scarcity-sensitive human expertise

The first allocation is not merely a technology budget. It purchases the infrastructure required to make an appropriate foundational support layer potentially available 24 hours a day, 365 days a year, across multiple languages and modes of interaction, without charging users for each appropriate use and without requiring an additional professional to become personally available for every additional interaction.

The second allocation protects the functions that technology should not replace. It purchases professional judgment, therapeutic and helping relationships, complex-care capacity, trauma-informed practice, risk assessment, escalation, intervention, supervision, quality assurance, ethical governance, and human accountability.

The proposed 50/50 structure is therefore not fundamentally a division between "technology" and "people." It is a division between two different kinds of public capacity. One purchases scalable availability. The other protects scarce expertise. The purpose of the demonstration is to determine whether combining the two can create more public value than either could create alone.

Why the Ratio Should Be Tested Rather Than Assumed

There is no reason to assume that 50/50 will prove to be the optimal long-term allocation. Indeed, one uniform ratio is unlikely to be appropriate across all populations, jurisdictions, service settings, risk profiles, languages, accessibility requirements, or forms of need.

Higher-risk environments may require substantially more professional involvement. Some populations may require more intensive human supervision or culturally specific human support. Certain languages may require additional quality assurance. Particular accessibility needs may require greater investment in interface design or human alternatives. Some low-complexity structured functions may eventually require proportionately less direct professional involvement. For that reason, the proposed ratio should be treated strictly as a testable starting point.

Future allocations, if the model advances, should be determined by evidence concerning factors such as computational and inference costs, model performance, multilingual quality, accessibility requirements, professional supervision ratios, privacy and cybersecurity obligations, regulatory requirements, jurisdictional differences, population needs, utilization patterns, escalation rates, user trust, professional workload, safety performance, workforce effects, and measurable outcomes.

The demonstration should answer concrete questions:

- Can a mixed funding model maintain strong professional employment and governance while substantially expanding public access and continuity?

- Can responsible automation of selected structured functions actually release professional time for higher-complexity needs?
- Can a substantially larger volume of appropriate foundational interactions be supported without professional labour having to rise in direct proportion to every additional interaction?
- Does the model maintain or improve user safety, privacy, trust, dignity, and access to professional care?
- Does it create additional workload elsewhere through supervision, escalation, technical governance, or administrative complexity?

And, ultimately:

- Can a comparable public funding envelope support a substantially broader and more continuous safety net than a predominantly episodic human-labour architecture alone?

The demonstration should not be designed to confirm that the answer is yes. Evidence should decide.

From Service Reimbursement to Hybrid Infrastructure Financing

Traditional mental-health financing frequently purchases discrete units or episodes of professional service. A funded appointment purchases a finite period of professional attention. Once that professional time has been used, the unit of capacity has been consumed.

Infrastructure financing operates differently. Infrastructure requires substantial upfront and ongoing investment, but once the underlying capability exists, it can often serve additional users without requiring an equivalent new unit of human labour for every additional use.

This does not mean digital infrastructure is free or infinitely scalable. A responsible public system would incur continuing costs for computing, model access, secure hosting, cybersecurity, privacy, accessibility, multilingual performance, testing, monitoring, human supervision, quality assurance, governance, evaluation, and technology renewal.

The important point is that the relationship between cost and additional use is different. Once a secure, multilingual, accessible, trauma-informed, human-supervised digital support infrastructure exists, an additional appropriate foundational interaction does not necessarily require another full unit of professional labour.

This creates the possibility of moving from an almost exclusively **service-reimbursement orientation** toward a hybrid **service-and-infrastructure financing model**. The distinction changes the fiscal question.

Under a predominantly episodic model, government may ask: How many counselling sessions can this funding envelope purchase?

Under a hybrid infrastructure model, government can ask: How much safe, meaningful, continuously available foundational support can this funding envelope sustain—and where should scarce professional expertise be concentrated within that system?

The first question is centred on purchasing finite units of service. The second concerns building and maintaining public capability. Both remain necessary. Professional services must continue to be funded where professional services are required. The innovation is that foundational availability need not always be financed exclusively through repeated one-to-one professional transactions. This shift could be consequential if Canada seeks to move from episodic access toward meaningful continuity.

The Economic Logic at a Glance

The difference between a predominantly human-labour model and the proposed hybrid architecture can be summarized as follows.

Predominantly Human-Labour Architecture	Proposed Human-AI Infrastructure Architecture
Additional supportive interactions generally require additional staff time.	Suitable structured interactions may increasingly be supported through scalable computational infrastructure.
Longer availability generally requires more staffing and additional shifts.	AI-enabled infrastructure can support continuous baseline availability, subject to technical capacity, safeguards, and human oversight.
Additional language access often requires additional direct human capacity.	AI may broaden baseline multilingual availability, while qualified humans remain essential for complex, culturally sensitive, or higher-risk situations.
Professionals may divide their time between routine and complex functions.	Selected routine functions may shift toward responsible AI assistance, protecting professional time for complexity and judgment.
Availability and expertise are largely purchased together through human labour.	**Availability and expertise can increasingly be financed through different but integrated forms of capacity.**
Expanding service volume generally requires proportionately greater labour expenditure.	Expansion of appropriate foundational support may not require professional labour to increase in direct proportion to every additional interaction.
Funding primarily purchases discrete episodes of service.	Funding can also purchase **continuously available public capability**.
Workforce shortages are addressed primarily through recruitment and training.	Workforce expansion can be combined with **better allocation of existing professional attention**.
Support may end when a funded episode or session allocation concludes.	A bounded foundational layer may remain available after one professional episode ends.

The policy choice is therefore not between artificial intelligence and therapists. It is between continuing to treat all forms of supportive capacity as if they have the same production economics, or recognizing that Canada now has access to two different but complementary resources: **computational capacity, which can create scalable availability;** and **professional human attention, which remains inherently scarce.**

A rational public-finance architecture should preserve both and allocate each to the functions where it creates the greatest public value.

Resource Reallocation as the Economic Foundation of Universal Continuity

The significance of this funding architecture extends beyond administrative efficiency. It may provide the economic foundation for a substantially larger public objective: **continuous access to an appropriate foundational safety net**.

The aspiration is straightforward. A person should be able to access an appropriate first layer of emotional support, guided reflection, non-clinical coaching, psychoeducation, life-skills assistance, or navigation without having to pay out of pocket for every use and without falling into complete absence of support simply because one funded program or professional encounter has ended.

Historically, such a promise has been constrained by labour economics. If every additional supportive interaction requires another paid professional hour, universal continuous availability becomes exceptionally difficult to finance. Government cannot employ an unlimited number of therapists. Counsellors cannot work continuously. No professional workforce can be simultaneously available across every community, language, accessibility need, time zone, and moment of emotional difficulty.

Artificial intelligence does not eliminate professional scarcity. It changes the economics of the support layer surrounding professional expertise. An appropriately governed system could potentially provide a bounded first-line resource continuously while a person waits for professional care, between appointments, while searching for an appropriate provider, after a short-term program concludes, or outside conventional service hours.

This is not unlimited psychotherapy. It is **continuous access to an appropriate foundational safety net**. That distinction is what may make the vision economically sustainable. Government would not need an unlimited supply of professional labour to ensure that every person always has some appropriate place to begin. It would need four complementary capacities:

- **scalable infrastructure for availability;**
- **sufficient professional human capacity for expertise, complexity, and intervention;**
- **qualified human supervision for safety and accountability; and**
- **public governance for standards, equity, and durability.**

Resource reallocation is therefore more than an accounting exercise. It is the mechanism that may allow continuity to become an economically credible public capability rather than a moral aspiration constrained by the limits of human labour.

From Per-Session Scarcity to a Continuity Guarantee

Traditional service architecture can require individuals to make repeated calculations.

Can I afford another session?

How many sessions do I have left?

What happens when the funded program ends?

What if the first provider is not right for me?

What if my next appointment is several weeks away?

What if I need support tonight?

What if I require another language?

What happens if a similar need returns months later?

These are partly clinical and service-design questions, but they are also economic questions because they arise from systems in which access is frequently organized around finite units of professional labour.

A more mature architecture should attempt to create a different baseline expectation: **There is somewhere appropriate and safe to which I can turn.** The professional system will continue to have limits. Some services will remain appointment-based. Professional care will continue to require finite human time. Specialized expertise will remain scarce. But a foundational safety net need not necessarily disappear every time one finite episode of professional care ends.

This is the economic basis of the proposed **Continuity Guarantee**: **No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.**

The complementary **Care Guarantee** addresses a different form of scarcity: **No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.**

These aspirations solve different economic problems. The Care Guarantee addresses the affordability of **expertise**. The Continuity Guarantee addresses the scarcity of **availability**. AI-enabled infrastructure may help make the second substantially more achievable without weakening the first. Indeed, if the model performs as intended, protecting foundational continuity could allow scarce professional resources to be directed more effectively toward fulfilling the Care Guarantee.

Matching Public Need to the Appropriate Form of Capacity

The economic architecture can also be understood by identifying the type of capacity best suited to address different public needs.

Public Need	Primary Capacity
Immediate baseline availability	AI-enabled infrastructure within defined boundaries
24/7/365 foundational continuity	AI-enabled infrastructure supported by human governance
Multilingual baseline access	AI-enabled infrastructure, subject to language-quality testing and human review
Text and audio-enabled interaction	AI-enabled infrastructure
Routine structured functions	AI-assisted or automated where evidence supports appropriateness
Guided reflection and non-clinical coaching	AI-enabled support within explicit scope boundaries
Basic service navigation	AI-enabled infrastructure with human escalation where necessary
Professional judgment	Qualified human professionals
Complex mental-health needs	Qualified human professionals
Therapeutic relationships	Qualified human professionals
Risk, ambiguity, and complex escalation	Qualified humans and relevant professionals
Safety and quality governance	Human oversight supported by technical monitoring
Ethical and professional accountability	Qualified humans, professional governance, and institutional accountability
Long-term durability and equitable access	Government and public institutions
Standards, evaluation, and public responsibility	Government in partnership with relevant jurisdictions, experts, and stakeholders

This produces a deliberately differentiated public-service architecture:

- AI provides availability.
- Professionals provide expertise.
- Human supervisors provide accountability.
- Government provides continuity, standards, and public responsibility.

The economic value lies not in maximizing any one form of capacity, but in aligning each form of capacity with the functions for which it has the greatest comparative advantage.

Fiscal Productivity, Not Crude Cost Reduction

The relevant fiscal question is not whether AI can make an individual conversation cheaper than a professional conversation. That comparison is too narrow to guide public policy.

A safe public system incurs costs that a simple per-interaction comparison can conceal: privacy, cybersecurity, human supervision, technical maintenance, multilingual quality assurance, accessibility, safety systems, professional consultation, independent evaluation, governance, escalation, referral, and infrastructure renewal.

The correct unit of analysis is therefore the complete human-AI service architecture. The central question should be: **Can the complete system produce greater access, continuity, safety, equity, and meaningful outcomes from a comparable funding envelope?** That is fiscal productivity.

Evaluation should therefore examine a broader set of measures, potentially including the cost per person meaningfully supported; cost per meaningful interaction; cost per successful navigation outcome; cost per person connected to appropriate professional care; proportion of professional hours devoted to structured versus complex functions; number of people meaningfully supported per professional hour; professional capacity released through responsible automation; infrastructure, computational, inference, supervision, and escalation costs; availability outside conventional operating hours; continuity while waiting for care; multilingual and accessibility performance; reduction in periods with no appropriate support; user trust and dignity; equity; safety outcomes; professional satisfaction; workforce effects; and total public expenditure relative to meaningful access and outcomes.

One metric deserves particular attention: **Where did the released human attention go?**

Suppose responsible automation reduces the professional time required for a particular standardized process by thirty minutes. A conventional efficiency calculation might value the saving primarily in labour-cost terms.

This proposal asks a more consequential question: **Was that thirty minutes of professional capacity redirected toward someone whose circumstances genuinely required professional expertise?**

If it was, the public value may extend far beyond the value of thirty minutes of wages. The system may have increased access to complex care. It may have reduced a wait. It may have allowed deeper therapeutic engagement. It may have improved supervision. It may have supported a higher-risk situation. It may have prevented professional expertise from being consumed by a function another form of capacity could safely provide. That is the kind of productivity gain this demonstration is designed to test. The objective is therefore not to maximize automation; it is to **maximize the public value created by finite resources**.

The Economic Test

The economic case for the proposed demonstration can ultimately be reduced to a single question. Within a defined or comparable public funding envelope: **Can Canada use computational capacity to**

create appropriate abundance in availability while protecting professional human attention for complex needs—and thereby generate substantially greater access, continuity, safety, equity, and public value?

The answer should not be presumed. If the model cannot demonstrate adequate safety, trust, equity, professional sustainability, fiscal value, or meaningful improvement in continuity, government should not scale it in its proposed form.

If, however, the evidence demonstrates that the architecture can meaningfully expand access without compromising professional standards or public trust, the implications would be significant. It would suggest that:

- improving mental-health accessibility does not always require professional labour to increase in direct proportion to every additional foundational interaction.
- public-system capacity can be expanded not only by adding professionals, but also by protecting and reallocating the attention of the professionals Canada already has.
- workforce shortages can be addressed through a combination of recruitment, retention, professional investment, and service redesign rather than through workforce expansion alone.
- public financing can move beyond purchasing isolated episodes of support toward maintaining a more durable continuum around professional care.

The core economic proposition can therefore be stated simply:

- Do not use scarce professional human attention where responsibly governed scalable capacity is sufficient.
- Do not use technology where professional human expertise is necessary.
- Fund each according to its comparative advantage.
- Evaluate the performance of the system as a whole.

This is not a proposal to reduce public responsibility. Rather, it is a proposal to exercise public responsibility more intelligently. The ultimate objective is not merely lower expenditure. It is to determine whether finite public resources can support something that has historically been difficult to make financially sustainable: a free-at-the-point-of-appropriate-use, continuously available, multilingual, accessible, text- and audio-enabled foundational mental-health and emotional-support safety net—backed by qualified humans, connected to professional care, governed in the public interest, and designed so that people are not left without an appropriate place to turn when professional care is temporarily unavailable.

Technology creates the possibility. Resource reallocation creates the economic pathway. Professional and human governance create the safeguards. Government creates the durability. The 50/50 demonstration is the mechanism for determining whether those elements can be combined into a credible, safe, and fiscally sustainable public capability.

7. Responsible AI, Privacy, Accessibility, and Trust

Trust, Safety, Privacy, and Human Dignity

Trauma-Informed AI Requires More Than Technical Accuracy

A trauma-informed virtual safe space cannot be evaluated solely on whether its outputs are factually accurate, linguistically fluent, technically coherent, or consistent with predefined safety rules. Those standards are necessary, but they are not sufficient.

The more important question is: **How might this interaction affect a person who has experienced trauma, rejection, coercion, abuse, discrimination, marginalization, fear, social exclusion, or institutional betrayal?**

A response can be technically plausible and still leave a person feeling dismissed, controlled, misunderstood, pathologized, exposed, blamed, or unsafe. In emotionally sensitive public infrastructure, those relational effects matter.

Trauma-informed design should therefore be treated as a core governance requirement rather than an optional feature of conversational style. The system should be designed around psychological safety, dignity, autonomy, informed choice, transparency, non-judgment, appropriate boundaries, privacy, non-exploitation, cultural sensitivity, avoidance of unnecessary re-traumatization, proportionate escalation, and accountable human oversight.

This is particularly important because some users may approach the service after difficult experiences with institutions, healthcare systems, workplaces, family systems, relationships, or other forms of authority. A public-facing system should not reproduce avoidable patterns of opacity, coercion, dependency, or loss of control.

Users should understand clearly what the system is, what it is not, what forms of support it can appropriately provide, what functions remain outside its scope, when professional care may be more appropriate, when human review may occur, what safety boundaries apply, and what alternative services remain available.

Artificial intelligence should never falsely represent itself as a therapist, psychologist, physician, social worker, counsellor, or other regulated professional when it is not one. It should not independently diagnose or prescribe outside lawful authority. It should not imply that psychotherapy, medical care, crisis intervention, or emergency services are unnecessary. And it should not encourage a person to substitute AI for professional care when professional care is clearly warranted.

The objective is not to create the appearance of professional authority. It is to create a **bounded, transparent, supportive environment that understands both its role and its limits**.

The more continuously available the system becomes, the more important those boundaries become. A service that may eventually operate across jurisdictions, languages, and large populations must be designed so that scale does not amplify ambiguity, dependence, inequity, or harm.

The governing principle should therefore be: **The more scalable the technology, the stronger the human governance around it.**

Psychological Safety Must Be Designed Into the Interaction

Psychological safety does not arise automatically because a service is free, accessible, or compassionate in intention. It must be experienced by the person using it.

A psychologically safer digital environment should allow people to begin without fear of judgment, ask questions without being shamed, express uncertainty without being pressured, decline suggestions, question the system's responses, understand its limitations, disengage at any time, request human support where available, and retain control over their own decisions.

The system should not encourage unhealthy dependency. It should not suggest that it is the only place capable of understanding the user. It should not create artificial exclusivity, pressure people to remain engaged, or blur the distinction between supportive interaction and a genuine reciprocal human relationship. This is especially important in emotionally vulnerable contexts, where an always-available system may feel unusually responsive or consistent. Availability itself can create attachment dynamics. Responsible design should therefore avoid features or language that intentionally cultivate dependence or imply a level of relational reciprocity that the technology cannot genuinely provide.

The goal is to make support more available **without making the user less autonomous**.

A well-designed system should leave the person with greater clarity, greater agency, more options, stronger understanding of appropriate next steps, and a better ability to reach human care where needed. It should not make people more isolated from the human world around them.

Privacy Must Be Treated as Infrastructure

A virtual safe space cannot become credible public infrastructure without trust, and trust cannot be sustained if people believe their most sensitive emotional disclosures may be casually retained, exposed, commercialized, repurposed, or monitored. Privacy should therefore not be treated as a legal notice positioned at the edge of the service. It should be treated as **core infrastructure**.

The demonstration should incorporate privacy by design, data minimization, meaningful and understandable consent, clear purpose specification, appropriate retention limits, strong cybersecurity, encryption where appropriate, role-based access, transparent disclosures, strict limits on secondary use, independent privacy auditing, clear rules for human review, complaint and correction mechanisms, breach-response procedures, and accountable governance.

Users should be able to understand in plain language what information is collected, why it is collected, what information is not necessary, who may access it, under what circumstances human review may occur, how long information is retained, whether information may be used to improve the system, what rights the user retains, and what human or non-digital alternatives remain available.

The privacy model must also reflect the sensitivity of the context. A conversation about abuse, grief, trauma, family conflict, fear, loneliness, identity, relationships, or emotional distress is not equivalent to an ordinary consumer transaction. Emotionally sensitive disclosures deserve a heightened standard of restraint.

The fact that information may be useful does not mean it should automatically be collected. The fact that it can be retained does not mean it should be retained indefinitely. The fact that AI can analyze sensitive material at unprecedented scale makes disciplined restraint more important, not less.

The service should communicate a simple principle through its architecture: **People should be able to seek support without surrendering unnecessary control over their private lives.**

AI Must Not Become a Tool of Individualized Surveillance

AI-enabled infrastructure may eventually generate useful knowledge about recurring barriers within the mental-health and support system. That potential should be explored carefully. It should also be subject to strict limits.

Private emotional conversations should not become individualized intelligence for government. Identifiable histories should not be distributed to decision-makers for generalized policy analysis. Private disclosures should not be converted into behavioural or psychological profiles unrelated to the immediate purpose of the service. They should not be sold, used for advertising, deployed for political targeting, or repurposed for unrelated administrative objectives without appropriate legal authority and meaningful authorization.

A government-backed virtual safe space must never evolve into an infrastructure for quietly monitoring the emotional lives of citizens. Where system learning is undertaken, it should rely on methods such as appropriate de-identification, aggregation, data minimization, purpose limitation, access controls, privacy review, independent oversight, and meaningful consent where required.

For example, aggregate analysis may help identify persistent waitlist problems, underserved languages, regional gaps, navigation failures, repeated affordability barriers, or recurring difficulty connecting people with appropriate professional care. Those insights could legitimately improve public policy, but the object of analysis must remain structural.

The governing principle should be: **The purpose is not watching people at scale. It is understanding systemic barriers at scale.** That distinction should be reflected not only in policy statements, but in technical architecture, law, procurement requirements, institutional culture, audit mechanisms, and accountability structures.

Frontline Learning Must Not Come at the Cost of Trust

One potential advantage of AI-enabled public infrastructure is its ability to detect recurring frontline patterns more quickly than conventional reporting systems. Today, lived experience often travels through several institutional layers. A person speaks to a frontline worker. The interaction becomes documentation. Documentation is summarized. Summaries become reports. Reports become performance indicators. Leadership eventually receives an aggregate number. At each stage, important context may be lost.

AI may make it possible to identify recurring system-level patterns earlier. That could be useful, but this capability must remain subordinate to the primary purpose of the service: **Supporting the individual seeking help**.

The user should never feel that emotional vulnerability is being converted into an unrestricted data-extraction exercise. Support must remain primary. System learning must remain secondary, proportionate, privacy-preserving, and purpose-limited.

A virtual safe space cannot plausibly ask people to trust it while simultaneously treating their private emotional lives as a broad institutional intelligence resource. Any frontline-learning function should therefore be governed by a principle of minimum necessary use: collect and analyze only what is needed to answer legitimate public-interest questions, and do so at the lowest level of identifiability compatible with that purpose.

Accessibility Must Be Built In From the Beginning

A national public-interest service should not be designed for an imaginary "average user." It should be designed around the diversity of people who may need it. That includes differences in language, disability, literacy, neurodivergence, culture, age, geography, technological access, communication preference, digital confidence, and familiarity with mental-health systems.

The demonstration should therefore incorporate accessibility as an architectural requirement from the outset. Where technically feasible and safe, this should include multilingual interaction, text-based access, audio-enabled interaction, screen-reader compatibility, simplified-language options, accessible navigation, adaptable interfaces, low-bandwidth functionality, captioning or transcription where appropriate, clear visual design, and transparent alternatives for people who cannot or do not wish to use AI.

This is directly connected to the larger policy objective. If the long-term vision is a continuously available foundational safety net, access cannot be meaningful only for people who speak fluent English or French, have high-speed internet, are comfortable composing long written messages, have no visual or cognitive accessibility needs, understand mental-health terminology, and already know how to navigate digital systems.

Universal availability without usable accessibility is not universal access. Accessibility should therefore be measured as an outcome, not simply confirmed through a compliance checklist. The demonstration should ask whether people can actually use the service, understand what it is telling them,

communicate in a mode that works for them, locate human assistance, understand the distinction between AI and professional care, and navigate the system without unnecessary confusion. These are not merely technical questions. They express an important public value: *You belong here.*

Multilingual Access Must Mean More Than Translation

Multilingual capability may become one of the most important advantages of AI-enabled public infrastructure. It has the potential to address a barrier that human-only service systems find especially difficult to solve at population scale. However, multilingual access should not be confused with literal translation. A system can produce grammatically correct language and still misunderstand the person using it.

Language operates within culture. People make sense of distress, responsibility, autonomy, family, privacy, shame, spirituality, identity, caregiving, authority, obligation, and help-seeking through cultural frameworks. A response developed for one cultural environment may therefore be linguistically accurate yet relationally inappropriate in another.

The demonstration should assess multilingual performance across several dimensions: linguistic accuracy, tone, cultural appropriateness, forms of address, family and relational context, assumptions about autonomy, stigma, spiritual or religious context where relevant, communication norms, and whether users themselves feel understood.

The relevant question is not only: Does the system speak my language? It is: **Does the system communicate with me in a way that respects how I understand my life?** Multilingual capability should therefore be treated as both an accessibility function and a quality-and-safety function.

Cultural Responsiveness Is a Safety Requirement

Cultural responsiveness should not be treated as an ancillary diversity objective. In emotionally sensitive services, it is a safety issue. Culturally inappropriate guidance can create confusion, reinforce mistrust, increase shame, discourage further help-seeking, or misinterpret dynamics that are central to a person's circumstances. This is especially important in conversations involving family expectations, intergenerational responsibility, collectivist and individualist values, religious or spiritual beliefs, migration, racialization, gender roles, caregiving, community obligation, stigma, and different understandings of mental health itself.

The demonstration should therefore incorporate community participation, culturally diverse professional input, bias review, multilingual evaluation, lived-experience participation, documentation of system limitations, and ongoing assessment of whether different groups experience the service as respectful, useful, dignified, and appropriate.

Cultural responsiveness should not mean uncritically validating every cultural norm or practice. Respecting cultural context is different from reinforcing coercion, abuse, discrimination, or harm. A responsible system should seek to understand context before making assumptions, while maintaining

clear safety and human-rights boundaries. This is one reason qualified human governance remains essential.

Relational Safety Matters Even in an AI-Enabled System

Although the proposed virtual safe space is technologically mediated, people will experience it relationally. Users respond not only to factual content, but also to tone, timing, language, perceived empathy, consistency, respect, and whether they feel heard. Relational safety should therefore be assessed alongside technical performance.

The system should avoid dismissive language, unnecessary confrontation, excessive certainty, moralizing, shaming, patronizing responses, artificial intimacy, dependency-promoting language, and claims to emotional experience or relationship that the system does not genuinely possess.

Compassionate design does not require deception. An AI system can communicate warmly while remaining transparent about what it is. It can acknowledge emotion without claiming to experience emotion itself. It can be supportive while maintaining boundaries. It can encourage connection while making clear that it is not a substitute for human relationships or professional care.

The objective is not to imitate a human being perfectly. It is to create an interaction that is **safe enough, respectful enough, transparent enough, and useful enough to help the person take the next constructive step**. That distinction is especially important for public infrastructure because the state should not rely on anthropomorphic design to manufacture trust. Trust should arise from appropriate conduct, clear boundaries, reliable safeguards, and accountability.

Human Review Must Be Proportionate, Purpose-Limited, and Transparent

Human oversight is essential to the proposed architecture, but human oversight itself creates privacy and governance questions. The demonstration should therefore establish explicit rules governing when human review can occur, what information reviewers may access, why access is permitted, whether review is triggered automatically or manually, what qualifications reviewers must possess, how access is logged, how long reviewed information is retained, and how inappropriate access is detected and addressed.

Users should not simply be told that conversations "may be reviewed." They should receive understandable information about the circumstances in which review may occur and the purposes for which it is permitted. At the same time, privacy should not be invoked as a reason to eliminate meaningful human accountability. A system in which no qualified person can ever examine problematic behaviour may be private in one sense while unsafe in another.

The correct objective is therefore neither unrestricted monitoring nor zero human review. It is: **Proportionate, purpose-limited, transparent human review under strong privacy safeguards.** This approach should be complemented by auditing, role-based permissions, minimum-necessary access, and meaningful consequences for misuse.

Safety Escalation Should Operate as a Continuum

Escalation should not be designed as a binary switch between "AI conversation" and "emergency services." Human needs exist on a continuum, and the support architecture should reflect that reality. Different circumstances may require different interventions. Some may call for a more cautious automated response. Others may require additional information or navigation. Some may warrant involvement from a trained human navigator, counsellor, or professional. Higher-risk situations may require crisis or emergency resources.

The demonstration should therefore test **graduated escalation pathways**. A responsible system should be able to recognize when its current mode of support is no longer appropriate and when the level of human involvement should increase. This is not evidence that the system has failed; it is evidence that the system understands its boundaries.

A well-designed safety architecture should know:

- when to continue;
- when to slow down;
- when to involve a human; and
- when to direct the person toward a higher level of care.

Escalation quality should also be evaluated in both directions. Missed escalation can create serious risk, but excessive or inappropriate escalation can undermine trust, overwhelm human services, and make people reluctant to disclose honestly. The demonstration must therefore assess both under-escalation and over-escalation.

Trust Must Be Measured as an Outcome

A technically sophisticated and fiscally productive system can still fail if people do not trust it enough to use it appropriately. Trust should therefore be treated as a measurable policy outcome rather than a branding objective.

The demonstration should examine psychological safety, dignity, perceived respect, usefulness, understanding of the respective roles of AI and humans, clarity about system limitations, confidence in privacy protections, willingness to return, willingness to recommend the service, cultural responsiveness, accessibility, comfort with escalation processes, perceived autonomy, and whether users feel more or less connected to appropriate human support after using the service.

One measure is particularly important: **Does the virtual safe space make people more connected to appropriate human care—or more distant from it?**

The model succeeds only if AI functions as a bridge rather than a barrier. It should support movement toward clarity, professional help where required, community connection, self-understanding, agency, and safer navigation of the broader support system. It should not create an isolated digital

ecosystem that gradually substitutes itself for real-world relationships, professional care, or community life. Trust should therefore be evaluated together with autonomy and connection.

Public Infrastructure Requires a Higher Trust Standard

If the virtual safe space ultimately becomes part of publicly supported infrastructure, it should be held to a higher standard than an ordinary consumer application. The public should be able to expect that emotionally sensitive conversations will not be exploited for advertising, that disclosures will not be sold, that the system will not become a mechanism for individualized government surveillance, that AI will not misrepresent itself as a professional, and that appropriate professional care will not be discouraged when it is needed.

People should also be able to expect meaningful human accountability, understandable privacy rules, accessible complaint and correction mechanisms, serious attention to disability and cultural responsiveness, and continued independent evaluation after deployment.

A government-supported emotional-support system should not ask citizens to trust it simply because government endorses it. It should earn trust through institutional design. The appropriate message is not: Trust us. It is: **Here are the boundaries. Here are the safeguards. Here is what information is used and why. Here is who is accountable. Here is how human review works. Here is what happens when something goes wrong. Here are your alternatives, and here is how the system is independently evaluated.** That is the standard appropriate for public infrastructure.

Safety and Accessibility Are the Same Investment

It would be a mistake to treat safety as a constraint on accessibility. In a public virtual safe space, the two are mutually reinforcing. A service is not meaningfully accessible if people do not trust it. It is not accessible if users fear surveillance. It is not accessible if they cannot understand the language, navigate the interface, or determine whether they are interacting with AI or a professional. It is not accessible if cultural assumptions make the service alienating, if disability prevents effective use, or if people believe that seeking support may expose them to unnecessary harm.

For this reason: **Safe AI and accessible support are the same investment.**

- Privacy creates access.
- Trust creates access.
- Cultural responsiveness creates access.
- Disability inclusion creates access.
- Clear boundaries create access.
- Human accountability creates access.

A virtual safe space becomes genuinely universal only when people can approach it without feeling that they must exchange dignity, autonomy, privacy, cultural identity, or safety for the ability to receive support.

The Trust Architecture at a Glance

Public Requirement	Design and Governance Response	Intended Public Value
Psychological safety	Trauma-informed interaction, autonomy, clear boundaries, non-judgment, transparent limitations, dependency safeguards	**Dignity and emotional safety**
Privacy	Data minimization, understandable consent, purpose limitation, retention controls, cybersecurity, access controls	**Confidentiality and trust**
Protection from surveillance	No individualized government intelligence, strict secondary-use limits, privacy-preserving aggregate analysis where appropriate	**Freedom from institutional overreach**
Accessibility	Multilingual, text and audio access, disability-inclusive design, low-bandwidth functionality, simplified navigation, non-AI alternatives	**Usable access across diverse populations**
Cultural responsiveness	Community participation, culturally diverse professional review, contextual evaluation, multilingual testing, bias review	**Relational and cultural safety**
Clear professional boundaries	Transparent AI identity, no unauthorized diagnosis or prescribing, explicit scope limitations and escalation pathways	**Informed and appropriate use**
Human accountability	Qualified supervision, proportionate review, escalation, quality assurance, complaints and redress mechanisms	**Safety and institutional responsibility**
Independent evaluation	External review of safety, privacy, bias, accessibility, trust, outcomes, and unintended consequences	**Public legitimacy**
Connection to human care	Navigation, graduated escalation, referral, professional partnerships, continuity while waiting	**AI as a bridge, not a destination**

A virtual safe space cannot become legitimate public infrastructure merely because it is technologically capable or continuously available. It becomes legitimate only if it is also:

- safe,
- private,
- accessible,
- culturally responsive,
- professionally governed,
- transparent,
- accountable, and
- worthy of public trust.

The long-term vision of free, continuous, multilingual, text- and audio-enabled foundational support depends on that trust.

If Canada ultimately wishes to offer people the assurance that **there is somewhere appropriate to turn and that their country has their back**, then the system must demonstrate through its design that having people's backs also means protecting **their dignity, autonomy, privacy, culture, safety, and right to appropriate human care**. That is not an ancillary safeguard around the proposed model. It is the trust architecture on which the entire model depends.

8. Governance, Legal Clarity, and Public Responsibility

A Multi-Stakeholder Public-Interest Governance Model

A national **Human-Supervised Virtual Safe Space Demonstration Initiative** should be governed through a multi-stakeholder framework designed explicitly around the public interest.

No single department, profession, nonprofit organization, technology company, AI vendor, or regulatory body possesses all of the expertise required to govern a system operating at the intersection of mental health, artificial intelligence, privacy, accessibility, digital public infrastructure, workforce transformation, public administration, social innovation, and professional regulation.

Depending on the final jurisdictional and implementation design, appropriate participation could include relevant federal departments and agencies, provincial and territorial governments, mental-health professional organizations and regulatory bodies, universities and research institutions, nonprofit and community organizations, people with lived experience, Indigenous partners where appropriate and consistent with Indigenous rights and governance, accessibility and disability experts, privacy and cybersecurity specialists, responsible-AI and safety researchers, legal and regulatory experts, cultural and linguistic specialists, social innovators, and Canadian technology organizations.

The purpose of such breadth would not be to create governance by committee for its own sake. It would be to ensure that decisions with clinical, ethical, technological, legal, cultural, fiscal, and social consequences are not made from a single institutional perspective.

Government and its partners would need to determine, among other matters, which functions AI may appropriately perform; which functions require qualified human oversight; where foundational non-clinical support ends and regulated professional care begins; what supervision standards are necessary; how escalation should operate; what privacy and data-governance safeguards should apply; what accessibility obligations should be built into the service; how cultural responsiveness should be assessed; how safety should be independently evaluated; how accountability should be distributed across service providers; and how responsibility should operate across jurisdictions and contractual relationships.

The governing principle should therefore be: **Government defines the public-interest standard; no single vendor defines it.**

Different participants will contribute different forms of expertise. A technology provider may supply infrastructure. A nonprofit organization may contribute frontline knowledge. Professional bodies may help define scope and practice boundaries. Researchers may evaluate outcomes. People with lived experience may expose barriers and unintended consequences. Privacy and accessibility specialists may identify risks that others overlook.

However, the underlying capability, public rights, safeguards, and accountability framework should remain larger than any individual participant. The objective is not to create permanent dependence on a particular platform. It is to establish a **technology-neutral public standard for continuously available, human-supervised foundational support**.

A Legal and Regulatory Blind Spot Around Emerging Social Innovation

The governance challenge extends beyond artificial intelligence. It also reflects a broader difficulty facing social innovation: new public-interest models can develop more quickly than the legal and regulatory categories through which conventional services are governed.

In my Forbes Nonprofit Council article, *Why Culturally Sensitive Social Entrepreneurship Demands a Legal Rethink*, I argued that socially innovative models can sit across boundaries that existing legal and professional systems understandably treat as distinct. A single initiative may combine elements of education, coaching, peer support, community care, mental-health support, technology, cultural practice, navigation, social services, and professional services. At the same time, legal and regulatory frameworks rely on important distinctions between clinical and non-clinical activity, regulated and unregulated practice, professional and peer support, medical and non-medical services, commercial and charitable activity, and technology provision and human-service delivery.

Those distinctions serve legitimate public-protection purposes. The challenge arises when a socially useful innovation spans several of them simultaneously. This becomes particularly complex in culturally responsive social entrepreneurship, where trust, relational norms, spirituality, family structures, communication practices, community expectations, and cultural understandings of support may influence how people experience and interpret a service.

An innovative model may therefore generate meaningful public value while also creating uncertainty concerning professional scope, supervision, duty of care, consent, privacy, data governance, cross-jurisdictional responsibility, AI accountability, and liability.

The appropriate response is not to exempt social entrepreneurs or technology-enabled services from accountability. Public-interest motivation should never create immunity from professional, legal, ethical, privacy, or safety obligations. The stronger response is to build **clearer, more proportionate, and more coherent governance frameworks for emerging forms of socially beneficial service delivery**.

Good regulation should protect the public while also making responsible experimentation possible. Innovators should not be forced to navigate major areas of legal uncertainty individually when the underlying policy questions are systemic.

Government Can Distribute Responsibility More Coherently

Government participation should not eliminate accountability. Nor should it. Its value lies in distributing responsibility more coherently across the institutions best positioned to carry it.

At present, an organization developing an AI-enabled emotional-support service may confront a wide range of difficult questions with limited authoritative guidance. It may need to determine what forms of AI-assisted emotional support are appropriate, where coaching ends and regulated care begins, what disclosures users should receive, when qualified human involvement becomes necessary, what level of supervision is appropriate, what safety signals justify escalation, what data may be retained, how consent should operate, when conversations may be reviewed, how culturally responsive practice should be governed, which professional standards apply across jurisdictions, and how responsibility should be divided when an AI system produces an inappropriate or harmful output.

These are not questions that should be answered independently and repeatedly by every nonprofit, startup, professional practice, or technology company entering the field. Such fragmentation can produce duplicated legal costs, inconsistent interpretations, regulatory uncertainty, uneven safety practices, barriers to responsible innovation, and potentially unequal protection for the public.

Government can help create a common framework. A national or intergovernmental demonstration could test and clarify areas such as permitted and prohibited functions, professional-boundary standards, minimum supervision requirements, privacy and cybersecurity expectations, consent, retention, human-review rules, escalation, accessibility, cultural responsiveness, complaint and redress mechanisms, independent evaluation, quality assurance, and the appropriate allocation of responsibility across organizations and professional roles.

Clearer standards can protect users and innovators at the same time. For users, they can establish minimum protections, transparent accountability, professional oversight, and meaningful complaint pathways. For responsible innovators, they can provide clearer operating boundaries, more predictable obligations, and less need to navigate foundational governance questions from first principles.

The objective is not deregulation. It is **better governance through clearer allocation of responsibility**.

Social Entrepreneurs Should Be Catalysts, Not Permanent Shock Absorbers

Social entrepreneurs can play an unusually valuable role in identifying needs before large institutions are positioned to respond to them. They can remain close to frontline experience. They can listen. They can experiment. They can challenge established assumptions. They can build prototypes, test unconventional models, and demonstrate possibilities that conventional systems have not yet explored. That capacity for experimentation is one of social entrepreneurship's principal contributions to the public interest. However, innovation and permanent infrastructure are different responsibilities.

A social entrepreneur should be able to identify an unmet need, test a solution, generate evidence, and demonstrate whether a new capability may have public value. They should not automatically become personally responsible for maintaining essential infrastructure indefinitely.

This distinction becomes especially important when services involve emotionally vulnerable populations or sensitive personal data. A small nonprofit or one or two founders may quickly become responsible for privacy, cybersecurity, data governance, crisis protocols, professional boundaries, AI safety, clinical or trainee supervision, volunteer management, accessibility, insurance, regulatory interpretation, quality assurance, technical risk, and substantial potential legal exposure.

As the service grows, those obligations can grow with it. This produces a difficult paradox: **The more public value an innovation creates, the greater the private burden the innovator may be required to carry.** That is not a sustainable architecture for essential public capability.

The concern becomes particularly acute where a social entrepreneur is not extracting substantial private profit from the service, but instead contributes considerable labour, expertise, financial resources, and personal effort in pursuit of a public-interest purpose.

A durable public system should not depend on indefinite founder sacrifice. Canada should encourage people willing to experiment responsibly in service of the public good. It should not build essential social capacity on the assumption that those individuals will absorb unlimited financial, operational, emotional, ethical, technological, and legal exposure forever.

The principle should therefore be: **Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.**

Public-Interest Innovation Should Not Depend on Indefinite Personal Sacrifice

There is a broader institutional principle behind this argument. Social innovation often begins precisely because established systems have not yet solved a persistent problem. An innovator enters the gap, contributes time and resources, experiments under uncertainty, develops a prototype, and frequently accepts financial, reputational, operational, and legal risk before society has determined whether the model deserves broader adoption.

That willingness to experiment can create substantial public value. Public policy, however, should not romanticize founder resilience to the point that institutional responsibility disappears. Compassion can motivate innovation. It should not require **indefinite personal sacrifice as the operating model for a public good**.

An infrastructure that depends permanently on one founder’s finances, health, stamina, willingness to accept legal exposure, or capacity to keep working indefinitely contains an obvious point of fragility. If a socially beneficial innovation addresses a widespread, persistent, and important public need, the next policy question should eventually become: **How can responsibility move from individual sacrifice toward durable institutional arrangements?**

That is the point at which social entrepreneurship becomes not only an innovation question, but a governance question. The purpose is not to transfer every private initiative automatically to government. Most innovations should remain private, charitable, community-based, or commercial according to their purpose and evidence. The principle applies where a demonstrated capability becomes sufficiently important, enduring, and broadly beneficial that depending indefinitely on personal sacrifice creates unacceptable risk for both the innovator and the public.

Public Need Should Create Public Responsibility

FASS and FASSLING™ can identify unmet needs, test models, demonstrate possibilities, generate frontline learning, and help reveal what emerging technologies may make possible. They should not, however, be understood as organizations that must carry permanent responsibility for resolving a systemic problem alone.

If rigorous evidence eventually demonstrates that Canadians benefit from access to a free-at-the-point-of-appropriate-use, continuous, multilingual, trauma-informed, accessible, human-supervised emotional-support safety net, government and society would then need to confront a larger institutional question: **Has this remained primarily a charitable innovation, or has the underlying capability become a legitimate public need?**

If the latter is demonstrated, responsibility should evolve accordingly. At some point, successful experimentation can mature into infrastructure. A socially necessary capability should not remain indefinitely dependent on one founder, one nonprofit organization, one donor, one grant, one volunteer network, or one technology company.

The broader principle is: **Public need should create public responsibility.** That principle does not require government to operate every component directly.

Government may choose among a range of models: public operation, contracting, partnership, regulation, certification, common standards, targeted funding, interoperable service networks, or combinations of these approaches. What matters is that government take responsibility for ensuring that the **public capability itself** is adequately governed and capable of enduring. If the need is persistent, the public should not lose the underlying support function solely because a particular founder, nonprofit, vendor, or funding arrangement disappears.

The Continuity Guarantee Requires Institutional Continuity

This principle becomes especially important because the larger ambition of the proposal is not the creation of another temporary program. It is the possible foundation of a **Continuity Guarantee**: **No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.**

A guarantee cannot be meaningful if the infrastructure supporting it can disappear whenever a founder burns out, a grant expires, a company changes strategy, a donor withdraws, a vendor fails, or a technology becomes obsolete.

A credible public guarantee requires **institutional continuity**. The technology may change. The AI model may change. The contractor may change. The nonprofit partner may change. The professional governance model may evolve as evidence, regulation, and technology develop. But if the underlying capability proves valuable, the public commitment should not disappear with those changes.

The enduring promise is not that one particular system will exist forever. It is: **There will continue to be somewhere appropriate and safe to turn.** That distinction transforms continuity from a product feature into a governance obligation.

Reducing Founder and Vendor Dependency Is Public Risk Management

Reducing dependence on individual founders is therefore not merely a question of fairness to social entrepreneurs. It is also a matter of **public risk management**.

Essential infrastructure should avoid unnecessary single points of failure. A service supporting large numbers of people becomes institutionally fragile if its continued existence ultimately depends on one individual, one small organization, one grant, one proprietary technology, one commercial vendor, or one funding relationship.

Founders may leave. Programs may end. Companies may fail. Technologies become obsolete. Procurement arrangements change. Funding priorities shift. Leadership turns over. A durable public capability must be able to survive those transitions.

Government, therefore, need not guarantee the continued existence of FASSLING™, a particular AI model, a particular nonprofit organization, or a particular technology provider. What it can protect is the underlying **technology-neutral capability**: **Free-at-the-point-of-appropriate-use, continuously available, multilingual, trauma-informed, accessible, human-supervised emotional-support infrastructure connected to appropriate professional care.**

This is the difference between funding a product and safeguarding a public capability. A product can be replaced. A public commitment should be designed to endure.

Government Should Protect the Promise, Not the Platform

Technology neutrality is particularly important in artificial intelligence because the technical landscape is evolving rapidly. It would be unwise to define a long-term national policy architecture around one commercial model, one AI company, one nonprofit platform, one proprietary system, or one vendor. Model capabilities will change. Costs will change. Safety methods will evolve. New providers will emerge. Existing providers may decline or disappear. Standards that appear sufficient today may prove inadequate tomorrow.

Government policy should therefore define the public capability and its requirements, not permanently enshrine the technology currently used to provide it. Those requirements should include public rights, scope, professional boundaries, privacy standards, accessibility obligations, human-supervision requirements, safety expectations, escalation pathways, interoperability where appropriate, evaluation criteria, transparency, procurement requirements, and accountability.

Technology providers should operate within those standards rather than define them. The governing principle is: **Government should protect the promise, not the platform.**

The promise is continuity, safety, equitable access, human accountability, and appropriate connection to professional care. The platform is merely one means of delivering it.

Legal Clarity Should Protect Responsible Social Innovation

Government participation could also reduce the disproportionate legal uncertainty faced by organizations experimenting responsibly in emerging areas. Social entrepreneurs, nonprofits, professionals, and technology providers should remain accountable for genuine wrongdoing or failure, including negligence where applicable, privacy breaches, misrepresentation, unsafe practices, professional-boundary violations, and other legally cognizable harms.

Public-interest intent should never become a substitute for accountability. At the same time, accountability is more effective when obligations are clear.

Innovators should not be expected to operate indefinitely in uncertainty about where emerging service categories fit, which legal and professional standards apply, what supervision is required, how cross-jurisdictional responsibility should operate, or how AI-specific responsibility should be divided among developers, service organizations, professional supervisors, contractors, and infrastructure providers.

Clear governance can produce two forms of protection simultaneously. For the public, it can establish minimum standards, defined accountability, complaint and redress mechanisms, professional oversight, and greater consistency. For responsible innovators, it can create clearer boundaries, more predictable obligations, reduced regulatory ambiguity, and a safer environment for experimentation.

The objective should therefore be neither blanket immunity nor disproportionate private exposure. It should be **proportionate accountability within a clear public framework.**

Any specific legal allocation of liability would require detailed analysis under applicable federal, provincial, territorial, professional, contractual, privacy, and common-law frameworks. The demonstration should help generate the evidence and governance experience needed to support that work.

Social Entrepreneurs Discover; Government Institutionalizes When Evidence Justifies It

The strengths of social entrepreneurs and governments are different. They are also complementary. Social entrepreneurs are often particularly effective at remaining close to frontline reality, identifying unmet needs early, listening to people who fall through institutional gaps, experimenting, moving rapidly, questioning inherited assumptions, and testing alternative forms of service delivery.

Government possesses a different set of capacities. It can establish standards, coordinate jurisdictions, protect equitable access, fund durable infrastructure, define accountability, support workforce transition, create regulatory frameworks, require independent evaluation, distribute responsibility across institutions, and maintain continuity over time.

The relationship should therefore not be framed as social entrepreneur versus government. A healthier model is sequential and complementary.

- Social entrepreneurs can discover.
- Evidence can determine whether the discovery has wider public value.
- Government can institutionalize the underlying capability when that evidence justifies doing so.

Institutionalization need not mean nationalization. It can take many forms. The essential point is that once a capability becomes part of a public commitment, its durability should no longer depend primarily on the persistence of the individual who first demonstrated it.

The Ideal Public-Innovation Cycle

The broader innovation pathway envisioned by this proposal can be understood as an eight-stage cycle.

First, frontline communities reveal a need. People living with a problem identify gaps that may not yet be visible through aggregate administrative data or established institutional categories.

Second, social entrepreneurs listen and experiment. They remain close to the problem, test assumptions, build prototypes, and explore alternative models.

Third, evidence begins to emerge. Experience, user feedback, operational data, safety information, and early outcomes begin to show where the innovation is useful, where it fails, and what would be required for responsible improvement.

Fourth, government evaluates. Public institutions assess safety, effectiveness, equity, accessibility, professional implications, workforce consequences, fiscal value, privacy, legal requirements, and broader social benefit.

Fifth, public standards are established. Governments, professionals, communities, people with lived experience, researchers, legal experts, accessibility specialists, and other relevant stakeholders define boundaries, rights, safeguards, obligations, and accountability.

Sixth, successful capabilities become durable infrastructure. Where evidence supports broader implementation, government helps ensure that the underlying capability is sustainably financed, governed, interoperable where appropriate, accessible, and continuously evaluated.

Seventh, responsibility becomes institutional rather than primarily personal. The public capability no longer depends disproportionately on the endurance, finances, or legal exposure of one founder or small organization.

Eighth, social entrepreneurs are freed to innovate again. Rather than spending the remainder of their careers personally maintaining one systemic solution, innovators can return to identifying new unmet needs and begin the cycle again.

This is a healthier division of institutional roles. The social entrepreneur remains a source of experimentation and frontline learning. Government becomes a source of durability, standards, and public accountability.

A Four-Way Division of Responsibility

The larger governance architecture of this proposal can therefore be summarized through four complementary forms of responsibility.

Artificial Intelligence Provides Scalable Capacity

AI can contribute continuous availability, multilingual access, structured support, repeatable functions, navigation, and scalable interaction within clearly defined boundaries. Its role is capacity.

Helping Professionals Provide Judgment and Human Care

Therapists, counsellors, psychologists, social workers, physicians, supervisors, and other qualified professionals contribute complexity management, professional judgment, therapeutic and helping relationships, risk assessment, ethics, intervention, supervision, and genuine human presence. Their role is expertise and accountable human care.

Social Entrepreneurs Provide Innovation and Frontline Learning

Social entrepreneurs identify unmet needs, remain close to communities, challenge assumptions, experiment, develop prototypes, and reveal possibilities that established systems may not yet recognize. Their role is discovery and innovation.

Government Provides Durability and Public Responsibility

Government contributes standards, funding architecture, regulatory clarity, equitable access, independent evaluation, workforce transition, accountability, institutional continuity, and public responsibility. Its role is durability and stewardship.

No one of these actors should be expected to carry the entire system. The governance objective is to assign responsibility according to institutional comparative advantage in the same way that the service model assigns functions according to human and technological comparative advantage.

From Founder-Led Innovation to a National Public Capability

The broader journey represented by this proposal can ultimately be understood as a transition: **From founder-led experimentation to institutionally supported public capability.**

FASS tested the value of an accessible human bridge before, between, and alongside professional care.

FASSLING™ explored whether technology could make an appropriate part of that safety net continuously available.

Frontline experience raised the question of how scarce professional attention is currently allocated.

The proposed 50/50 demonstration establishes a hypothesis for reorganizing public resources between scalable availability and protected human expertise.

The three-layer service model defines the respective roles of AI, human supervision, and professional care.

The trust and safety architecture defines the conditions under which such a system could become worthy of public confidence.

The remaining question is institutional: **Who should carry responsibility when an innovation becomes too socially important to depend on private sacrifice alone?**

The answer should not be: One founder forever. It should be a distributed public architecture in which:

- **AI scales appropriate availability;**
- **professionals safeguard judgment, complexity, and human care;**
- **social entrepreneurs continue to innovate; and**
- **government ensures that the public capability can endure.**

The central governance principle is therefore: **Essential public infrastructure should not depend indefinitely on individual sacrifice.**

If Canada ultimately determines—through rigorous evidence—that people should have access to a free-at-the-point-of-appropriate-use, continuous, multilingual, accessible, text- and audio-enabled, human-supervised foundational emotional-support safety net, that public commitment should be backed by institutions capable of outlasting any particular founder, organization, technology, vendor, or funding cycle.

The objective is not to preserve FASSLING™ indefinitely. It is to preserve, where evidence supports doing so, the public capability that FASS and FASSLING™ were developed to explore **that people should not be left with nowhere appropriate and safe to turn.**

Social innovation can reveal that possibility. Government can determine whether the evidence justifies making it durable.

9. Pilot Design and Evidence Agenda

Demonstration, Evaluation, and the Evidence Path to Scale

A Controlled, Staged, Multi-Year Demonstration

The Government of Canada should begin with a **controlled, staged, independently evaluated, multi-year demonstration** of the Human-Supervised Virtual Safe Space model. The first objective should be **learning, not scale**.

The initiative should not begin by maximizing user volume, automating the greatest possible number of functions, or attempting immediate national deployment. Those objectives would be premature before government has established whether the proposed architecture is sufficiently safe, useful, equitable, trusted, professionally sustainable, and fiscally productive.

The demonstration should instead answer a more consequential public-policy question: **Can Canada safely, equitably, and sustainably combine AI-enabled availability, qualified human oversight, professional care, and public accountability in a way that produces greater access and continuity from finite public resources?**

The initiative should therefore be designed as an evidence-generating public-policy experiment rather than a technology rollout.

The unit of evaluation should not be the AI model alone but the **complete human-AI service architecture**: technology, human supervision, professional pathways, escalation mechanisms, privacy and cybersecurity safeguards, accessibility infrastructure, workforce arrangements, governance structures, and funding model operating together as one system.

This distinction is critical. A technically capable AI system could still produce poor public value if it increases professional workload, weakens connection to human care, creates unacceptable privacy risks, fails culturally or linguistically, generates excessive escalation, or proves difficult for people with disabilities to use. Conversely, a system that is imperfect as a standalone technology may still create significant value if it operates effectively within a well-governed human support architecture.

The demonstration should therefore assess, at minimum, **accessibility, continuity, professional capacity, workforce sustainability, safety, fiscal productivity, user trust and dignity, frontline learning, and legal and governance clarity**.

It should also directly test the central economic proposition of this submission: **Can responsible automation of suitable structured, repetitive, informational, navigational, and highly text- or audio-mediated functions release scarce professional human attention for circumstances in which human expertise creates substantially greater value?**

A further question follows: **Can that reallocation support a free-at-the-point-of-appropriate-use, multilingual, accessible, human-supervised, continuously available foundational safety net without requiring professional labour to increase in direct proportion to every additional supportive interaction?**

The demonstration should include sufficient variation to determine where the model works, for whom, and under what conditions. Appropriate experimental or quasi-experimental variation could include different languages, geographic settings, hours of use, accessibility configurations, levels of user need, human-supervision ratios, escalation models, structured service functions, cultural contexts, AI infrastructure arrangements, and allocations between computational and human capacity.

The approximate 50/50 funding model proposed elsewhere in this submission should therefore be treated as **one testable starting configuration**, not as a predetermined national formula.

Expansion should occur only where preceding stages demonstrate acceptable performance in safety, privacy, trust, accessibility, equity, cultural responsiveness, professional sustainability, governance, usefulness, and public value.

The governing principle should be: **Evidence before expansion, safeguards before scale, and public value before permanence.**

A Five-Stage Demonstration Path

A staged implementation model would allow government to build evidence systematically while limiting avoidable risk. Each stage should have defined objectives, evaluation criteria, and explicit conditions that must be met before progression.

Stage 1 — Co-Design, Governance, and Baseline Development

The first stage should establish the institutional and evaluative architecture before broad public deployment begins. Government should convene appropriate professional advisers, people with lived

experience, privacy and cybersecurity experts, accessibility specialists, cultural and linguistic experts, legal and regulatory advisers, responsible-AI and safety researchers, research institutions, social-innovation representatives, and Indigenous partners where appropriate and consistent with Indigenous rights, governance, and data principles.

This stage should define the functions the AI system may perform, the functions it may not perform, professional and service boundaries, escalation thresholds, human-review protocols, consent requirements, privacy and data-retention rules, minimum supervision requirements, accessibility obligations, cultural-responsiveness expectations, complaints and redress mechanisms, independent auditing requirements, and criteria for advancing to later phases.

The evaluation framework should be established at this stage rather than constructed retrospectively. Where feasible, government should also establish **baseline measures or relevant comparison conditions** using existing service architectures. These may include current wait times, service utilization, professional workload, operating-hour constraints, accessibility performance, multilingual availability, navigation outcomes, cost patterns, and continuity gaps.

The relevant question is not merely whether the demonstration appears useful in isolation. It is: **Does the proposed architecture provide meaningful improvement relative to appropriate existing alternatives or benchmarks?** Baseline development is therefore essential to credible evaluation.

Stage 2 — Controlled Launch

The second stage should begin with clearly bounded functions, limited populations or service settings where appropriate, close human oversight, and intentionally conservative scope.

Initial AI-enabled functions should focus on areas where responsible automation appears most plausible, such as structured reflection, routine psychoeducation, non-clinical coaching, guided exercises, basic service navigation, goal clarification, and appropriately bounded emotional support through text or audio-enabled interaction.

Higher-risk, clinically complex, diagnostic, treatment-related, or otherwise regulated functions should remain outside automated scope unless expressly authorized within an appropriate professional and legal framework.

During this phase, qualified supervisors and independent evaluators should closely monitor boundary adherence, safety, privacy, escalation behaviour, user understanding, accessibility, multilingual quality, cultural responsiveness, interaction quality, and unintended consequences.

Special attention should be paid to whether users understand that the system is AI, whether they understand what it can and cannot provide, and whether the service strengthens rather than weakens appropriate pathways to professional care.

This stage should deliberately favour **controlled learning over rapid growth**.

Stage 3 — Responsible Expansion

Expansion should occur only where evidence supports it. Potential expansion could include additional languages, accessibility functions, geographic regions, rural and remote populations, new service-navigation pathways, different operating contexts, new structured support functions, or revised human-supervision models.

Each expansion should be treated as a **new hypothesis**, not as an automatic extension of earlier success. For example, satisfactory performance in English should not be assumed to imply equivalent safety or quality in another language. Success in an urban population should not automatically establish suitability for rural or remote communities. Safe use for routine reflection should not justify expansion into clinically complex functions. A supervision ratio that works in one context may be insufficient in another.

Generalization should therefore be demonstrated rather than presumed. **Scale should be earned through evidence.**

Stage 4 — Independent Evaluation and Comparative Analysis

The fourth stage should evaluate the complete architecture after sufficient operational experience has accumulated. Independent evaluators should assess accessibility, continuity, user outcomes, safety, privacy, cybersecurity, professional workload, professional satisfaction, workforce transformation, accessibility, cultural and linguistic performance, user trust, escalation quality, legal and governance clarity, AI infrastructure performance, frontline learning, fiscal productivity, and unintended consequences.

Where feasible, comparative analysis should examine the mixed human-AI model against relevant existing service models, benchmarks, or counterfactuals.

The evaluative question should not be: Did people use the AI? Nor should it be: Did the AI generate acceptable responses?

The relevant question is: **Did the complete human-AI architecture produce greater public value?**

Public value should be judged across access, continuity, safety, equity, professional capacity, user dignity, workforce sustainability, governance, and fiscal productivity.

Stage 5 — Public-Infrastructure Decision

The final stage should determine what, if anything, government should do next. Possible conclusions should include continuation, redesign, targeted expansion, restriction of certain functions, additional testing, revision of the funding model, or discontinuation.

The initiative should be structured so that **discontinuation is a legitimate policy outcome** if the evidence does not justify expansion. If results are sufficiently strong, government could then consider establishing a **technology-neutral public-interest standard** for continuously available, multilingual, trauma-informed, accessible, privacy-preserving, human-supervised emotional-support infrastructure connected to appropriate professional care.

A successful demonstration should not automatically convert the original platform, vendor, or organizational arrangement into permanent infrastructure. Government should consider institutionalizing **the demonstrated capability and public standard**, not necessarily the technology or organization through which the capability was first tested.

Nine Evaluation Hypotheses

The demonstration should be organized around **nine explicit and falsifiable evaluation hypotheses**. Together, these hypotheses test whether the proposed architecture creates meaningful value for Canadians, helping professionals, government, and the broader public-interest innovation ecosystem.

Comprehensive Nine-Hypothesis Evaluation Framework

No.	Evaluation Domain	Central Hypothesis / Policy Question	Illustrative Measures	What Positive Evidence Would Suggest
1	**Accessibility**	Can human-supervised AI provide meaningful first-line emotional support, reflection, non-clinical coaching, and navigation at materially greater practical reach?	Time to first access; availability outside conventional hours; multilingual utilization; rural and remote reach; disability accessibility; text/audio use; low-bandwidth performance; user-reported lack of alternative immediate support.	AI-enabled infrastructure can broaden the practical reach of an appropriate foundational safety net beyond conventional constraints of operating hours, geography, language, and professional availability.
2	**Continuity**	Does the virtual safe space reduce periods in which people have nowhere appropriate to turn before, between, or after professional encounters?	Use while waiting; between appointments; during provider transitions; after short-term programs; length of unsupported gaps; successful connection or reconnection to care; user-reported continuity.	The model can create meaningful continuity rather than simply adding another isolated episode of service.
3	**Professional Capacity**	Can suitable structured functions be AI-assisted while increasing professional attention available for complex human needs?	Professional time spent on routine versus complex functions; time released; case complexity; users supported per professional hour; supervision burden; referral and intervention workload.	Responsible automation can protect scarce professional attention rather than merely reduce labour expenditure.

No.	Evaluation Domain	Central Hypothesis / Policy Question	Illustrative Measures	What Positive Evidence Would Suggest
4	**Workforce Sustainability**	Can helping professionals transition toward supervision, escalation, complex care, quality assurance, and governance without diminishing job quality, autonomy, or sustainability?	Workload; burnout; satisfaction; retention; professional autonomy; perceived meaningfulness; complex-care capacity; training requirements; emergence of new roles.	Human-AI integration can support **job evolution rather than indiscriminate job elimination**.
5	**Safety**	Can the system maintain appropriate boundaries, trauma-informed design, safe escalation, privacy, accessibility, and human intervention as availability expands?	Missed escalation; unnecessary escalation; harmful-output patterns; boundary adherence; privacy/security incidents; accessibility failures; cultural failures; timeliness and quality of human intervention; user understanding of limitations.	Scale can expand without producing unacceptable risk when supported by sufficiently strong human governance.
6	**Fiscal Productivity**	Can the complete architecture produce greater access and continuity from a comparable funding envelope?	Cost per person meaningfully supported; cost per meaningful interaction; navigation outcomes; infrastructure costs; inference costs; human-supervision costs; escalation costs; users supported per professional hour; released professional capacity; total expenditure relative to meaningful outcomes.	Resource reallocation can generate greater overall public value rather than merely reducing unit costs.
7	**User Trust and Dignity**	Do users experience the service as safe, understandable, respectful, useful, private, accessible, and culturally responsive?	Psychological safety; trust; dignity; clarity about AI/human roles; privacy confidence; autonomy; willingness to return; cultural responsiveness; accessibility; connection to human care.	The service can earn sufficient legitimacy for voluntary use as part of public infrastructure.
8	**Frontline Learning**	Can privacy-preserving aggregate analysis identify recurring service barriers earlier without converting personal conversations into surveillance?	Recurring navigation failures; language gaps; geographic disparities; accessibility gaps; waitlist patterns; time from emerging issue to institutional attention; absence of inappropriate individualized profiling.	AI-enabled system learning can bring leadership closer to frontline reality while preserving dignity and privacy.
9	**Legal and Governance Clarity**	Can a shared public framework reduce uncertainty regarding scope, supervision, accountability, privacy, liability, and cross-jurisdictional responsibility?	Unresolved scope questions; boundary clarity; supervision standards; privacy ambiguities; complaints; liability gaps; jurisdictional conflicts; stakeholder confidence; areas requiring new rules.	Government can replace fragmented private interpretation with clearer and more coherent public standards.

These hypotheses should not be interpreted as predetermined conclusions. Their purpose is to make the proposal **testable**. The demonstration should be credible enough to show not only where the

model succeeds, but where it fails, for whom it performs poorly, and under which circumstances greater human involvement is required.

How the Nine Hypotheses Fit Together

The nine hypotheses are not independent technical indicators. Together, they form an integrated theory of public value.

Accessibility, continuity, and user trust test whether Canadians actually receive a more usable and dependable support architecture.

Professional capacity and workforce sustainability test whether AI releases scarce human attention and supports higher-value professional roles rather than simply transferring burden or reducing employment.

Safety and legal/governance clarity test whether increased technological capacity can remain bounded, accountable, professionally governed, and institutionally coherent.

Fiscal productivity and frontline learning test whether government obtains greater value from finite resources while becoming more responsive to recurring service barriers.

The model can be summarized as follows:

Policy Foundation	Primary Test	Intended Result
AI-enabled availability	Accessibility	More people can obtain an immediate appropriate first layer of support
Continuity-oriented design	Continuity	Fewer people experience unsupported gaps
Responsible automation	Professional Capacity	Scarce professional attention shifts toward more complex human needs
Workforce redesign	Workforce Sustainability	Professionals move toward higher-value roles without unsustainable burden
Human-supervised safeguards	Safety	Availability expands without unacceptable loss of control or accountability
50/50 resource-allocation hypothesis	Fiscal Productivity	Comparable funding may generate greater overall public value
Trauma-informed, privacy-preserving design	User Trust and Dignity	People can use the service voluntarily while understanding its role and limits
Privacy-preserving aggregate learning	Frontline Learning	Leadership identifies systemic barriers earlier without individualized surveillance
Multi-stakeholder public governance	Legal and Governance Clarity	Responsibility, scope, and accountability become more coherent

The nine hypotheses therefore evaluate the proposal as a **public-service system**, not as a standalone technology.

The Central Evaluation Chain

The demonstration should explicitly test the causal chain underlying the proposal.

Responsible automation of suitable structured functions

↓

less professional time consumed by appropriate repetitive or highly standardized work

↓

more professional attention available for judgment, complexity, therapeutic relationships, intervention, supervision, and safety

↓

AI-enabled infrastructure creates greater foundational availability

↓

more people can obtain an appropriate first layer of support without paying for each interaction

↓

fewer people experience complete gaps while waiting, searching, or moving between services

↓

professional expertise becomes more concentrated where professional expertise creates the greatest value

↓

the complete system produces greater access and continuity from a comparable funding envelope

↓

evidence determines whether a durable Continuity Guarantee is operationally, professionally, and fiscally credible

Every link in this chain should be measured.

It would be insufficient to demonstrate that AI can perform selected activities if the released professional time does not translate into improved complex-care capacity.

It would be insufficient to show increased access if users become less connected to human services.

It would be insufficient to reduce unit cost if supervision, escalation, safety failures, or administrative complexity offset the apparent savings.

The demonstration must therefore determine whether **resource reallocation actually produces the public outcomes the proposal predicts**.

Measures of Success: Public Value, Not Activity

Success should not be measured by the number of AI conversations.

Nor should raw interaction volume, tokens consumed, downloads, logins, screen time, or repeat usage be treated as primary indicators of policy success.

Those are activity measures.

Activity is not the same as public value.

A service could generate millions of interactions and still fail if people do not trust it, if privacy is compromised, if professional workload increases, if accessibility remains poor, if referrals deteriorate, if dependency increases, or if users become more disconnected from appropriate professional and community care.

Success should instead be defined through meaningful outcomes.

These should include greater and more equitable access; materially shorter time to an appropriate first layer of support; stronger continuity before, between, and after professional encounters; fewer periods during which people report having nowhere appropriate to turn; improved navigation; appropriate connection to professional and community services; safe and proportionate escalation; effective maintenance of professional boundaries; measurable release of professional time from suitable routine functions; greater professional capacity for complex needs; sustainable professional roles; user trust and dignity; cultural responsiveness; accessibility; privacy and cybersecurity performance; fiscal productivity; clearer professional and legal governance; earlier recognition of recurring systemic barriers; and stronger—not weaker—connection to appropriate human care.

The demonstration should also test the proposal against its two larger national aspirations. It should ask:

Does the model make the Care Guarantee more attainable by protecting professional capacity for those who genuinely require professional care?

and:

Does it make the Continuity Guarantee more operationally and economically realistic by reducing complete gaps in support?

Measuring the Reallocation of Human Attention

One metric deserves particular emphasis because it captures the central workforce and economic proposition of the proposal: **Where does professional human attention go after suitable functions are automated or AI-assisted?**

It is not enough to report that a digital system "saved" a certain number of professional hours. Those hours should not simply disappear from the workforce budget and automatically be classified as a productivity gain.

The demonstration should determine whether released capacity is actually redirected toward complex care, therapeutic relationships, trauma-informed professional intervention, risk assessment, difficult navigation, escalation, supervision, quality assurance, professional judgment, or other functions in which human expertise is especially valuable.

For example, if a redesigned process releases 10,000 professional hours over a defined period, evaluation should examine how those hours were subsequently used.

- Did more people receive complex professional care?
- Did waiting periods for high-value professional functions decrease?
- Did practitioners report that a greater proportion of their work involved functions requiring professional expertise?
- Did quality improve?
- Did professional burnout improve or worsen?
- Did new supervision obligations consume much of the apparent capacity gain?
- Did escalations create more work than anticipated?

These questions distinguish between two very different outcomes: **labour eliminated** and **professional attention reallocated**. The second is the objective of this proposal. This distinction should be reflected explicitly in the evaluation framework.

Evaluating the 50/50 Funding Hypothesis

The approximate 50/50 funding allocation should itself be treated as an experiment. Government should not simply determine whether the demonstration remained within its assigned budget. It should assess **what each side of the funding envelope purchased and how the two interacted**.

On the AI-infrastructure side, evaluation should examine model and inference expenditures, computational capacity, technical operating costs, continuous availability, multilingual quality,

accessibility performance, security, privacy protections, number and type of appropriate interactions supported, marginal costs as utilization changes, and the technical resources required to maintain acceptable performance.

On the human-capacity side, government should evaluate professional hours, supervision requirements, complex-care capacity, intervention workload, escalation, safety review, quality assurance, training, professional satisfaction, governance costs, and the quality and timeliness of human responses.

Different allocations should be tested where appropriate. One service setting might function best with approximately 40% of resources directed to infrastructure and 60% to human capacity. Another may operate safely and effectively with the reverse. A population with greater complexity may require substantially more professional involvement. A low-risk, highly structured use case may require proportionately less.

The purpose is not to prove that the 50/50 hypothesis is correct. It is to identify the allocation—or range of allocations—that **maximizes meaningful public value while preserving safety, professional quality, equity, trust, and access to human care**.

Decision Gates for Responsible Expansion

Each stage of expansion should be subject to explicit decision gates. The demonstration should not progress simply because a predetermined implementation schedule has reached its next date. Progression should depend on evidence. Before adding significant new populations, languages, functions, geographic areas, or jurisdictions, decision-makers should assess at least the following:

Decision Gate	Core Question
Safety	Are risks being identified, contained, escalated, and reviewed appropriately?
Privacy	Are sensitive disclosures adequately protected and secondary uses tightly controlled?
Trust	Do users understand the service and trust it sufficiently for voluntary and informed use?
Accessibility	Can people with diverse needs actually access, understand, and navigate the service?
Cultural Responsiveness	Is the system functioning appropriately across relevant cultural and linguistic contexts?
Professional Capacity	Is AI assistance genuinely releasing useful professional attention rather than merely shifting workload?
Workforce Sustainability	Are professional roles manageable, meaningful, adequately trained, and appropriately supported?
Fiscal Performance	Is the complete system generating greater public value relative to its total cost?
Governance	Are scope, responsibility, complaints, review, redress, and accountability sufficiently clear?
Connection to Human Care	Is the system strengthening rather than weakening pathways to appropriate professional, community, crisis, and emergency services?

Failure at a major decision gate should not be treated as an inconvenience to the implementation schedule. It should trigger redesign, additional safeguards, narrower scope, further evaluation, or suspension where necessary. A public demonstration earns credibility by having the institutional capacity to stop.

Closing the Frontline Gap

One of the potential benefits of the proposed infrastructure extends beyond individual service delivery. It concerns the distance between frontline experience and institutional decision-making. In large systems, lived experience often travels through multiple layers before it reaches senior decision-makers.

- A person speaks with a frontline worker.
- The interaction becomes documentation.
- Documentation becomes a summary.
- The summary becomes a report.
- The report becomes an indicator.
- Leadership eventually receives a number.
- Each transformation is necessary for scale, but context can be lost along the way.
- Nuance may disappear.
- Weak signals may disappear.
- A problem may affect large numbers of people before it becomes visible enough to alter an aggregate indicator.

This can create a **frontline gap**: the distance between what people are experiencing now and what institutional leadership can see through conventional reporting structures. Appropriately governed AI may help narrow that distance.

With strong privacy protections, de-identification, aggregation, data minimization, purpose limitation, and independent oversight, the system may be able to identify recurring patterns such as failed navigation pathways, underserved languages, regional disparities, repeated waitlist barriers, accessibility problems, common difficulties finding appropriate providers, or emerging forms of unmet need.

The objective must remain structural. It is not to convert intimate conversations into individualized government intelligence. The principle remains: **The purpose is not watching people at scale. It is understanding systemic barriers at scale.**

Bringing Support Closer to the Public and Leadership Closer to Reality

The proposed infrastructure may therefore provide two complementary public functions. The first is to **bring support closer to the public** by providing an appropriate first layer of support, reflection, navigation, and non-clinical coaching when conventional services are not immediately available. The second is to **bring leadership closer to frontline reality** by identifying privacy-preserving aggregate patterns that may reveal where service systems repeatedly fail.

This second function should never replace direct engagement with communities. If technology identifies a recurring barrier affecting a particular population, public leaders should spend **more**, not less, time engaging with the people affected.

The appropriate principle is: **AI identifies patterns. Humans investigate meaning.** Aggregate signals can direct attention. They cannot substitute for listening.

The Human Policy-Learning Loop

Policy learning should therefore remain fundamentally human. A responsible learning cycle would proceed as follows:

frontline experience

↓

privacy-preserving pattern recognition

↓

leadership attention

↓

direct engagement with affected users, professionals, and communities

↓

policy or service adjustment

↓

frontline implementation

↓

user and professional evaluation

↓

continuous learning

The purpose of AI is not to replace consultation. It is to help identify where consultation may be needed sooner. AI can summarize patterns at scale. Human leaders must still listen. This reflects the principle that shaped the proposal itself: **The innovation did not begin with AI. It began with listening.**

Technology should help public institutions listen more effectively at scale without losing the human relationships through which meaning is ultimately understood.

The Final Evidence Question

At the conclusion of the demonstration, government should be able to answer a clear set of questions.

- Did more people obtain meaningful and appropriate support?
- Did they obtain it faster?
- Did access improve across language, geography, disability, and conventional service hours?
- Did fewer people experience periods in which they had nowhere appropriate to turn?
- Did the virtual safe space strengthen rather than weaken connections to professional and community care?
- Did responsible automation release professional capacity?
- Was that capacity actually redirected toward people and functions requiring higher levels of human expertise?
- Did professional job quality, autonomy, sustainability, and satisfaction remain acceptable?
- Were escalation and human intervention appropriate and timely?
- Was privacy protected?
- Did cybersecurity remain adequate?
- Did users understand the distinction between AI support and professional care?
- Did culturally diverse users experience the service as appropriate and respectful?
- Did the system generate useful aggregate learning without creating individualized surveillance?
- Did professional, legal, and institutional responsibilities become clearer?
- And, taken as a whole, did the model generate greater public value from a comparable funding envelope?

Only after those questions have been answered should government consider the larger one: **Has Canada demonstrated a sufficiently safe, equitable, trusted, professionally sustainable, and economically credible pathway toward making a free-at-the-point-of-appropriate-use, multilingual, accessible, text- and audio-enabled, continuously available foundational mental-health and emotional-support safety net a durable public capability?**

That is the ultimate purpose of the demonstration. The pilot is not intended to establish that artificial intelligence exists or that it can sustain a conversation. Those questions are too narrow. The demonstration is intended to determine whether Canada can **govern AI responsibly, combine it with qualified human expertise, allocate public resources around it intelligently, protect professional capacity, preserve human dignity, and use the resulting architecture to achieve a public objective that a human-only model has historically struggled to make continuously available at population scale**. The desired endpoint is therefore not more AI. It is:

- greater access
- greater continuity

- better protection and use of professional human attention
- stronger public accountability
- greater trust and equity

and—only if the evidence supports it—

- a credible path toward ensuring that no Canadian is left with nowhere appropriate and safe to turn.

10. The Four-Way Public-Value Model

The Public-Value Proposition: Benefits for People, Professionals, Government, Social Innovators, and Canada

The proposed **Human-Supervised Virtual Safe Space Demonstration Initiative** should ultimately be judged not by the sophistication of the technology deployed, but by the public value the complete system creates. Its central proposition is that several interests commonly treated as competing may, under the right architecture, become mutually reinforcing.

A well-designed model could potentially provide:

- greater access and continuity for the public;
- better use of professional expertise for the helping workforce;
- stronger fiscal productivity and system capacity for government and taxpayers;
- a more sustainable innovation environment for social entrepreneurs; and
- a credible international leadership opportunity for Canada.

These outcomes are not guaranteed. They are precisely what the proposed demonstration should test. The policy opportunity lies in determining whether responsible resource reallocation can produce a system in which computational capacity expands appropriate availability without degrading professional care; professional human attention is protected for complexity and judgment; governments obtain greater value from finite resources; social innovators are not required to carry essential public infrastructure indefinitely; and the public benefits from a more continuous and dependable support architecture.

If the evidence supports that proposition, the value of the model would extend considerably beyond any particular AI platform. It would represent a new approach to organizing public capacity.

Public Value for Canadians: From Episodic Access to Greater Continuity

The most important beneficiary of the proposed model is the person seeking support. If the demonstration succeeds, Canadians could eventually gain access to an appropriate foundational support layer that is free at the point of appropriate use, immediately available, accessible 24 hours a day and 365 days a year, multilingual, available through text and audio-enabled interaction, and designed to remain accessible before, between, and alongside episodes of professional care. Such a model would broaden the meaning of mental-health accessibility.

Traditional measures of access often focus appropriately on whether a person can obtain an appointment, enter a program, receive a referral, or access a defined number of funded sessions. Those measures remain essential. They do not, however, fully capture what happens in the intervals surrounding formal services.

A person may still reasonably ask: What happens tonight? What happens while I wait? What happens if the first therapist is not the right fit? What happens when a short-term counselling allocation ends? What happens if the same difficulty returns months later? What happens if I need support in another language? What happens if I recognize that I am struggling but do not yet know what kind of service I need?

The proposed architecture is designed to reduce the frequency with which those questions produce the same outcome: **Nothing appropriate is available right now.**

The objective is not unlimited AI psychotherapy, automated diagnosis, or unrestricted digital clinical care. It is **continuous access to an appropriate, bounded foundational layer of emotional support, guided reflection, non-clinical coaching, life-skills assistance, routine psychoeducation, and service navigation—subject to clear safety boundaries, privacy protections, human governance, technical capacity, and connection to appropriate professional care**.

That distinction is fundamental. Professional services will continue to have finite capacity. Some forms of care will continue to require appointments, specialized expertise, and direct human relationships. The innovation is that the surrounding foundational safety net may not have to disappear every time those limits are reached.

For the person using the system, the value can be expressed simply: **There is somewhere appropriate I can turn to right now.** At its most human level, the desired public experience is: *My country has my back.* That phrase should not imply that government can eliminate every waitlist or make every professional continuously available. It represents a more credible commitment: **Temporary limits in professional availability should not automatically become complete absence of support**.

From Improved Access to Two Complementary Public Guarantees

The larger public-value proposition is captured by two complementary aspirations.

The Care Guarantee: No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.

The Care Guarantee protects access to professional expertise. It addresses the financial barriers that may prevent people from obtaining assessment, treatment, psychotherapy, specialized intervention, or other appropriate human care.

The Continuity Guarantee: No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.

The Continuity Guarantee addresses a different problem: the gaps surrounding professional care. These guarantees should not compete. The first recognizes that professional expertise must remain accessible when it is required. The second recognizes that professional expertise will remain finite and cannot be continuously available to every person at every moment.

The proposed human-AI architecture could make the second aspiration more operationally and economically plausible by combining

- Scalable availability through digital infrastructure.
- Professional expertise where professional expertise is necessary.
- Qualified human supervision where accountability is required.
- Government stewardship to ensure continuity and durability.

For the individual, the institutional complexity behind the model should ultimately feel simple.

- Professional care when professional care is needed.
- Appropriate support while waiting.
- Somewhere to turn between appointments.
- A place to begin again if the first provider is not the right fit.
- No requirement to pay out of pocket for every appropriate foundational interaction.
- Text or audio-enabled access depending on what is usable.
- Multilingual availability where the system can safely support it.
- A safety net that does not automatically disappear because one funded episode has ended.

That is the public value the Continuity Guarantee is intended to create.

Public Value for Helping Professionals: Protecting Human Expertise

Helping professionals are not peripheral to the proposed model. They are central to it. Indeed, one of the core premises of this submission is that therapists, counsellors, psychologists, social workers, physicians, supervisors, and other qualified professionals should be treated as **more valuable, not less, in an AI-enabled service environment**.

Many professional roles currently contain combinations of highly specialized and relatively standardized functions. Some activities depend heavily on judgment, therapeutic relationship, ethical reasoning, risk assessment, cultural interpretation, or sustained human presence. Others may be more structured, repetitive, informational, procedural, or text-based.

Where evidence demonstrates that selected structured functions can be responsibly AI-assisted, professionals may be able to devote a greater proportion of their limited time to work that depends on distinctly human expertise. That includes deep listening, therapeutic relationship-building, complex trauma, professional judgment, risk assessment, ethical decision-making, culturally nuanced care, ambiguous circumstances, difficult family or relational dynamics, human intervention, escalation, supervision, quality assurance, and professional governance.

The relevant workforce question is therefore not: How many professional jobs can AI eliminate? It is **How much professional attention can be protected and redirected toward work that genuinely requires professional expertise?** That is a fundamentally different workforce philosophy.

Professional human attention has an opportunity cost. If a highly trained professional spends substantial time manually reproducing a standardized function that another appropriately governed form of capacity could safely support, the system does not incur only the financial cost of that hour. It may also leave another person waiting for expertise that cannot be replicated through automation.

Responsible AI assistance can therefore become a mechanism for protecting professional capacity. The desired outcome is not less human care. It is less professional time consumed by functions technology can appropriately support, and more professional time available for functions technology cannot replace.

Job Evolution and the Future of Helping Professions

The transition may also create new or expanded professional responsibilities. As AI becomes more deeply integrated into emotionally sensitive public services, Canada will require professionals not only to provide direct care, but also to govern, evaluate, supervise, and improve the systems surrounding that care.

Potential roles may include AI-support supervisors, escalation specialists, AI-safety reviewers, trauma-informed AI evaluators, clinical AI advisers, professional-boundary reviewers, quality-assurance specialists, culturally responsive AI reviewers, digital mental-health governance specialists, human-machine service designers, AI-assisted service navigators, privacy and ethics advisers, and complex-care specialists.

Some of these roles already exist in limited forms. Others may develop as public systems gain experience with hybrid service delivery. The workforce objective should therefore remain: **job evolution, not job elimination.**

Success should not be measured by the number of professionals removed from service delivery. It should be measured by whether professionals spend more time applying their education and judgment, retain meaningful and sustainable employment, gain appropriate new competencies, increase their capacity to serve complex needs, and remain central to the governance of emotionally sensitive technologies.

AI should create room for **better human work**. It should not simply create less human work.

Public Value for Government and Taxpayers: Greater Value From Finite Resources

For government and taxpayers, the proposed architecture creates a different kind of productivity opportunity. The demonstration could test whether Canada can broaden availability, extend support beyond conventional office hours, improve multilingual access, better reach rural and remote populations, reduce avoidable service gaps, strengthen navigation, protect professional hours, support workforce evolution, improve frontline learning, and maintain strong safety and governance within a comparable overall funding envelope.

The relevant measure is not the lowest possible cost per AI conversation. That metric would be inadequate. A low-cost interaction that produces little meaningful benefit has little public value. A system that lowers labour expenditure while increasing risk, weakening privacy, reducing trust, or disconnecting people from professional care would not represent a successful modernization of public services.

The correct fiscal question is: **Does the complete human-AI architecture produce better public outcomes from finite public resources?**

Government should therefore evaluate the model across access, continuity, safety, equity, professional capacity, user trust, privacy, accessibility, workforce effects, navigation, human-care connection, and overall cost.

Fiscal productivity should mean **more public value per dollar—not simply fewer dollars spent.** This distinction is essential. The proposed initiative is not an austerity strategy disguised as technological innovation. Its purpose is to examine whether scarce financial and human resources can be organized more intelligently.

Public Value Through Better Allocation of Human Attention

Government may also obtain a form of productivity that conventional financial metrics often fail to capture: **better allocation of professional human attention**. If responsible automation releases

thousands of hours of professional capacity, the government should not evaluate the result solely by converting those hours into salary savings.

The more important questions are:

- Where did those hours go?
- Did waiting periods for complex care decline?
- Did professionals gain more time for trauma-related work, difficult cases, or therapeutic relationships?
- Did access to human intervention improve?
- Did supervisors gain capacity for meaningful oversight?
- Did professional workload become more sustainable?
- Did the quality of human care improve?

The most important productivity gain may therefore be **not eliminating the professional hour, but moving the professional hour to where it creates greater human value**. That is the resource-reallocation principle at the centre of this submission. A system that releases professional time but does not redirect it toward unmet human needs has not fully realized the intended public benefit. The demonstration should therefore treat **attention reallocation** as a measurable outcome in its own right.

Public Value for Social Entrepreneurs: Innovation Without Indefinite Systemic Burden

Social entrepreneurs also stand to benefit from a more mature public architecture. They play an important role in social innovation precisely because they are often willing to enter gaps that larger institutions have not yet addressed. They listen to communities, experiment, challenge assumptions, accept uncertainty, build prototypes, and reveal possibilities before conventional systems are ready to act.

That contribution is valuable. But the role of innovator should not automatically become the role of permanent infrastructure provider. This distinction is particularly important when social entrepreneurs operate in the public interest rather than primarily for private financial gain.

A founder may contribute years of unpaid or underpaid labour, professional expertise, financial resources, reputational capital, emotional energy, and substantial personal effort to create a service intended to assist people who otherwise might have nowhere to turn. As the service grows, the same founder or small nonprofit may become responsible for privacy, cybersecurity, data governance, AI safety, professional boundaries, supervision, crisis protocols, accessibility, insurance, quality assurance, regulatory interpretation, ethical oversight, and potential legal liability.

This creates a structural paradox: **The greater the public benefit created, the greater the private burden the innovator may be required to absorb.** That is not a sustainable model for essential social infrastructure.

Public policy should not assume that socially useful innovation can be sustained indefinitely through the financial capacity, legal tolerance, emotional endurance, or personal sacrifice of one or two founders. The principle should remain: **Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.**

Reducing Founder Dependency Is Also Public Risk Management

Reducing dependency on individual founders is not only a matter of fairness. It is also prudent **public risk management**. A capability becomes fragile when its continued existence depends disproportionately on one person, one nonprofit organization, one grant, one vendor, one donor, or one proprietary technology.

Founders may leave. Organizations may close. Funding arrangements may change. Technologies can become obsolete. Vendors can alter their strategies. Leadership and procurement priorities can shift. If a capability has become socially important, it should be designed to survive those transitions.

Government therefore need not guarantee that FASSLING™, or any specific organization or technology, operates indefinitely. The longer-term objective should be to preserve the underlying technology-neutral public capability: free-at-the-point-of-appropriate-use, continuous, multilingual, trauma-informed, accessible, human-supervised emotional-support infrastructure connected to appropriate professional care.

The platform may change. The AI model may change. The provider may change. The governance model may evolve. The public promise should be capable of enduring. There will continue to be somewhere appropriate and safe to turn. That is the difference between preserving a product and preserving a public capability.

The Four-Part Public-Value Model at a Glance

Beneficiary	What the Model Could Change	Primary Public Value
The Public	Free-at-the-point-of-appropriate-use foundational access; 24/7/365 availability; multilingual text and audio-enabled support; continuity before, between, and alongside professional services; navigation to appropriate care	**Access, continuity, dignity, and reduced support gaps**
Helping Professionals	Less time on suitable standardized functions; more time for judgment, relationships, complex care, safety, intervention, supervision, and AI governance	**Protected professional attention and higher-value professional work**
Government and Taxpayers	Scalable infrastructure complements finite professional capacity; broader reach from a comparable funding envelope; improved system learning and service navigation	**Fiscal productivity, system capacity, and stronger public outcomes**
Social Entrepreneurs	Innovators can demonstrate public-value solutions without indefinitely carrying legal, financial, technological, ethical, and operational responsibility alone	**More sustainable innovation and reduced systemic fragility**

These benefits are interconnected. Greater scalable availability can reduce unnecessary pressure on scarce professional capacity. Better allocation of professional attention can strengthen human care. Greater fiscal productivity can make broader access more sustainable. Institutional continuity can allow

social entrepreneurs to return to experimentation rather than spending their careers personally maintaining one public-interest solution. This is why the proposal should be evaluated as a **system architecture**, not as an AI product.

A Fifth Public Value: Canadian Leadership in Human-Centred Institutional Design

There is also a broader national opportunity. Governments around the world are confronting a common question as artificial intelligence becomes capable of performing functions that historically required human labour.

Much of the public debate focuses on substitution: Which jobs will disappear? Which tasks can be automated? How much labour expenditure can technology remove? Canada has an opportunity to frame a different question: **How can artificial intelligence create forms of appropriate public abundance that were previously constrained by the scarcity of human attention—and how can that abundance be governed in ways that strengthen rather than diminish human care?**

If the proposed demonstration succeeds, Canada could offer a valuable international example. It could show that responsible AI adoption need not mean abandoning professionals, privatizing public responsibility, weakening social protection, or replacing meaningful human relationships with machines.

Instead, Canada could demonstrate an architecture in which:

- AI expands appropriate availability;
- professionals provide depth, judgment, safety, and genuine human connection;
- social innovators identify and test new possibilities; and
- government protects equitable access, standards, accountability, and institutional durability.

That would be a contribution not merely to AI adoption, but to **human-centred institutional design**.

From Technological Leadership to Social-Infrastructure Leadership

National leadership in artificial intelligence should not be measured only by the number of AI firms established, the scale of computing infrastructure, the volume of patents, or the technical sophistication of models. A country can also demonstrate leadership through **how effectively it organizes technology around public purpose**.

If Canada can demonstrate that responsible resource allocation can support free foundational access, continuous availability, multilingual and accessible interaction, strong privacy, meaningful human supervision, appropriate professional care, and continuity when that professional care is temporarily unavailable, the achievement would extend beyond technological capability.

It would demonstrate **institutional capability**. It would show that public systems can adapt when technology changes the scarcity assumptions around which previous service architectures were designed.

The relevant form of leadership is therefore not having the "best AI." Technology evolves too rapidly for such a claim to remain meaningful. A more durable achievement would be demonstrating how a democratic state can use emerging technology while maintaining professional standards, privacy, social solidarity, fiscal discipline, human accountability, and public trust.

Canada as an International Exemplar

If the evidence supports the model, Canada could provide a transferable example for other jurisdictions considering how artificial intelligence might be integrated responsibly into mental-health and human-service systems. The international lesson should not be: Canada replaced therapists with AI. It should be: **Canada preserved professional human care while using computational capacity to ensure that people were less likely to be left without an appropriate first place to turn.**

That distinction matters. Other countries face many of the same structural challenges: professional shortages, growing demand, constrained public budgets, rural and remote access difficulties, language diversity, finite program models, workforce burnout, fragmented navigation, and populations that remain unsupported while waiting for formal care. A successful Canadian demonstration could therefore generate several transferable policy principles:

- separate availability from expertise;
- use automation selectively for appropriate structured functions;
- protect scarce professional human attention;
- maintain qualified human supervision;
- embed privacy, accessibility, cultural responsiveness, and safety into the infrastructure;
- evaluate resource reallocation at the system level; and
- institutionalize the public capability rather than dependence on one platform.

Canada would therefore not simply be demonstrating how to deploy AI. It would be demonstrating **how a democratic public system can govern AI in service of human dignity**.

National Strength Includes Institutional Reliability

The ability of a country to provide its people with confidence that **there will continue to be somewhere appropriate to turn** is itself a form of national strength. National capacity is not expressed only through economic performance, technological development, security, or international rankings.

It is also expressed through the reliability of institutions. A resilient public system gives people confidence that support will not disappear entirely because a private session is unaffordable, a short-term program has ended, the next appointment is several weeks away, the office is closed, the first professional was not the right fit, a language barrier exists, or a difficult moment occurs outside the boundaries of a conventional program.

A continuously available foundational safety net could communicate something important about the relationship between Canadians and their public institutions:

- You are not abandoned simply because the system is temporarily at capacity.
- Your access to a first place of support does not necessarily disappear when an appointment ends.
- Public responsibility for continuity is not limited to conventional office hours.

That is not a promise of unlimited professional service. It is a promise of institutional reliability.

A Canadian Model of Responsible AI

If the demonstration succeeds, the resulting Canadian approach could be summarized through four commitments.

Abundance Without Abandoning Humans

Use AI to expand appropriate supportive availability while preserving human professionals for functions requiring professional expertise, judgment, relationship, and accountability.

Innovation Without Deregulation

Support technological and social innovation while maintaining clear professional boundaries, privacy protections, accessibility requirements, safety standards, and public accountability.

Productivity Without Dehumanization

Use responsible automation to protect and redirect scarce human attention rather than treating people simply as costs to eliminate.

Continuity Without Pretending Professional Capacity Is Infinite

Acknowledge honestly that professional human attention will remain finite while designing a system in which finite professional capacity does not automatically leave people with nowhere appropriate to turn.

Together, these commitments offer a more constructive framework for responsible public-sector AI. The policy question is no longer simply: What can AI automate? It becomes: **What forms of public value become possible when automation is deliberately used to expand appropriate availability while protecting the human capabilities that should remain scarce, relational, and professionally accountable?**

The Broader Public-Value Proposition

Taken together, the proposed model creates a potential five-part public-value architecture.

- For Canadians, it could provide greater access, continuity, multilingual availability, affordability, navigation, dignity, and reassurance.
- For helping professionals, it could protect scarce professional attention, concentrate expertise on complex human needs, create more meaningful roles, and support new professional pathways.
- For government and taxpayers, it could increase system capacity, strengthen fiscal productivity, improve frontline learning, support responsible AI governance, and potentially generate greater public value from finite resources.
- For social entrepreneurs, it could create a healthier innovation ecosystem in which innovators can identify unmet needs and demonstrate solutions without being expected to personally maintain essential infrastructure indefinitely.
- For Canada, it could provide an opportunity to demonstrate internationally how responsible artificial intelligence can be used to expand social protection rather than merely automate labour.

The larger vision is therefore not fundamentally about the virtual safe space as a product. It is about a model of public administration in which **technological capacity, professional human expertise, social innovation, and public responsibility reinforce one another**.

If the demonstration succeeds, Canada would not need to pretend that professional human labour can become unlimited in order to make a meaningful promise of continuity. It would need:

- sufficient professional care for circumstances requiring professional expertise;
- scalable technological infrastructure for appropriate forms of availability;
- qualified human governance to protect safety, dignity, and accountability; and
- public institutions willing to ensure that the resulting capability can endure.

That architecture could support a simple but consequential long-term public promise: **Wherever a person is, and whenever an appropriate first layer of support is needed, there should be somewhere appropriate and safe to turn.** It should not depend on the ability to pay for every interaction. It should not disappear solely because conventional office hours have ended. It should not disappear simply because a short-term program has concluded. It should strive to be multilingual and accessible through different modes of communication. It should remain connected to real human and professional care. It should be safe, transparent, accountable, and publicly governed.

Furthermore, if Canada can demonstrate that such a model is safe, effective, equitable, trusted, and fiscally sustainable, it could show—both to Canadians and to other countries—what technological progress looks like when innovation, fiscal responsibility, professional expertise, and human dignity are organized around a simple public principle: A country should have its people's backs**.**

11. The Care Guarantee and the Continuity Guarantee

Two National Guarantees: Care and Continuity

Canada should work toward two complementary long-term public guarantees that address different dimensions of mental-health accessibility. The first concerns **access to professional expertise**. The second concerns **access to an appropriate foundational layer of support when professional expertise is not immediately available**. Together, they offer a more complete understanding of what meaningful access should ultimately mean.

A mature mental-health support architecture should seek to answer two fundamental questions:

- **Can I obtain the professional care I need?**
- **Do I have somewhere appropriate and safe to turn right now?**

Canada should aspire, over time, to answer **yes** to both.

The distinction is important. Expanding access to professional care addresses one form of scarcity. Building continuity around that care addresses another. Neither objective is sufficient on its own. The proposed Human-Supervised Virtual Safe Space Demonstration Initiative is intended to test whether responsible AI-enabled infrastructure can help make the second aspiration operationally and economically achievable while strengthening, rather than weakening, the first.

The Care Guarantee

The Principle

> **No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.**

The Care Guarantee expresses the principle that access to necessary professional expertise should not depend primarily on an individual's financial capacity. Where a person requires psychotherapy, psychological assessment, psychiatric or medical care, social work, counselling, trauma treatment, specialized intervention, or another form of appropriate professional service, affordability should not become the principal barrier separating that person from care.

Financial barriers can affect access in several ways. They can delay the decision to seek help. They can interrupt care that has already begun. They can discourage a person from trying another provider after an unsuccessful match. They can leave people feeling compelled to remain with a practitioner who is not the right clinical, cultural, linguistic, or relational fit because restarting the search would impose additional financial cost.

Financial resources can also influence how much opportunity a person has to search for the right care. Someone with greater means may be able to consult multiple providers, try different modalities, or remain in longer-term care until an appropriate therapeutic relationship is established. Someone without those resources may have far less room for trial, error, and choice.

A commitment to equitable mental-health access should therefore continue to include the expansion and protection of appropriate publicly supported professional care. The introduction of AI does not weaken the rationale for that commitment. If anything, it makes the distinction more important.

The purpose of AI-enabled infrastructure is not to create a lower-cost substitute for care that should properly be delivered by a qualified professional. Its purpose is to expand appropriate availability around professional care and to protect scarce professional attention for the circumstances in which professional expertise is genuinely required.

The service architecture should therefore preserve a clear principle:

- **AI can support availability and continuity.**
- **Professionals provide professional care.**

Where a person's circumstances require assessment, therapeutic depth, specialized treatment, complex judgment, risk evaluation, trauma expertise, ethical decision-making, or direct professional intervention, the system should facilitate connection to appropriately qualified human care.

The Care Guarantee protects that boundary.

Why the Care Guarantee Alone Is Not Sufficient

Expanding public coverage would address one of the most significant barriers to professional mental-health care: **cost**. That objective remains essential. But affordability is not the same as availability. Even in a hypothetical system in which every appropriate professional mental-health service were fully publicly funded, the scarcity of human time would remain.

A therapist can see only a finite number of people. A counsellor can conduct only a finite number of sessions. A psychiatrist has limited appointment capacity. A social worker can safely manage only a finite number of complex cases. A qualified professional cannot be simultaneously available to every person at night, on weekends, across every geographic region, in every language, and at every moment between scheduled encounters.

Public coverage can reduce **financial scarcity**. It cannot make **professional human attention unlimited**. As a result, people may continue to encounter meaningful gaps even when cost is not the primary barrier. They may be waiting for an appointment. They may be searching for an appropriate provider. The first professional relationship may not have been a good fit. They may be between sessions. Their time-limited counselling program may have concluded. They may require support outside conventional office hours. They may need another language or a more accessible form of communication. They may live in a rural or remote community. They may be uncertain about what form of care is appropriate. Or their need may recur long after a formal episode of service has ended.

These are not primarily problems of payment. They are problems of **continuity and availability**. This is why a second guarantee is necessary.

The Continuity Guarantee

The Principle

> **No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.**

The Continuity Guarantee addresses the spaces surrounding formal services. Those spaces may arise before the first appointment, while a person is waiting for professional care, between scheduled encounters, during a transition between providers, after an unsuccessful therapeutic match, after the end of an Employee Assistance Program or short-term counselling allocation, after completion of a structured digital intervention, outside conventional service hours, during a later recurrence of difficulty, or at a moment when a person recognizes a need for support without yet knowing what form of service would be appropriate.

These intervals should not be treated as peripheral to mental-health access. For many people, they constitute a substantial part of the help-seeking journey.

The Continuity Guarantee begins from a basic reality: **human need does not begin and end according to program boundaries.** The appointment ends. The person's life continues. The funded sessions end. The relationships and circumstances that shaped the original need may continue. A structured intervention ends. Grief, loneliness, stress, uncertainty, trauma, family conflict, or other life difficulties may remain—or return later in a different form. A professional may be temporarily unavailable. The need for somewhere appropriate to turn does not disappear. The Continuity Guarantee is intended to address that gap.

What the Continuity Guarantee Would Mean in Practice

The Continuity Guarantee should not be misunderstood as a promise of unlimited psychotherapy. It would not mean unlimited automated clinical treatment. It would not mean that every form of emotional difficulty should be medicalized. And it would not imply that professional human capacity can somehow become infinite. Its meaning is narrower and more practical: **An appropriate foundational layer of support remains available even when professional care is temporarily unavailable.**

If evidence ultimately supports the proposed model, that foundational layer could be free at the point of appropriate use, available 24 hours a day and 365 days a year, multilingual, accessible through text and audio-enabled interaction, trauma-informed, privacy-preserving, disability-inclusive, geographically accessible, human-supervised, and connected to professional, community, crisis, or emergency services where required.

The public experience would be straightforward: **There is somewhere appropriate I can turn right now.**

A person should not necessarily have to wait in complete isolation simply because the appropriate professional is unavailable. They should not have to pay out of pocket for every appropriate foundational interaction. They should not automatically lose all support because a time-limited program has ended. They should not need to identify the perfect provider before receiving any assistance at all. They should not have to restart the entire support journey from zero every time formal service access pauses. The foundational safety net remains underneath the formal system. That is the practical meaning of continuity.

The Economic Foundation of the Continuity Guarantee

The Continuity Guarantee becomes more plausible because artificial intelligence changes the economics of availability. Historically, continuous support required continuous human labour. More interactions required more staff time. Longer operating hours required more shifts. Additional languages required additional human personnel. Greater demand generally required a corresponding increase in human capacity. That relationship placed a natural limit on the possibility of continuously available support.

AI changes that production relationship for some carefully bounded functions. An appropriately governed digital infrastructure can potentially provide substantial quantities of structured reflection, basic emotional support, non-clinical coaching, routine psychoeducation, service navigation, multilingual interaction, and continuity without requiring another full unit of professional labour for every additional interaction.

This does not make the system free to operate. Government would still need to fund model access, token and inference usage, computing capacity, secure infrastructure, privacy protections, cybersecurity, multilingual quality assurance, accessibility, human supervision, testing, maintenance, governance, escalation pathways, professional involvement, and independent evaluation.

The relevant difference is that these costs can scale differently from one-to-one professional labour. This is why the proposal's resource-allocation model matters. It may provide the economic architecture through which the Continuity Guarantee becomes substantially more feasible.

Instead of attempting to finance unlimited professional availability—which is neither economically nor operationally realistic—government can test whether it is possible to finance **scalable computational capacity for appropriate availability** while protecting **scarce professional human attention for complex human needs.**

The outcome would not be unlimited professional care. It would be the possibility of **continuous access to an appropriate foundational safety net within clear service and safety boundaries**.

Two Guarantees Address Two Different Forms of Scarcity

The Care Guarantee and Continuity Guarantee address related but distinct structural barriers.

Guarantee	Core Question	Principal Scarcity Addressed	Primary Public Response
Care Guarantee	**Can I obtain the appropriate professional care I need?**	Financial barriers to professional expertise	Expand equitable access to appropriate publicly supported professional care
Continuity Guarantee	**Do I have somewhere appropriate and safe to turn right now?**	Availability gaps created by finite professional capacity	Maintain a free-at-the-point-of-appropriate-use, continuously available, human-supervised foundational support layer

The Care Guarantee primarily addresses **financial scarcity**. The Continuity Guarantee primarily addresses **availability scarcity**.

These forms of scarcity frequently overlap, but they are not interchangeable. A person may be able to afford therapy and still wait months for an appropriate provider. A person may already receive publicly funded care and still need appropriate support between appointments. A person may have insurance and still struggle to find a therapist with the right specialization, language, cultural understanding, or relational fit. A person may successfully complete a short-term program and still encounter another difficulty months later. A comprehensive mental-health architecture should therefore address both.

Two Guarantees, One Continuum of Support

The two guarantees should not become separate silos. They should form a single continuum in which foundational support and professional care reinforce one another. An illustrative journey might proceed as follows:

A person recognizes a need

↓

Immediate access to an appropriate foundational support layer

↓

Emotional support, reflection, non-clinical coaching, life-skills assistance, psychoeducation, or navigation

↓

Qualified human supervision and escalation where appropriate

↓

Connection to professional care where professional care is required

↓

Psychotherapy, counselling, assessment, treatment, or other specialized human support

↓

Foundational support remains available while the person waits or between professional encounters

↓

Return to professional care whenever necessary

↓

The foundational safety net remains available after one formal episode ends

This creates a continuum in which professional care and AI-enabled support are complementary rather than competing. The person should not be forced to choose between **AI support** and **a therapist**. The appropriate objective is to use each at the point where it creates the greatest value.

Technology provides continuity within scope. Qualified professionals provide care where expertise, judgment, relationship, and accountability are required.

The Division of Responsibility Behind the Two Guarantees

The two-guarantee model becomes possible because responsibilities are deliberately distributed among different forms of capacity.

Professional Care Provides Depth

Therapists, counsellors, psychologists, social workers, physicians, psychiatrists, and other appropriate professionals provide judgment, therapeutic relationships, clinical depth, complexity management, risk assessment, specialized treatment, ethical responsibility, and genuine human presence. Their role is expertise and professional care.

AI-Enabled Infrastructure Provides Availability and Scale

Appropriately bounded AI can support immediate availability, structured reflection, basic emotional support, non-clinical coaching, routine psychoeducation, navigation, multilingual interaction, text and audio-enabled access, and continuity surrounding formal services. Its role is scalable foundational availability.

Qualified Human Supervision Provides Accountability

Human supervisors and reviewers provide governance, safety review, quality assurance, professional-boundary oversight, escalation, intervention where appropriate, and connection to human services. Their role is accountability.

Government Provides Durability

Government can provide public standards, funding architecture, privacy requirements, accessibility obligations, regulatory clarity, independent evaluation, equitable access, technology neutrality, and institutional continuity. Its role is stewardship and durability.

The architecture can therefore be summarized as:

- Professional care for depth.
- AI-enabled infrastructure for availability and scale.
- Qualified humans for accountability.
- Government for continuity and public responsibility.

From "How Many Sessions?" to "Will I Still Have Somewhere to Turn?"

Traditional mental-health financing often organizes support around a session-based logic. How many sessions are covered? How many remain? What does another session cost? What happens when the allocation ends? That structure reflects the underlying scarcity of professional human time. Where professional expertise is required, finite sessions may remain an unavoidable aspect of service delivery. But if an appropriate foundational support layer can be financed partly as infrastructure, the public-policy question can begin to shift.

Instead of asking only: How many supportive interactions can this funding arrangement purchase? government can also ask: How can we ensure that a person does not fall into complete absence of appropriate support when finite professional capacity is temporarily unavailable? This represents a shift from an exclusively episodic understanding of access toward a **continuity-oriented model**.

Professional care can remain episodic where clinically and operationally appropriate. The foundational safety net can remain available around it.

"My Country Has My Back" as a Public-Service Standard

The human meaning of the two guarantees can ultimately be expressed in a simple expectation. A person should be able to know:

- If I require appropriate professional mental-health care, inability to pay should not be the principal reason I cannot obtain it.

- If I am waiting, searching, between appointments, outside conventional service hours, or temporarily without the right professional, I will not be left with absolutely nowhere appropriate and safe to turn.

This would create a different relationship between people and public institutions. It would mean:

- professional care remains available as the appropriate destination when professional expertise is required;
- foundational support remains available during waiting periods and transitions;
- people are not required to pay for every appropriate first-line interaction;
- text and audio-enabled access can accommodate different preferences and needs;
- multilingual support can reduce linguistic barriers;
- and the entire safety net does not disappear simply because one funded program has ended.

In the most human terms: *"My country has my back."*

That phrase should not remain rhetorical. If adopted as a long-term aspiration, it should be translated into measurable standards of availability, accessibility, privacy, safety, continuity, professional connection, and public accountability.

The Two-Guarantee Model at a Glance

Element	Care Guarantee	Continuity Guarantee
Primary Need	Appropriate professional mental-health care	An appropriate place to turn when professional care is not immediately available
Primary Barrier Addressed	Affordability	Availability
Principal Scarcity	Access to professional expertise	Gaps created by finite professional human attention
Core Resource	Qualified professional capacity	Scalable foundational support infrastructure
Role of AI	Complementary; helps protect professional capacity and continuity around care	Central to scalable availability for bounded non-clinical and structured functions
Role of Professionals	Direct professional assessment, treatment, counselling, psychotherapy, specialized intervention, and judgment	Supervision, intervention, escalation, and professional care whenever required
Role of Government	Expand equitable access to appropriate professional care	Establish standards and ensure durable access to the foundational safety net if evidence supports the model
Long-Term Public Promise	**Professional care should not be denied primarily because of inability to pay**	**People should not be left with absolutely nowhere appropriate and safe to turn when professional care is temporarily unavailable**

Together, these guarantees represent a fuller understanding of accessibility.

Accessibility is not only: Can I afford care? It is also: Can I reach support? Can I reach it when I need it? Can I understand it? Can I access it in a usable format? Can I obtain support while searching for an appropriate professional? Can I return if my need re-emerges? A mature mental-health system should aspire to address all of these questions.

A National Promise That Technology May Help Make Achievable

The significance of this proposal is not that artificial intelligence has solved mental health. It has not. Nor does the proposal assume that technology can reproduce the therapeutic depth, professional judgment, accountability, and relational value of human care.

Its significance is more specific: Technology may have changed one of the economic assumptions that historically made continuous supportive availability exceptionally difficult to finance.

For the first time, government can seriously test whether availability can become substantially more abundant while professional expertise remains protected. If so, Canada would not have to choose between broader access and professional quality. It could pursue both:

- continuous foundational availability where technology can responsibly provide scale; and
- professional human care where human expertise is indispensable.

That is the deeper meaning of the two-guarantee model.

The Care Guarantee protects people from financial exclusion from professional expertise. The Continuity Guarantee protects people from complete support abandonment during gaps in professional availability.

Together, they articulate a more complete long-term public principle:

- **No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because they cannot afford it.**
- **No Canadian should be left completely unsupported while searching for, waiting for, moving between, or temporarily unable to access that care.**

Professional capacity will remain finite. Public resources will remain finite. Technology will remain imperfect. But if responsible AI, deliberate resource allocation, qualified human governance, professional care, and durable public responsibility can be combined successfully, one important constraint may no longer have to remain permanent: **The safety net does not have to disappear when professional capacity is temporarily unavailable.**

That is the continuum Canada should test—and, if the evidence supports it, work toward building.

12. Requested Government Action

I respectfully recommend that the Government of Canada establish a disciplined pathway for determining whether a **Human-Supervised Virtual Safe Space** can become a safe, equitable, sustainable, and publicly accountable component of Canada's mental-health and social-support infrastructure.

The requested action is not immediate nationwide deployment. It is to create a structured progression from: **frontline innovation → controlled demonstration → independent evidence → public standards → durable infrastructure, if justified.**

The first objective should be to determine whether the proposed architecture works, for whom it works, under what conditions it works, what risks it creates, what safeguards it requires, and whether it produces sufficient public value to justify further development.

Government should therefore treat the initiative as both a **service-design experiment and a public-policy experiment**. It should test not merely whether AI can support certain interactions, but whether a complete architecture combining computational capacity, professional human attention, human supervision, privacy, accessibility, safety, and public governance can improve access and continuity without weakening professional care.

Establish a National Human-Supervised Virtual Safe Space Demonstration

The Government of Canada should establish a **time-limited, staged, independently evaluated Human-Supervised Virtual Safe Space Demonstration Initiative** focused on evidence generation, responsible design, fiscal productivity, institutional learning, and public value.

The demonstration should test whether Canada can provide an appropriate first layer of free-at-the-point-of-appropriate-use emotional support, guided reflection, non-clinical coaching, holistic life-skills assistance, routine psychoeducation, service navigation, multilingual interaction, text and audio-enabled access, and 24/7 continuity within clearly defined boundaries.

The initiative should maintain an explicit distinction between this foundational support layer and psychotherapy, diagnosis, treatment, crisis care, emergency services, prescribing, and other regulated or specialized professional functions.

Qualified humans should remain central to supervision, escalation, intervention, quality assurance, professional-boundary oversight, ethical governance, safety review, and connection to appropriate professional or community services.

The demonstration should begin conservatively. Its first objective should be **learning rather than maximum user volume**. Government should begin with clearly bounded use cases, evaluate

performance rigorously, introduce new populations or functions only when evidence supports doing so, and retain the ability to redesign, restrict, or discontinue the model where necessary.

Test a 50/50 Resource-Allocation Hypothesis

Within a defined existing or comparable public funding envelope, government should test an approximate **50/50 allocation between AI-enabled public digital infrastructure and human professional capacity and governance**.

For demonstration purposes, approximately half of the envelope could support the infrastructure required to purchase availability and scale: AI model access, token and inference usage, computing capacity, secure infrastructure, multilingual functionality, text and audio-enabled interaction, privacy, cybersecurity, accessibility, trauma-informed design, automated safety systems, testing, maintenance, monitoring, and independent technical evaluation.

The remaining portion could support the human capacity required to purchase judgment, safety, professional depth, and accountability: therapists, counsellors, psychologists, social workers, qualified supervisors, clinical advisers, intervention and escalation specialists, safety and quality reviewers, service navigators, AI-support supervisors, cultural reviewers, professional governance, and other necessary human functions.

The proposed allocation should **not** be treated as a permanent national formula. It is a demonstration hypothesis. The long-term balance, if any, should be determined empirically.

The central economic question should be: **Can Canada use computational capacity to create greater abundance in appropriate availability while protecting scarce professional human attention for complex human needs?**

Measure and Reallocate Professional Human Attention

The demonstration should explicitly examine how professional human attention is currently used and whether selected functions can be redesigned without compromising safety or quality. Many structured supportive-service environments contain functions that may be repetitive, informational, protocol-guided, standardized, or highly text-based. These can include routine psychoeducation, structured exercises, guided reflection, standard follow-up, basic navigation, standardized goal-setting, journaling prompts, and similar activities. Government should evaluate these functions individually rather than assuming that either all should remain human-delivered or all should become automated.

The appropriate question is: **Which functions require professional expertise, which can be responsibly AI-assisted, which may be suitable for bounded automation, and when must responsibility return to a qualified human?**

The objective should not be merely to reduce labour expenditure. The more important objective is to determine whether professional capacity released through appropriate automation is actually redirected toward complex cases, therapeutic relationships, trauma-informed care, risk assessment, ethical reasoning, difficult navigation, human intervention, escalation, supervision, and other functions in which professional expertise creates substantially greater value.

The key productivity question should therefore be: **Where did the released human attention go?** A professional hour should count as a productivity gain only if its reallocation produces meaningful public value.

Protect and Strengthen Professional Employment

Workforce transition should be designed into the initiative from the outset rather than treated as an incidental consequence of technology adoption. The demonstration should follow the principle of **job evolution, not job elimination.**

Helping professionals should remain central to system governance, professional-boundary development, safety, escalation, complex care, quality assurance, cultural review, human intervention, supervision, and the continuing determination of which functions are appropriate for automation. Government should also assess emerging professional roles such as AI-support supervision, trauma-informed AI evaluation, AI-safety review, clinical AI advising, professional-boundary oversight, quality assurance, culturally responsive AI review, human-machine service design, digital mental-health governance, and complex-care specialization.

Workforce evaluation should examine whether professionals experience greater use of their expertise, meaningful employment, manageable workload, appropriate autonomy, adequate training, stronger complex-care capacity, and sustainable professional roles.

The intended outcome is not fewer humans in the system. It is **better use of human expertise within the system**.

Establish Trauma-Informed Safety Standards Before Scale

Safety requirements should be established before broad deployment rather than added reactively after problems emerge. At minimum, the demonstration should require trauma-informed design, transparent AI identity, clear limits on system scope, explicit differentiation between AI-enabled support and professional care, prohibition on unauthorized diagnosis or prescribing, qualified human oversight, graduated escalation pathways, human intervention where appropriate, privacy by design, cybersecurity protections, understandable consent, professional-boundary review, complaint and correction mechanisms, and independent evaluation.

The governing principle should be: **The more scalable the technology becomes, the stronger the human accountability surrounding it must become.**

Safety should not be treated as a barrier to accessibility. For emotionally sensitive services, safety is a precondition of meaningful access. A system that is technically available but psychologically unsafe, opaque, culturally inappropriate, or poorly governed is not genuinely accessible.

Make Privacy a Core Infrastructure Requirement

Emotionally sensitive conversations should receive a heightened standard of privacy protection. The initiative should require data minimization, clear purpose limitation, meaningful and understandable

consent, appropriate retention limits, role-based access, transparent human-review rules, strong cybersecurity, auditability, complaint pathways, and strict controls on secondary use.

Private disclosures should not be sold, used for advertising, transformed into unrelated commercial profiles, or converted into individualized government intelligence. Users should be able to understand in plain language what information is collected, why it is collected, what information is unnecessary, who can access it, when human review may occur, how long it may be retained, and what alternatives remain available.

A national virtual safe space cannot earn trust if privacy is treated as a secondary compliance matter. Privacy should therefore be regarded as **infrastructure, not fine print.**

Make Accessibility Fundamental to the Public Standard

Accessibility should be built into the service from the beginning. The demonstration should evaluate multilingual interaction, text access, audio-enabled interaction, screen-reader compatibility, simplified-language options, low-bandwidth functionality, accessible navigation, disability-inclusive design, neurodivergent accessibility, cultural responsiveness, geographic reach, and usability in rural and remote communities.

Government should also ensure that people who cannot or do not wish to use AI retain appropriate human or non-AI alternatives. Universal availability is not meaningful if the service is technically present but practically unusable.

Accessibility should therefore be measured not simply by whether features exist, but by whether people can actually understand, navigate, trust, and use them. The public standard should communicate an important value: **You belong here.**

Connect the Foundational Layer Directly to Human and Professional Care

The virtual safe space should be designed explicitly as a **bridge**, not as a destination for every user. It should complement rather than compete with licensed professional care, specialized mental-health services, community supports, social services, crisis resources, and emergency services. The demonstration should therefore include reliable pathways for navigation, referral, escalation, human intervention, and continuity while users await further care.

Referral should not automatically terminate all support. If a person is connected to an appropriate therapist but the next available appointment is several weeks away, the foundational support layer should remain available within its defined scope. This is one of the most important distinctions in the model. The goal is not to create another isolated program. It is to create a **continuum around professional care**.

Test the Care Guarantee and Continuity Guarantee Together

The demonstration should examine how two complementary long-term public aspirations can reinforce one another.

The Care Guarantee

No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.

The Continuity Guarantee

No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.

The Care Guarantee concerns access to professional expertise. The Continuity Guarantee concerns the gaps surrounding that expertise. The demonstration should assess whether scalable foundational support can strengthen the broader professional system by reducing complete gaps, improving navigation, and protecting professional capacity for needs that genuinely require professional care.

The long-term system should aspire to answer *yes* to both questions:

- **Can I obtain the professional care I need?**
- **Do I have somewhere appropriate and safe to turn right now?**

Use the Demonstration to Establish Legal and Governance Clarity

The initiative should be used to identify and address legal and regulatory uncertainty surrounding AI-supported emotionally sensitive services.This work should include professional scope, supervision, privacy, consent, human review, data governance, cross-jurisdictional responsibility, complaint and redress mechanisms, cultural responsiveness, AI-specific accountability, professional-boundary rules, escalation standards, and appropriate allocation of liability.

The objective should not be to exempt any actor from legitimate accountability. It should be to replace fragmented, organization-by-organization interpretation with **clearer and more coherent public standards**.

Government should help clarify:

- **what AI may do**
- **what AI may not do**
- **when human involvement is required**

- **what minimum standards apply**
- **how responsibility is allocated when something goes wrong.**

Clear governance can protect users, professionals, responsible innovators, service organizations, vendors, and public institutions simultaneously.

Protect Responsible Social Innovation Through Proportionate Accountability

Government should develop a governance environment in which nonprofits and social entrepreneurs can identify public needs, test innovative responses, and generate evidence without being left to carry indefinitely disproportionate legal, financial, ethical, technological, operational, and professional burdens.

Social entrepreneurs should remain accountable for negligence, misrepresentation, unsafe practices, privacy failures, professional-boundary violations, and other genuine harms where applicable.

Public-interest motivation should not create immunity. At the same time, essential public innovation should not depend on one or two founders permanently absorbing the systemic risk associated with a service that addresses a broader public gap.

The principle should remain: **Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.**

Where rigorous evidence shows that an innovation serves a widespread, durable, and legitimate public need, responsibility should gradually move from individual sacrifice toward institutional accountability.

Treat Founder and Vendor Dependency as a Public-Risk Issue

Excessive dependence on one founder, nonprofit, vendor, donor, platform, proprietary model, or funding arrangement should be treated as a public-risk concern. Essential infrastructure should avoid avoidable single points of failure.

Founders may leave. Organizations may close. Grants may expire. Technology may become obsolete. Vendors may change strategy. Procurement arrangements may change. Government therefore should not promise that one particular platform or organization will operate indefinitely. It should protect the underlying **technology-neutral capability**: **Free-at-the-point-of-appropriate-use, continuous, multilingual, trauma-informed, accessible, human-supervised emotional-support infrastructure connected to appropriate professional and human care.**

The provider may change. The AI model may change. The governance arrangements may evolve. But if the capability has been established as a legitimate public good, the public promise should be designed to survive those transitions: **There will continue to be somewhere appropriate and safe to turn.**

Measure Outcomes Rather Than Activity

The demonstration should not define success primarily through the number of AI conversations, logins, tokens consumed, downloads, repeat engagements, or total user volume. Those are activity measures. Activity is not public value.

The evaluation framework should instead measure accessibility, continuity, equity, safety, privacy, cultural responsiveness, disability inclusion, navigation success, connection to appropriate professional care, professional workload, professional capacity released through appropriate automation, workforce sustainability, user trust, user dignity, escalation quality, fiscal productivity, governance clarity, legal clarity, cybersecurity, and unintended consequences.

The central question should be: **Did the complete human-AI architecture improve the experience of people seeking support and strengthen the public system around them? not: Did people use the technology?**

Test Privacy-Preserving Frontline Learning

The demonstration should assess whether appropriately de-identified, aggregated, privacy-preserving patterns can help government identify recurring public-service barriers more quickly. Potential areas of learning could include persistent waitlist problems, underserved languages, rural and regional disparities, poor navigation pathways, accessibility failures, recurring provider-matching difficulties, and forms of unmet need not readily visible through existing reporting.

Such analysis must remain subject to strict purpose limitation, minimization, governance, and independent oversight. Private emotional conversations should never become individualized surveillance.

The principle should remain: **The purpose is not watching people at scale. It is understanding systemic barriers at scale.**

Where aggregate patterns identify a problem, AI analysis should trigger human investigation rather than replace it. **AI may identify patterns. Human leaders must still listen.**

Require Independent Evaluation and Transparent Public Reporting

Evaluation should be conducted by independent researchers and experts rather than relying solely on vendor, contractor, or program self-assessment. The evaluation should address safety, privacy, cybersecurity, bias, accessibility, cultural responsiveness, user outcomes, professional outcomes, workforce effects, escalation quality, fiscal performance, legal and governance clarity, public trust, and unintended consequences.

Where consistent with privacy, security, and legitimate confidentiality requirements, findings should be reported publicly. Government should establish predefined decision gates for continuation, redesign, expansion, restriction, further testing, or termination. The initiative should advance because the evidence is strong, not merely because the technology is popular or politically attractive.

Evaluate the 50/50 Hypothesis Empirically

The approximate 50/50 funding allocation should itself be an object of study. Government should assess how much funding is required for computational infrastructure; the cost of model access, inference, and security; the level of professional supervision needed; how escalation affects staffing; how multilingual expansion changes cost; how accessibility requirements affect infrastructure investment; how much professional capacity is genuinely released; and how different funding ratios affect safety, quality, trust, and outcomes.

The optimal long-term allocation may be 40/60, 50/50, 60/40, or another configuration entirely. Different contexts may require different ratios. The purpose is not to defend the number. It is to identify **the allocation that produces the greatest meaningful public value while preserving safety, trust, professional quality, equity, and sustainable human employment.**

Determine Whether the Capability Merits Public-Infrastructure Status

If the evidence is positive, government should examine whether the underlying capability warrants treatment as a durable component of Canada's digital and social infrastructure. That assessment should address long-term financing, governance, interoperability, professional oversight, privacy and cybersecurity standards, procurement, technology neutrality, accessibility, public reporting, accountability, evaluation, and the mechanisms required to preserve continuity when providers or technologies change.

The objective should not be to institutionalize one product. It should be to institutionalize, if justified, **the public capability**. That capability is an appropriate first layer of support that can remain available before, between, and alongside professional services and does not disappear simply because a particular program, provider, funding cycle, or platform ends.

Develop a Technology-Neutral Canadian Public Standard

If the demonstration produces sufficiently strong evidence, Canada should consider developing a technology-neutral public-interest standard for human-supervised virtual safe-space infrastructure. Such a standard could establish minimum expectations for privacy, professional boundaries, trauma-informed design, accessibility, multilingual performance, human supervision, safety architecture, escalation, user rights, evaluation, transparency, interoperability where appropriate, complaint mechanisms, and vendor accountability.

The standard should define the capability rather than the vendor. No single AI company, nonprofit organization, technology platform, model, or founder should permanently determine the architecture.

The governing principle should be: **Government should protect the promise, not the platform.**

Position Canada for International Leadership in Responsible AI-Enabled Social Infrastructure

If the demonstration succeeds, Canada should document and share the governance, workforce, safety, accessibility, privacy, funding, and evaluation lessons generated through the initiative. Many jurisdictions face similar pressures: rising demand, workforce shortages, rural and remote access challenges, language diversity, finite program models, constrained budgets, workforce burnout, and uncertainty about how AI should be integrated into emotionally sensitive public services. Canada could contribute an alternative to a global AI conversation dominated by labour replacement.

The Canadian lesson should not be: Canada replaced therapists with AI. It should be: **Canada preserved professional human care while using responsible AI and strategic public investment to ensure that people were less likely to be left without an appropriate first place to turn.**

That would represent leadership not simply in technology adoption, but in **human-centred public institutional design**.

Build Toward a Durable National Promise

The long-term purpose of these actions is larger than the demonstration itself. If evidence supports the model, Canada should work toward a system in which people can reasonably expect two things.

1. **If I require appropriate professional mental-health care, inability to pay should not be the primary reason I cannot obtain it.**

2. **If I am waiting, searching, between services, outside conventional office hours, or temporarily unable to reach the appropriate professional, I will not be left with absolutely nowhere appropriate and safe to turn.**

The resulting foundational capability should aspire to be free at the point of appropriate use, available 24 hours a day and 365 days a year, multilingual, accessible through text and audio-enabled interaction, trauma-informed, privacy-preserving, disability-inclusive, human-supervised, and connected to appropriate professional, community, crisis, or emergency services when required.

This is not a promise the Government of Canada should make before the evidence exists. It is the **public objective the demonstration should determine whether Canada can responsibly achieve**.

Requested Government Action at a Glance

Government Action	Purpose
Establish a controlled national demonstration	Test the complete human-AI service architecture before considering broader deployment
Test the approximate 50/50 funding hypothesis	Determine whether separating availability from expertise can increase public value
Measure and reallocate professional attention	Protect scarce professional expertise for complexity, judgment, relationships, and intervention
Fund workforce transition	Support job evolution, training, supervision, governance, and meaningful professional roles
Establish trauma-informed safety standards	Ensure that greater scale is accompanied by stronger human accountability
Make privacy foundational	Protect sensitive disclosures and create conditions for public trust
Build accessibility from the outset	Make multilingual, disability-inclusive, rural, text, audio, and low-bandwidth access meaningful in practice
Connect the foundational layer to professional care	Ensure that AI functions as a bridge rather than a substitute for appropriate human services
Test the Care and Continuity Guarantees together	Address both affordability and availability within one continuum
Clarify legal and governance responsibilities	Establish clearer scope, accountability, supervision, escalation, privacy, and liability frameworks
Protect responsible social innovation	Allow innovators to demonstrate solutions without carrying systemic public burdens indefinitely
Reduce founder and vendor dependency	Remove avoidable single points of failure
Measure public value rather than volume	Evaluate meaningful outcomes instead of technology activity
Test privacy-preserving frontline learning	Identify recurring systemic barriers without individualized surveillance
Require independent evaluation	Generate credible evidence and public accountability
Evaluate public-infrastructure status	Determine whether the capability merits durable institutional support
Develop technology-neutral public standards	Preserve the capability rather than dependence on one provider
Share evidence internationally	Position Canada as a leader in human-centred AI-enabled social infrastructure

The Central Government Decision

The decision before government is ultimately not whether artificial intelligence should replace mental-health professionals. The more consequential question is: **Can Canada reorganize finite public resources so that technology creates greater continuity where scalable availability is appropriate, while scarce professional human attention is protected for the moments in which judgment, complexity, therapeutic depth, safety, ethical responsibility, and genuine human connection make human beings irreplaceable?**

If the answer is no, the model should not proceed in its proposed form. If the answer is yes, the implications extend beyond a single demonstration.

Canada could begin moving from a system in which supportive access is constrained almost entirely by sessions, staffing, geography, language, and operating hours toward one in which **professional care remains deeply human while an appropriate foundational safety net can remain continuously available around it**.

- Social entrepreneurs could remain free to innovate without carrying essential public burdens indefinitely.
- Helping professionals could spend more of their scarce time on the work that genuinely requires their expertise.
- Government could obtain greater access and continuity from finite resources while strengthening public accountability rather than weakening it.
- Canadians could have greater confidence that temporary limits in professional capacity do not automatically translate into complete absence of support.

The requested action is therefore deliberately disciplined:

- **test the model;**
- **establish safeguards before scale;**
- **measure the complete system independently;**
- **involve professionals and people with lived experience from the beginning;**
- **evaluate the 50/50 hypothesis rather than assuming it;**
- **follow the evidence;** and, only if the evidence warrants it,
- **transform a demonstrated innovation into durable, technology-neutral public capability.**

The long-term public promise is equally clear:

- **Professional care when professional care is needed.**
- **An appropriate place to turn when that care is not immediately available.**
- **No per-interaction out-of-pocket barrier to the foundational safety net.**

- **Multilingual, accessible, text and audio-enabled entry points.**
- **Qualified human supervision and accountability.**
- **Continuity beyond one appointment, one program, one provider, one platform, or one founder.**
- **At the broadest level, a Canadian public system capable of demonstrating through its institutions—not merely through rhetoric—that a country should have its people's backs.**

13. Long-Term National Impact: Building a Lasting Canadian Legacy of Care

The significance of this proposal extends far beyond the immediate delivery of mental-health and emotional-support services. If the proposed model proves safe, useful, trusted, equitable, professionally sustainable, accessible, and fiscally responsible, its most important contribution may not be the creation of one additional digital service or the success of one particular technology. Its deeper contribution may be the demonstration of a different way to organize public healthcare capacity: one in which scalable infrastructure, protected professional human expertise, public financing, social innovation, and institutional responsibility are deliberately connected around the practical needs of the person seeking care.

The proposed **Human-Supervised National Virtual Safe Space** should therefore be understood simultaneously as a potential public capability and as a concrete policy experiment. At the immediate level, it would test whether Canada can establish a free-at-the-point-of-appropriate-use, multilingual, accessible, trauma-informed, human-supervised foundational layer that remains available before, between, and alongside professional care. At the broader level, it would provide a bounded and measurable environment in which government could test a possible healthcare budgeting-reallocation model: whether different forms of public capacity can be funded according to the functions they are best suited to perform while still operating as one coherent continuum of support.

This proposal did not begin with artificial intelligence. Its intellectual and practical roots lie in more than a decade of frontline observation, advocacy, social innovation, service delivery, experimentation, and sustained listening to people from different socioeconomic circumstances describe what access to care actually looked like from where they stood. Over time, it became increasingly clear that affordability, while fundamental, is only one dimension of accessibility. People also wait. They search. They encounter provider mismatch, limited operating hours, language barriers, cultural barriers, geographic barriers, disability-related barriers, fragmented navigation, short-term program limits, and the emotional burden of repeatedly beginning again when an earlier attempt to find help does not work.

Those experiences gradually revealed a distinction that became central to the proposal: **formal coverage does not automatically produce practical accessibility**. A service can technically exist and still be difficult to reach. A person can be formally covered and still spend months waiting. A referral can be clinically appropriate while the person remains unsupported before the next appointment. A professional can exist somewhere in the healthcare system while being inaccessible because of geography, language,

disability, specialization, scheduling, cost, or relational mismatch. A program can conclude successfully from an administrative perspective while the circumstances that brought the person into care continue long after the program ends.

For A Safer Space (FASS) was an early practical response to this gap. It sought to create a free human bridge for people who might otherwise have nowhere to turn while they were waiting, searching, navigating, or trying to understand what kind of support they actually needed. FASS demonstrated the importance of accessible human support and the value of continuity around professional care. It also revealed an unavoidable structural constraint: human beings can provide extraordinary compassion, wisdom, judgment, accountability, therapeutic depth, and presence, but human availability is finite.

A counsellor can speak with only one person at a time. A therapist has only so many professional hours in a day. A supervisor can responsibly oversee only so much work. Another language requires another form of human capacity. Longer service hours require additional staffing. Greater demand creates additional needs for recruitment, training, administration, coordination, supervision, and quality assurance. When the financial barrier is removed for the person receiving support, the underlying cost of human labour does not disappear. Someone else still has to carry it.

That realization eventually contributed to the development of **FASSLING™**, which explored another dimension of the same access problem. The relevant question was not whether artificial intelligence could replace psychotherapy, counselling, psychology, social work, medicine, or professional judgment. It was whether appropriately bounded forms of emotional support, guided reflection, non-clinical coaching, psychoeducation, structured exercises, service navigation, multilingual interaction, and continuity could become more continuously available without requiring one additional professional to become personally available for every additional appropriate interaction.

Taken together, FASS and FASSLING™ therefore provide more than the history of two social innovations. They offer practical cases through which a larger healthcare-system question became visible. FASS demonstrated the value and the limitations of a free human bridge. FASSLING™ explored whether technological capacity could extend certain forms of appropriate availability beyond some of those human-capacity limits. Together, they led to a broader conclusion: **the healthcare-access problem is not only about how much capacity exists; it is also about how different forms of capacity are allocated, financed, connected, and used**.

The proposed National Virtual Safe Space is the public-infrastructure expression of that insight. It is not intended merely to address a temporary shortage of appointments or capitalize on a short-term technological opportunity. It responds to a structural reality that is unlikely to disappear even if Canada substantially expands professional mental-health capacity: **professional human attention will remain finite, while the need for appropriate support can arise at any time**. Canadians will continue to experience grief, stress, trauma, loneliness, relationship difficulties, family conflict, employment transitions, caregiving pressures, financial uncertainty, migration, illness, aging, loss, identity questions, and unexpected disruption across the life course. Professional models will evolve and technologies will change, but the underlying human need for somewhere appropriate and safe to begin will remain.

Even a substantially expanded professional system cannot make every qualified professional continuously available to every person across every location, language, communication mode, specialty, and hour of the day. The National Virtual Safe Space would therefore seek to address a specific form of

scarcity: the periods in which appropriate professional care is not immediately available, but the person's need for some appropriate form of support has not disappeared.

Its purpose would not be to replace psychotherapy, counselling, psychiatric care, social work, medicine, crisis response, emergency services, or specialized professional intervention. Its purpose would be to create a foundational layer around those services so that temporary professional scarcity does not automatically become complete absence of support. A person waiting for an appointment should not necessarily have nothing. Someone searching for a better professional fit should not necessarily have to navigate that search entirely alone. A newcomer should not have to understand the entire healthcare system before having somewhere appropriate to begin. A person whose short-term program has ended should not be treated as though the end of the program means the end of the circumstances affecting their life. **The program may end; the person's life does not.**

This is why the National Virtual Safe Space matters as potential public infrastructure. Its value would lie not only in the number of interactions it could provide, but in whether it could make continuity itself more dependable. Instead of appropriate foundational support being determined almost entirely by geography, conventional office hours, the duration of individual programs, the availability of one specific provider, or the person's ability to pay for another interaction, continuity could become a more deliberate part of public healthcare architecture.

The proposal also uses the National Virtual Safe Space as a concrete environment in which a broader budgeting-reallocation hypothesis can be tested. Historically, many different forms of healthcare capacity have been purchased through human labour. More users required more professional or para-professional hours. Longer operating hours required more staffing. More languages required more human capacity. Broader geographic coverage often required parallel organizational and workforce expansion. In practical terms, availability and expertise were frequently purchased through the same scarce input.

Emerging technological capacity creates an opportunity to reconsider that relationship for some carefully bounded functions. This does not mean that artificial intelligence makes healthcare costless, nor does it mean that digital infrastructure is inherently inexpensive. Safe public infrastructure itself requires substantial and continuing investment in computation, model access, secure hosting, privacy architecture, cybersecurity, multilingual quality, accessibility, safety systems, monitoring, governance, qualified human supervision, escalation pathways, maintenance, and independent evaluation.

The relevant distinction is therefore not between "expensive human care" and "cheap AI." It is between **different categories of public expenditure that purchase different forms of capacity**. Some expenditure purchases professional expertise. Some purchases availability. Some purchases continuity. Some purchases infrastructure. Some purchases safety and supervision. Some purchases navigation. Some purchases accessibility. Some purchases governance. The broader healthcare-budgeting question is whether those different functions must always remain bundled together or whether, in appropriate contexts, government can finance them more deliberately.

This is the purpose of the proposed approximate **50/50 funding hypothesis**. The ratio itself is not the policy objective. It is an experimental starting point. One portion of a defined or comparable funding envelope could support scalable digital infrastructure that purchases availability, continuity, multilingual reach, accessibility, navigation, structured support, computation, and scale. Another portion could protect professional human capacity and governance for judgment, therapeutic depth, complex care, safety, intervention, supervision, ethical accountability, and genuine human connection.

The 50/50 allocation should not be interpreted as a recommendation that half of Canada's mental-health budget, much less half of the overall healthcare budget, should permanently be allocated to artificial intelligence. It is a testable hypothesis. A population with greater complexity may require substantially more professional capacity. A lower-risk and highly structured use case may be able to rely more heavily on scalable infrastructure. Different languages, disability needs, cultural contexts, service environments, professional requirements, and risk levels may require entirely different balances.

If the evidence demonstrates that 50/50 is the wrong allocation, the allocation should change. The enduring value of the demonstration would not be the validation of one percentage. It would be the development of a more disciplined **resource-allocation method**: identify what form of capacity each function genuinely requires, protect scarce human expertise where it is indispensable, use scalable infrastructure where appropriate abundance can be created responsibly, connect those forms of capacity into a coherent continuum, and evaluate the complete architecture according to the public value it produces.

This distinction matters because the scarce resource being protected by the proposal is not only public money. It is also **professional human attention**. A therapist's hour can only be spent once. A counsellor, psychologist, social worker, physician, or other professional who spends that hour performing a highly standardized function cannot simultaneously use the same hour with someone whose circumstances require complex assessment, trauma-informed judgment, ethical reasoning, therapeutic depth, risk management, specialized intervention, or sustained human presence.

The true cost of professional time therefore includes an opportunity cost. The question is not simply what a professional hour costs financially. It is also what society gives up when that hour is used for one function rather than another. For that reason, this proposal should not be interpreted as a labour-reduction agenda. Its objective is the opposite: to recognize that **professional human attention may be too valuable to use inefficiently**.

Artificial intelligence should therefore not be evaluated primarily according to how many professional tasks it removes. A more meaningful measure is whether technological capacity can make certain appropriate forms of availability more scalable while enabling scarce professional expertise to become more available where human expertise genuinely matters most. The underlying principle is simple: **AI should not replace human care. It should protect human attention for the moments when human care matters most.**

If the National Virtual Safe Space proves effective, its value could also extend across the life course rather than being confined to one episode of need. A student might first use the service while experiencing academic stress or uncertainty about where to seek help. Years later, the same person might return during a career transition, relationship difficulty, period of grief, or caregiving responsibility. They might use professional psychotherapy during one stage of life while relying on the foundational layer for appropriate continuity between appointments. Much later, they might return while navigating aging, bereavement, family responsibility, retirement, or another major transition.

A newcomer to Canada might use multilingual navigation while learning how Canadian health and social-service systems operate. A person living in a rural or remote community might use the service when the appropriate professional cannot be reached immediately. Someone searching for a therapist with the right cultural or linguistic understanding might remain connected to a foundational layer while continuing that search.

This is what distinguishes infrastructure from a short-term intervention. A program is generally designed to address a defined need for a defined period. Infrastructure creates **availability that can remain relevant as needs change**. The potential benefit is therefore not merely continuity between appointments, but continuity across different stages of life.

The timing of access may also matter. People may delay seeking help because they do not know where to begin, cannot immediately afford private care, are uncertain about what type of service they need, face language or accessibility barriers, or have become discouraged after an unsuccessful attempt to find an appropriate provider. Some may remain unsupported until distress becomes substantially more difficult to manage.

An always-available foundational layer cannot prevent every crisis or resolve every underlying problem, and this proposal should not imply otherwise. Its contribution may be more modest but still meaningful: it may reduce the friction involved in taking an earlier first step. A person might use the service to organize their thoughts, receive basic psychoeducation, understand available options, prepare for a professional appointment, remain engaged while on a waitlist, or continue searching after an unsuccessful provider match. At population scale, these relatively modest functions could contribute to earlier help-seeking, better mental-health literacy, more informed use of professional services, and fewer prolonged periods of complete unsupported distress.

These outcomes should be evaluated rather than assumed. If demonstrated, however, they would represent an additional form of preventive and social value.

The broader ambition of the proposal is ultimately larger than mental-health continuity. It is to contribute to a more complete understanding of **universal healthcare itself**. Universal healthcare has traditionally focused, appropriately, on coverage and financial protection. But formal coverage does not necessarily guarantee practical access.

A person can be covered and still wait. A person can be covered and still be unable to find the right professional. A person can be covered and still face geographic, linguistic, cultural, disability-related, scheduling, or navigational barriers. A person can be covered and still fall into the spaces between programs. A service can be publicly funded while the person who needs it still cannot reach it at the right time.

The next evolution of universal healthcare should therefore ask not only whether a person is covered, but whether that person can **actually reach the appropriate care**. The aspiration is to move, as far as responsibly and realistically possible, from **universal coverage toward universal practical accessibility**.

That means a healthcare system in which financing, professional expertise, technology, infrastructure, navigation, and continuity are designed around the practical experience of the person seeking care. It means reducing unnecessary out-of-pocket barriers where appropriate while refusing to assume that every access problem can only be solved by adding more funding to an unchanged delivery architecture. It means asking not only how much is being spent, but what that spending is actually purchasing and whether those resources could be organized more intelligently.

It also means asking where professional expertise creates the greatest value, where technology can responsibly create appropriate abundance, where people are lost between services, where systems

duplicate effort, and how different components can be connected so that the person experiences something closer to one coherent continuum rather than a collection of disconnected programs.

The deeper aspiration behind the proposal is therefore not technological. It is human.

No healthcare system can eliminate every waitlist, make every specialist continuously available, remove every provider mismatch, or eliminate every form of scarcity. Governments do not possess unlimited resources. Professional human attention will remain finite. Technology will remain imperfect.

But finite capacity does not necessarily have to result in complete abandonment.

Scarcity does not always have to become abandonment.

A person may have to wait for specialized professional care, but they should not necessarily have to wait with absolutely nowhere appropriate to turn. A program may end, but the person's life does not. A professional may temporarily be unavailable, but the entire support continuum does not necessarily have to disappear.

The goal is therefore not to promise an impossible system in which every service is continuously available to every person. It is to design a healthcare-delivery system that, **as far as responsibly and realistically possible, leaves fewer people behind because of the way its resources and services are structured**.

This is the larger purpose of the National Virtual Safe Space demonstration. It is not simply a national chatbot. It is not simply another mental-health program. It is not primarily an AI-adoption initiative. It is a test of whether Canada can deliberately connect **scalable availability, protected professional expertise, qualified human supervision, public financing, accessibility, continuity, and institutional responsibility** into a more complete healthcare architecture.

Perhaps the most significant long-term effect would therefore be institutional rather than technological. Public institutions shape how people understand their relationship with society. A country communicates its commitments not only through legislation, policy statements, budgets, and strategies, but through the capabilities people can reliably access during difficult moments.

A durable National Virtual Safe Space could communicate a simple principle: temporary scarcity should not automatically become complete abandonment. A person should not require substantial financial means merely to have somewhere appropriate to begin. The end of conventional office hours should not necessarily mean the disappearance of every form of appropriate support. An unsuccessful search for the right professional should not automatically mean a return to isolation. The completion of a finite program should not imply that the person no longer deserves any appropriate place to turn.

In the simplest human terms, the desired experience remains:

"My country has my back."

That phrase should not be interpreted as a promise that government can eliminate suffering, abolish every waitlist, or guarantee a particular professional appointment on demand. Its meaning is institutional: **the**

public system has been deliberately designed so that temporary scarcity does not automatically become complete abandonment.

If successful, the initiative could contribute to a distinctive Canadian standard: **professional care when professional care is required, and appropriate foundational support when professional care is not immediately available**. Such a standard would combine equitable access to professional care, responsible technological capacity, qualified human governance, privacy protection, accessibility, cultural responsiveness, multilingual availability, workforce sustainability, technology neutrality, evidence-based funding allocation, and public accountability.

This would not fundamentally be a technological model. It would be a **healthcare-delivery and public-service model**. Its defining feature would be the rejection of a false choice between professional human care and technological scale. Canada would instead test whether certain appropriate forms of availability can become more abundant while professional expertise remains protected for the situations in which judgment, relationship, responsibility, specialized knowledge, intervention, and genuine human presence are indispensable.

The measure of success would therefore not be how much automation entered the system. It would be whether practical access and continuity improved **without diminishing human care**.

The structural pressures motivating this proposal are not unique to Canada. Governments around the world face limited professional capacity, rising demand, long waitlists, rural inequalities, language barriers, workforce strain, fragmented service pathways, short-term program models, fiscal pressures, and growing questions about how artificial intelligence should be used in human services.

Much of the international conversation asks: **What can AI replace?** Canada has an opportunity to test a more useful question: **How can new forms of technological capacity be used to create appropriate availability while scarce professional human expertise is protected for the functions in which human beings remain indispensable?**

That is not merely an AI-policy question. It is a resource-allocation question.

If new technologies change the scarcity assumptions under which older delivery systems were designed, governments gain an opportunity to reconsider how different forms of capacity are financed and organized. The relevant question becomes not whether automation is technologically possible, but whether system redesign produces greater human and public value.

Canada’s strongest international leadership opportunity would therefore not be to claim that it possesses the most advanced conversational AI. Such a claim would quickly become obsolete. The more durable opportunity is leadership through **responsible institutional design**.

If Canada can demonstrate how free public access, continuous availability, multilingual interaction, accessibility, privacy, trauma-informed design, professional governance, appropriate escalation, workforce evolution, deliberate resource allocation, technology neutrality, and independent evaluation can be integrated into one accountable architecture, it could offer internationally relevant evidence about how technological capacity can strengthen social protection without diminishing professional human care.

Other jurisdictions may use different technologies. They may have different financing models, professional structures, constitutional arrangements, regulatory systems, and cultural expectations. They may choose very different allocations between technological infrastructure and human professional capacity. That is precisely why the transferable contribution should not be one platform or one funding ratio.

The transferable contribution is the underlying method: **identify what type of capacity each function genuinely requires, protect human expertise where it is indispensable, use scalable infrastructure where it can responsibly create appropriate abundance, connect the parts into one continuum, and evaluate the complete architecture according to human outcomes and public value**.

A successful National Virtual Safe Space demonstration could therefore generate evidence on questions that governments around the world will increasingly need to answer: which functions are suitable for technological assistance, which should remain human-led from the outset, how much human supervision different activities require, how escalation should be designed, what privacy protections are necessary for emotionally sensitive interactions, how multilingual and culturally responsive performance should be evaluated, how professional roles can evolve rather than simply be displaced, how public funding should be allocated between computational infrastructure and professional human capacity, and how successful social innovation can move from founder-led experimentation toward durable public capability.

Perhaps most importantly, the demonstration could test whether technology-supported services **strengthen people's connection to human care rather than weaken it**.

The National Virtual Safe Space should function as a bridge, not a closed destination.

The most consequential outcome of this proposal would therefore not be the number of interactions generated during the demonstration, whether FASSLING™ remains the technology used, whether one particular AI model survives, or whether the original 50/50 funding ratio proves correct. A meaningful legacy would be the creation of a public capability—and a resource-allocation method—that remains valuable long after the original technology, procurement arrangement, funding formula, or founding organization has changed.

If successful, future Canadians could inherit an architecture in which professional care is protected, financial barriers are reduced, appropriate foundational support remains available around finite formal encounters, multilingual and accessible support becomes more normal, technology expands human capacity rather than merely displacing labour, privacy and dignity remain central, and successful social innovation can mature into durable public capability without requiring founders to become permanent infrastructure providers.

They could also inherit a broader principle of healthcare financing: **public healthcare should be evaluated not only by how much money is spent, but by how effectively those resources are converted into practical access, continuity, protected professional capacity, safety, equity, and meaningful health outcomes**.

That would be the larger institutional legacy of the demonstration.

The ultimate measure of long-term success is therefore not whether one technology survives. It is whether the **public promise** survives.

Technology will change. AI models will change. Providers will change. Governments will change. Professional standards will evolve. Funding structures may change. But the underlying aspiration can remain: if a person requires appropriate professional care, financial hardship should not be the principal reason they cannot obtain it; and if that professional care is not immediately available, the person should not be left with absolutely nowhere appropriate and safe to turn.

That commitment can outlast a budget cycle. It can outlast a government. It can outlast a vendor. It can outlast FASS. It can outlast FASSLING™. It can outlast the technology through which it was first demonstrated.

If Canada can establish, through rigorous evidence, that this type of architecture can be delivered safely, equitably, professionally, accessibly, and sustainably, the country could demonstrate something more consequential than another technological innovation. It could demonstrate how a specific **FASS- and FASSLING™-inspired National Virtual Safe Space** became the testing ground for a broader healthcare-budgeting idea: that different forms of capacity can be funded, connected, and deployed according to where they create the greatest meaningful health value.

Mental health would be the demonstration environment. The National Virtual Safe Space would be the practical public-infrastructure case. The approximate 50/50 allocation would be the initial testable budgeting hypothesis.

But the larger ambition would be **healthcare-system redesign**.

Universal healthcare should not merely mean that everyone is covered. It should mean that the system is designed so that people can actually reach the care they need.

And the ultimate aspiration is a healthcare-delivery system that, as far as responsibly and realistically possible, **leaves no one behind because the architecture of the system failed to connect them to the right care**.

That could become the enduring legacy of the initiative: not simply a new mental-health service, but a stronger and more dependable model of universal healthcare—one in which innovation, fiscal responsibility, professional expertise, institutional accountability, and human dignity are organized around the same purpose.

Conclusion: When Innovation Meets an Enduring Public Need, It Should Be Capable of Becoming Infrastructure

My understanding of mental-health accessibility has evolved through more than a decade of frontline experience, advocacy, service delivery, social entrepreneurship, experimentation, technology development, and sustained listening to people across a wide range of socioeconomic circumstances. I began with a relatively straightforward assumption: that one of the central access problems was an

insufficient supply of mental-health professionals and insufficient public coverage of professional care. That assumption led me to advocate for broader access to professional mental-health services, and the principle underlying that advocacy remains unchanged: no person who requires appropriate professional mental-health care should be prevented from obtaining it primarily because they cannot afford it. If a person requires psychotherapy, counselling, psychological care, social work, psychiatric care, medical assessment, trauma treatment, or another appropriate professional intervention, financial hardship should not be the principal reason that care remains beyond reach.

Frontline experience, however, gradually revealed that affordability is only one dimension of accessibility. People also encounter waitlists, provider mismatch, limited operating hours, language barriers, geographic barriers, disability-related barriers, cultural mismatch, fragmented navigation, finite program limits, repeated intake processes, and the emotional burden of beginning again after an unsuccessful attempt to find help. These gaps can affect people across socioeconomic circumstances. Most importantly, people encounter discontinuities before, between, and after formal episodes of care. A person can therefore live within a publicly funded healthcare system, possess theoretical access to services, and still have nowhere appropriate to turn at the moment support is needed. Formal coverage is essential, but formal coverage alone does not necessarily produce practical accessibility. A publicly funded service can exist while the person who needs it continues to wait; a qualified professional can exist somewhere within the system while remaining inaccessible because of geography, language, disability, specialization, scheduling, cost, or relational fit; and a program can conclude successfully from an administrative perspective while the circumstances affecting the person continue.

Over time, the question therefore became larger than how to expand one service or add another funding stream. I began asking whether the architecture of healthcare delivery itself could be redesigned. The central question was no longer simply, **How do we put more money into healthcare?** It became: **How do we redesign the way existing and future healthcare resources are allocated, connected, and delivered so that every public dollar, every professional hour, and every appropriate form of technological capacity can create greater meaningful health value?** This proposal emerged from that shift in perspective. It does not argue that additional funding is unnecessary. Genuine shortages of professionals, specialized services, infrastructure, or other forms of capacity may require additional public investment. Rather, it asks a complementary and equally important question: whether the resources already being committed to healthcare—and the new forms of capacity now becoming available—can be organized more intelligently, effectively, and equitably.

For A Safer Space (**FASS**) was my first practical response to one part of this problem. FASS sought to create a free human bridge between the moment a person realized **"I need help"** and the moment they reached appropriate professional care. It was designed around the principle that a person should not be left entirely unsupported simply because they could not afford another session, had reached the end of a short-term program, were waiting for care, or had not yet found the right helping professional. FASS demonstrated that such a bridge can matter, but it also revealed an unavoidable structural limitation of a human-only model: people can provide extraordinary compassion, judgment, care, wisdom, accountability, and presence, but they cannot be infinitely available. Counsellors have finite hours, therapists have finite caseloads, supervisors can responsibly oversee only so much work, additional language coverage requires additional human capacity, and every expansion of service generates further requirements for recruitment, training, supervision, administration, coordination, and funding. Removing the financial barrier for the person receiving support does not eliminate the underlying cost of human labour; it changes who carries that cost.

That realization contributed to the development of **FASSLING™**, which explored another dimension of the same accessibility problem. FASSLING™ emerged from the recognition that the need for appropriate support may be continuous even when human availability cannot be. Its purpose was never to replace psychotherapy, psychology, counselling, social work, medicine, or professional judgment. Instead, it asked whether appropriately bounded forms of emotional support, guided reflection, non-clinical coaching, psychoeducation, navigation, multilingual interaction, structured exercises, and continuity could become more continuously available without requiring another professional to become personally available for every additional appropriate interaction. FASS and FASSLING™ therefore became more than two complementary social innovations. Together, they revealed the contours of a broader systems problem: **the healthcare-access problem is not only about how much capacity exists; it is also about how different forms of capacity are financed, allocated, connected, and used**.

The proposed **Human-Supervised National Virtual Safe Space** is the public-infrastructure expression of those lessons. It is inspired by the experience of FASS and FASSLING™, but it should not be understood simply as the scaling of either organization or platform. It is a proposed national demonstration of whether Canada can create a free-at-the-point-of-appropriate-use, multilingual, accessible, trauma-informed, human-supervised foundational layer around professional care while simultaneously testing a broader healthcare budgeting hypothesis. The National Virtual Safe Space would therefore serve two connected purposes. First, it would test whether Canadians can have somewhere appropriate to turn before, between, and alongside professional care, particularly during periods when the right professional is not immediately available. Second, it would create a bounded and measurable environment in which government could examine whether different forms of public capacity can be financed differently while still functioning as one coherent continuum of care. In this sense, the National Virtual Safe Space is not simply another mental-health program and not simply an AI initiative. It is the practical demonstration case through which a possible budgeting reallocation can be evaluated.

Historically, many different forms of supportive capacity have been purchased through the same scarce input: human labour. Greater availability required more people, longer service hours required additional staffing, broader language coverage required more human capacity, and expanded reach often required parallel organizational and workforce growth. Professional expertise and basic availability were frequently financed through the same mechanism because there was no viable alternative. Emerging technological capacity changes that relationship for some carefully bounded functions. It creates the possibility that certain forms of availability, structured support, navigation, multilingual interaction, information, continuity, and repeatability may scale differently from complex professional care. This does not mean artificial intelligence makes support costless. Safe public digital infrastructure itself requires substantial and continuing investment in computation, model access, secure hosting, privacy architecture, cybersecurity, accessibility, multilingual quality, safety systems, qualified human supervision, professional involvement, governance, maintenance, and independent evaluation. The relevant distinction is therefore not between “expensive humans” and “cheap AI,” but between **different categories of public expenditure that purchase different forms of capacity**.

This is the rationale behind the proposed approximate **50/50 funding hypothesis**. One portion of a defined or comparable demonstration envelope could purchase scalable availability through continuous access, multilingual reach, accessibility, navigation, structured support, computation, infrastructure, and scale. Another portion could protect professional human capacity and governance for judgment, therapeutic depth, complex care, safety, intervention, supervision, accountability, and genuine human connection. The 50/50 ratio is not the policy objective and should never be interpreted as a permanent formula. It is a testable starting point. If evidence demonstrates that a particular population requires 60

percent, 70 percent, or more of the envelope to remain in human professional capacity, the allocation should change. If a carefully bounded, low-risk, highly structured function can safely rely more heavily on scalable infrastructure, that should also be evaluated. Different populations, languages, disability needs, cultures, service environments, clinical risks, and professional requirements may require substantially different balances. The larger hypothesis is not that 50/50 is correct; it is that **public funding may create greater value when different forms of capacity are financed according to the functions they are best suited to perform**.

The demonstration therefore asks whether Canada can, where appropriate, separate the cost of availability from the cost of expertise; use scalable infrastructure where appropriate abundance can be created responsibly; protect professional human capacity for functions in which judgment, relationship, accountability, and human presence are indispensable; connect those capacities into one continuum; and evaluate the complete architecture according to meaningful human outcomes. This is fundamentally different from simply continuing to put more money into an unchanged system. Additional funding may absolutely be necessary where genuine shortages exist, and this proposal does not argue otherwise. The point is that the existence of legitimate funding needs should not prevent government from asking whether resources already being spent are structured in the most effective, intelligent, and equitable way. Fiscal responsibility and human care should not automatically be treated as opposing principles. Better resource stewardship does not have to mean austerity, fewer services, or less professional care. It can mean examining whether each dollar is purchasing the form of capacity most appropriate to the function it is intended to support.

This distinction is especially important because public money is not the only scarce resource in the healthcare system. **Professional human attention is also scarce.** A therapist's hour can only be spent once. A counsellor, psychologist, social worker, physician, or other professional who spends that hour performing a highly standardized function cannot simultaneously use the same hour with someone whose circumstances require complex assessment, trauma-informed judgment, therapeutic depth, ethical reasoning, significant risk management, intervention, or sustained human presence. The cost of professional time therefore includes not only its financial price, but also the **opportunity cost of human attention**. For that reason, the objective of this proposal is not to eliminate professional labour; it is to protect and better deploy it. The relevant policy question is not **How many therapists can artificial intelligence replace?** That framing is too narrow and misunderstands the opportunity. The more consequential question is: **How can Canada redesign the allocation of human, technological, and financial capacity so that scalable forms of support become more continuously available, scarce professional human attention is protected for the work that genuinely requires human expertise, and fewer people are left without an appropriate and safe place to turn when professional care is unavailable, delayed, or temporarily out of reach?**

The underlying principle is straightforward: **use each form of capacity where it creates the greatest meaningful health value**. Within that architecture, the role of artificial intelligence becomes much clearer. AI should not replace human care; **it should protect human attention for the moments when human care matters most**. The National Virtual Safe Space could make that principle concrete by testing whether an appropriate foundational layer can remain available 24 hours a day and 365 days a year, without a per-interaction out-of-pocket charge, across multiple languages and accessible modes, while people are waiting for professional care, moving between appointments, searching for the right provider, or adjusting after a finite episode of service has ended. This would not mean unlimited psychotherapy, automated diagnosis, or the medicalization of ordinary human difficulty. It would mean something more precise: **continuous access to an appropriate foundational safety net within clear**

service, safety, privacy, accessibility, and professional boundaries. Professional care would remain essential whenever professional care is required; the National Virtual Safe Space would exist around that care, not instead of it.

A mature architecture should therefore be capable of answering two questions simultaneously: **Can I obtain the professional care I genuinely need?** and **Do I have somewhere appropriate and safe to turn right now?** Those questions are reflected in the proposal's two complementary national aspirations. The **Care Guarantee** addresses the affordability of professional expertise: no Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay. The **Continuity Guarantee** addresses immediate availability: no Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn. Neither guarantee replaces the other. Professional care provides depth, scalable infrastructure can provide appropriate continuity, qualified human supervision provides safety and accountability, and government provides standards, equitable access, financing architecture, governance, evaluation, and institutional durability. Together, these components create a more complete vision of accessibility.

That vision points toward a larger evolution in how universal healthcare itself is understood. Universal healthcare has traditionally focused, appropriately, on coverage and financial protection, but formal coverage does not guarantee practical access. A person can be covered and still wait. A person can be covered and still be unable to find the right provider. A person can be covered and still face linguistic, geographic, disability-related, cultural, scheduling, or navigational barriers. A person can be covered and still fall between programs. A publicly funded service can exist while the person who needs it remains unable to reach it when it matters. The next evolution of universal healthcare should therefore ask not only whether a person is covered, but whether the system enables that person to **actually reach appropriate care**. The aspiration of this proposal is to move, as far as responsibly and realistically possible, from **universal coverage toward universal practical accessibility**.

Achieving that goal requires more than money alone. It requires better connections between services, more deliberate use of professional expertise, appropriate technological infrastructure, continuity across transitions, accessible navigation, and clearer decisions about what different categories of public funding are intended to purchase. It means asking not only **How much are we spending?**, but also **What are we buying with that spending, and could the same or comparable resources be organized more effectively?** It means asking where professional expertise creates the greatest value, where technology can responsibly create appropriate abundance, where people are lost between services, where duplication occurs, where continuity breaks down, and how the system can be redesigned so that people experience something closer to a coherent continuum of care rather than a collection of disconnected programs.

This is why the larger ambition of the proposal is not simply the creation of a National Virtual Safe Space. The National Virtual Safe Space is the **demonstration case**. FASS and FASSLING™ provide the **frontline and social-innovation foundations** from which that case emerged. The approximate 50/50 allocation is the **initial testable budgeting-reallocation hypothesis**. Mental health is the **first demonstration environment**. But the larger subject is **healthcare-delivery redesign**. The proposal does not assume that the same technology, funding ratio, or architecture should automatically be applied to every other area of healthcare. Different domains involve different clinical risks, professional scopes, regulatory requirements, technologies, evidence standards, and populations. The potentially transferable contribution is therefore not a universal formula but a method of thinking: identify what form of capacity each healthcare function genuinely requires, determine where professional expertise is indispensable,

determine where availability or infrastructure can responsibly be made more scalable, examine whether different functions are being purchased through the same scarce form of capacity simply because systems were historically designed that way, and connect those capacities so that people experience one healthcare system rather than a series of institutional handoffs.

The deeper ambition is a healthcare-delivery system that is more holistic, more continuous, more person-centred, more fiscally deliberate, and more responsive to the practical realities of how people seek care. It is a system in which public resources are evaluated not solely by how much is spent, but by how effectively those resources are converted into **practical access, continuity, protected professional capacity, safety, equity, and meaningful health outcomes**. It is also a system in which the interests of patients, professionals, taxpayers, government, and social innovators need not automatically be treated as competing. Patients benefit when access and continuity improve. Professionals benefit when scarce expertise is protected for the functions in which it is genuinely needed. Government and taxpayers benefit when finite resources create greater public value rather than simply greater activity. Social entrepreneurs benefit when they can remain innovators rather than becoming permanent private substitutes for public infrastructure.

My own experience as a social entrepreneur has made that last distinction particularly important. Social entrepreneurs can identify needs that institutions have not yet recognized, remain close to frontline communities, experiment, challenge assumptions, build prototypes, and demonstrate what may be possible. But innovation and permanent infrastructure are not the same responsibility. A founder should not become the permanent private shock absorber for a systemic gap simply because they were willing to enter it first. **Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.** FASS and FASSLING™ can reveal possibilities, and a National Virtual Safe Space can transform those possibilities into a formally evaluated public-infrastructure demonstration. If rigorous evidence ultimately establishes that the underlying capability serves a persistent and legitimate public need, responsibility for maintaining that capability should not remain indefinitely dependent on one founder, nonprofit, vendor, technology company, funding cycle, or AI model. At some point, successful innovation can mature into infrastructure. **Public need should create public responsibility.**

That does not necessarily mean direct government operation. Government may regulate, fund, procure, contract, certify, coordinate, establish standards, evaluate, or use a combination of institutional arrangements. What matters is that the capability itself, if demonstrated to have enduring public value, can survive changes in technologies, vendors, organizations, founders, funding arrangements, and governments. **Government should protect the promise, not the platform.** The attached proposal already establishes this transition from prototype to durable public capability: if the underlying need proves persistent, infrastructure should not remain indefinitely dependent on one founder or organization.

The long-term significance of such a capability could also extend across the life course. A student might first use an appropriate foundational layer during a difficult academic period and return years later during a career transition or relationship difficulty. The same person might use professional therapy during bereavement and rely on the foundational safety net between appointments, then return much later when navigating caregiving responsibilities or aging. A newcomer could use multilingual navigation while learning how the Canadian healthcare system operates. A person in a rural or remote community might have somewhere appropriate to begin while waiting for specialized care. Someone seeking a professional with the right cultural or linguistic fit could remain connected to an appropriate foundational layer while

continuing that search. This is what distinguishes infrastructure from a temporary program: programs have defined endpoints, while **public capability can remain relevant as needs change**.

The broader international significance of the proposal would likewise extend beyond artificial intelligence. Many countries face similar constraints: limited professional capacity, rising demand, geographic inequalities, language diversity, workforce strain, fragmented pathways, fiscal pressures, and uncertainty about how emerging technologies should be integrated into human services. The international lesson should not be that Canada found a way to replace therapists with artificial intelligence. It should be that Canada tested whether **responsible technological capacity, protected professional expertise, deliberate public financing, and accountable institutional design could be combined to produce greater practical access without weakening human care**. Other countries would not need to adopt the same technology, the same 50/50 starting point, or the same institutional structure. The transferable contribution would be the underlying resource-allocation method: determine what form of capacity each function genuinely requires, finance that capacity deliberately, protect scarce human expertise, use scalable infrastructure where appropriate, connect the different layers, and evaluate the complete system according to meaningful human outcomes.

The ultimate ambition of this proposal is therefore not technological. **It is human.** No healthcare system can eliminate every waitlist, provider mismatch, geographic barrier, or form of scarcity. Governments do not have unlimited resources. Professional human attention will remain finite. Technology will remain imperfect. But **scarcity does not always have to become abandonment**. A person may have to wait for a particular professional, but they should not necessarily have to wait with absolutely nowhere appropriate to turn. A program may end, but **the person's life does not**. The goal is not to promise an impossible healthcare system in which every service is instantly available to everyone. The goal is to design healthcare delivery, as far as responsibly and realistically possible, so that fewer people are left behind simply because the architecture of the system failed to connect them to the right form of care.

That is the deeper purpose of this proposal. **Universal healthcare should not merely mean that everyone is covered. It should mean that the system is designed so that people can actually reach the care they need.** The National Virtual Safe Space is one concrete way to test that principle. The experience of FASS and FASSLING™ helped reveal the problem and demonstrate possible responses. The proposed budgeting reallocation provides a mechanism through which government can test whether different forms of capacity can be used more intelligently. The larger healthcare-reform question is whether Canada can organize the resources it already has—and the new forms of capacity now becoming available—in a way that is smarter, more effective, more equitable, and more human-centred.

The ultimate measure of success should therefore not be the number of AI interactions generated, the amount of computational capacity purchased, whether FASSLING™ remains the platform used, or whether the initial 50/50 ratio proves correct. It should be whether Canada can create a more dependable relationship between people and the healthcare system: one in which professional care is available when professional care is genuinely required, appropriate continuity exists when that care is temporarily unavailable, scarce human expertise is protected for the moments when human beings matter most, and public resources are organized according to the meaningful human value they create.

Technology will change. Providers will change. Funding models will change. Governments will change. Professional standards will evolve. But the public aspiration can remain: **No one should be left behind simply because the architecture of the healthcare system failed to connect them to the right care.**

If Canada can make that aspiration real—not simply as a policy statement, but as reliable, accountable, human-centred public infrastructure—then the enduring significance of the initiative will be much larger than a mental-health technology project. It will demonstrate how social innovation can reveal a systemic problem, how evidence can test a different architecture, how professional expertise can safeguard what must remain human, and how government can use public resources more intelligently to turn formal coverage into practical accessibility.

And perhaps, from the perspective of the person seeking help, the meaning of that entire architecture will ultimately be much simpler:

My country has my back.

Appendix A: Key Definitions

The following definitions establish the terminology used throughout this proposal. They are intended to clarify scope, professional boundaries, governance responsibilities, and the distinction between foundational support and regulated or specialized professional care.

Virtual Safe Space

A **Virtual Safe Space** is a free-at-the-point-of-appropriate-use, accessible, trauma-informed, privacy-preserving digital environment designed to provide an appropriate foundational layer of emotional support, guided reflection, non-clinical coaching, life-skills support, routine psychoeducational information, and service navigation.

Its purpose is not to replace psychotherapy, counselling, regulated mental-health care, medical treatment, crisis intervention, emergency services, or professional judgment. Rather, it is intended to ensure that a person has **somewhere appropriate and safe to turn before, between, and alongside formal professional services**, including while waiting for care, searching for the right provider, navigating an unsuccessful provider match, moving between services, or experiencing a gap after a time-limited program ends.

A Virtual Safe Space should operate within clearly defined service boundaries and should incorporate transparent limitations, trauma-informed design, multilingual capability, text and audio-enabled interaction, accessibility, privacy by design, qualified human oversight, graduated escalation pathways, and reliable connections to professional, community, crisis, or emergency services where required.

Its foundational public promise is: **There is somewhere appropriate I can turn right now.**

Foundational Support

Foundational support refers to appropriate, non-emergency and non-diagnostic assistance available at the beginning of a help-seeking journey or during gaps surrounding formal professional care.

Depending on service scope and the person's circumstances, foundational support may include basic emotional support, guided reflection, non-clinical coaching, goal clarification, routine psychoeducation, structured exercises, life-skills support, basic navigation, and assistance understanding available service options.

Foundational support is not a substitute for professional assessment, diagnosis, psychotherapy, medical treatment, crisis intervention, emergency services, or other regulated and specialized care where such services are required.

Its purpose is narrower: **To reduce the likelihood that a person is left with no appropriate support at all simply because professional care is not immediately available.**

Human-Supervised

A **human-supervised** model is one in which qualified people remain responsible for defining service scope, governing system behaviour, maintaining professional boundaries, overseeing safety, reviewing quality, managing escalation, and ensuring institutional accountability.

Human supervision may include professional governance, AI-output quality review, review of predefined safety signals under strict privacy controls, direct intervention, escalation, ethical oversight, professional-boundary review, culturally responsive review, referral, service navigation, and connection to professional, crisis, or emergency services where necessary.

Human supervision does not mean that a professional must manually participate in every AI-enabled interaction. Rather, it means that **scaled technological availability remains embedded within accountable human governance**.

Professional Care

Professional care refers to services for which regulated, licensed, specialized, or otherwise professionally accountable expertise is required.

Depending on jurisdiction, professional scope, and individual need, this may include services provided by psychologists, psychotherapists, physicians, psychiatrists, social workers, counsellors, trauma specialists, and other appropriately credentialed professionals.

Professional care may include assessment; diagnosis where legally authorized; psychotherapy; treatment planning; complex risk evaluation; trauma treatment; therapeutic relationship-building; professional interpretation; ethical decision-making; specialized intervention; and other functions in which professional expertise, accountable judgment, and genuine human presence are central to the service itself.

Under the proposed architecture, professional care remains fundamentally human-led.

AI does not remove the need for professional care. It is intended to help protect scarce professional capacity for it.

Continuity

Continuity refers to the sustained availability of appropriate support before, between, and after formal professional encounters so that finite professional capacity does not automatically result in complete absence of support.

Continuity may include appropriate support before a first appointment; while waiting for professional care; between appointments; while searching for a better provider fit; after an unsuccessful professional match; after the end of a short-term counselling program; following completion of an Employee Assistance Program or structured intervention; outside conventional service hours; or during a later recurrence of emotional or practical need.

Continuity does not mean that professional care becomes unlimited. It means that **the surrounding safety net does not necessarily disappear when professional services are temporarily unavailable or one formal episode of care ends**.

Care Guarantee

The **Care Guarantee** is the long-term public aspiration that:

> **No Canadian who requires appropriate professional mental-health care should be prevented from obtaining it primarily because of inability to pay.**

It addresses the affordability of professional expertise.

Continuity Guarantee

The **Continuity Guarantee** is the long-term public aspiration that:

> **No Canadian who is searching for, waiting for, moving between, or temporarily unable to access appropriate professional care should be left with absolutely nowhere appropriate and safe to turn.**

It addresses gaps created by finite professional availability.

The Care Guarantee and Continuity Guarantee solve different problems and are intended to reinforce one another.

Computational Abundance

Computational abundance refers to the ability of appropriately governed digital infrastructure to provide large-scale availability, structured interaction, multilingual access, repeatable supportive functions, and continuity without requiring professional labour to increase in direct proportion to every additional interaction.

The term does not imply that digital infrastructure is free or infinitely scalable.

Computational abundance requires continuing investment in model access, token or inference usage, computing capacity, secure hosting, privacy, cybersecurity, accessibility, multilingual functionality, safety systems, testing, maintenance, monitoring, human supervision, governance, and independent evaluation.

Its economic significance lies in the fact that these costs may scale differently from one-to-one professional human labour.

Scarce Professional Human Attention

Scarce professional human attention refers to the finite time, judgment, expertise, relational capacity, ethical responsibility, and professional accountability of therapists, counsellors, psychologists, social workers, physicians, supervisors, and other qualified helping professionals.

Professional attention is treated throughout this proposal as a valuable public resource because: **a professional hour can only be spent once.**

The proposal therefore asks whether appropriately bounded structured, repetitive, informational, navigational, and text-based functions can increasingly be supported through technology so that professional attention can be redirected toward complex needs, therapeutic relationships, risk assessment, trauma-informed care, ethical ambiguity, professional judgment, intervention, escalation, supervision, and other circumstances in which human expertise is indispensable.

Opportunity Cost of Professional Attention

The **opportunity cost of professional attention** is the public value forgone when a qualified professional uses scarce time on a function that could be safely and appropriately supported through another form of capacity. The relevant cost is therefore not limited to wages, fees, or reimbursement.

If a professional spends time on a highly standardized function that could be responsibly AI-assisted, that same time is unavailable for a person whose needs require complex judgment, therapeutic depth, risk assessment, intervention, or human relationship.

This concept is central to the proposal's workforce and public-finance logic.

Resource Reallocation

Resource reallocation refers to the deliberate redistribution of a defined or comparable public funding envelope between **scalable AI-enabled infrastructure** and **human professional capacity and governance**.

The purpose is not simply to reduce expenditure. It is to test whether public resources can generate greater access, continuity, safety, equity, professional capacity, and public value when different forms of capacity are allocated according to their comparative advantage.

The proposed approximate 50/50 division is a demonstration hypothesis rather than a permanent funding formula.

Cost of Availability

The **cost of availability** refers to the expenditure required to ensure that an appropriate foundational support layer can be reliably accessed across time, geography, languages, communication modes, and differing levels of demand.

Under the proposed model, this may include expenditure on computing, model access, inference, secure infrastructure, multilingual capability, accessibility, cybersecurity, privacy protections, maintenance, safety systems, monitoring, testing, and related digital infrastructure.

Cost of Expertise

The **cost of expertise** refers to the expenditure required to provide qualified professional judgment, therapeutic depth, specialized treatment, supervision, escalation, safety review, complex-care intervention, ethical oversight, and accountable human presence.

The proposal's economic thesis rests in part on the possibility that **availability and expertise no longer need to be purchased exclusively through the same form of capacity**.

Technology-Neutral Public Capability

A **technology-neutral public capability** is a public commitment to maintain an underlying service function regardless of which platform, AI model, vendor, nonprofit organization, contractor, or technical architecture provides it at a particular time.

Under this proposal, government would not need to guarantee the permanent operation of FASSLING™, a particular AI model, one nonprofit organization, or one technology company. The public commitment would instead preserve the underlying capability: **Free-at-the-point-of-appropriate-use, continuous, multilingual, trauma-informed, accessible, privacy-preserving, human-supervised emotional-support infrastructure connected to appropriate professional care.**

Technology may change. Providers may change. The public promise can endure. This is the practical meaning of the principle: **Government should protect the promise, not the platform.**

Trauma-Informed AI

Trauma-informed AI refers to AI-enabled support designed not only around technical accuracy, but also around psychological safety, dignity, autonomy, transparency, choice, non-exploitation, privacy, avoidance of unnecessary re-traumatization, cultural context, appropriate boundaries, and accountable escalation.

A trauma-informed system should ask not only: **Is this response technically plausible?** It should also ask: **How might this interaction affect a person who has experienced trauma, coercion, rejection, discrimination, marginalization, abuse, fear, or institutional betrayal?**

Trauma-informed AI should not imitate professional authority it does not possess, promote dependency, obscure its identity, or discourage appropriate human care.

Frontline Learning

Frontline learning refers to the responsible use of privacy-preserving, de-identified, and appropriately aggregated patterns to identify recurring systemic barriers that may not yet be visible through conventional reporting systems.

Potential patterns could include persistent waitlists, underserved languages, navigation failures, rural or regional disparities, recurring affordability barriers, accessibility gaps, repeated provider-matching difficulties, and forms of unmet support.

Frontline learning must never become individualized surveillance.

Its governing principle is: **understanding systemic barriers at scale, not watching people at scale.**

Human-AI Symbiosis

Human-AI symbiosis refers to a public-service architecture in which technological and human capabilities are deliberately combined rather than treated as substitutes.

Within this proposal:

- **AI contributes availability, repeatability, multilingual reach, structure, and scale.**
- **Human supervisors contribute accountability, safety, governance, and escalation.**
- **Qualified professionals contribute judgment, therapeutic depth, complex care, intervention, and genuine human presence.**
- **Government contributes standards, funding, equitable access, evaluation, and institutional durability.**

The objective is not maximum automation. It is the most responsible allocation of each form of capacity.

Appendix B: Summary Logic Model

The following logic model summarizes the theory of change underlying the proposed **Human-Supervised Virtual Safe Space Demonstration Initiative**. It shows how public resources would be converted into activities, outputs, outcomes, and—only if supported by evidence—a durable public capability.

The model should be read as a set of propositions to be tested, not as a prediction that all intended outcomes will necessarily occur.

Summary Logic-Model Table

Logic-Model Element	Illustrative Content
Public Need / Problem	Affordability barriers; waitlists; provider mismatch; operating-hour constraints; language barriers; rural and remote access; disability-related barriers; fragmented navigation; finite program durations; and finite professional human attention.
Inputs	Defined public funding envelope; AI model access; token and inference capacity; computing and secure hosting; professional workforce; supervisors; governance structures; privacy and cybersecurity expertise; accessibility specialists; cultural and linguistic expertise; people with lived experience; legal and regulatory expertise; Indigenous participation where appropriate; social innovators; researchers; and independent evaluators.
Resource-Allocation Hypothesis	Approximate 50/50 demonstration allocation between AI-enabled digital infrastructure and human professional capacity/governance, with the ratio treated as an empirical starting point rather than a permanent formula.
Core Activities	Deliver bounded 24/7 AI-assisted foundational support; provide multilingual text and audio-enabled interaction; automate or augment appropriate structured functions; supervise system performance; review safety issues under strict privacy rules; escalate and intervene when required; connect users to professional and community care; train professionals for emerging roles; test accessibility and multilingual performance; monitor privacy and cybersecurity; evaluate cultural responsiveness; test different human/computational allocations; and identify de-identified aggregate system patterns.
Immediate Outputs	Faster access to foundational support; multilingual interactions; text and audio-enabled availability; referrals and navigation outcomes; human interventions; professional supervision; accessibility improvements; safety and privacy findings; workforce-training outputs; governance recommendations; and data on infrastructure, supervision, escalation, and professional-capacity costs.
Short-Term Outcomes	Greater continuity; fewer unsupported gaps; improved navigation; clearer AI-human boundaries; stronger user trust; better accessibility and language reach; more appropriate use of professional time; improved connection to professional care; clearer escalation pathways; and earlier identification of recurring system barriers.
Medium-Term Outcomes	Greater professional capacity for complex care; sustainable emerging professional roles; improved allocation of public funding; stronger fiscal productivity; reduced dependence on repetitive human delivery; more equitable geographic and linguistic access; stronger privacy and safety standards; clearer legal and governance frameworks; improved public-sector learning; and reduced founder/vendor dependency.

Logic-Model Element	Illustrative Content
Long-Term Outcome, If Evidence Supports It	A durable Canadian public capability combining continuously available digital support, qualified professional expertise, human accountability, privacy, accessibility, public funding, technology neutrality, and institutional responsibility.
Long-Term Public Impact	Progress toward the Care Guarantee and Continuity Guarantee; reduced periods of complete support absence; better protection of professional attention; stronger public trust; more sustainable social innovation; durable national support infrastructure; and transferable Canadian expertise in human-centred AI-enabled public services.

Comprehensive Logic Model

Public Need

Canada faces multiple overlapping access barriers:

cost • waitlists • provider mismatch • limited operating hours • language barriers • rural and remote access • disability-related barriers • fragmented navigation • short-term program limits • finite professional human attention

↓

Inputs

Public funding

- **AI model, inference, computing, and secure infrastructure**
- **Therapists, counsellors, psychologists, social workers, supervisors, navigators, and other qualified humans**
- **Privacy, cybersecurity, accessibility, cultural, linguistic, legal, regulatory, and AI-safety expertise**
- **People with lived experience, community participants, Indigenous partners where appropriate, and social innovators**
- **Independent research and evaluation**

↓

Resource-Allocation Hypothesis

Approximate 50/50 Demonstration

≈ 50% AI-Enabled Infrastructure
purchasing availability, continuity, multilingual reach, accessibility, secure computation, structured support, and technical safeguards

≈ 50% Human Capacity and Governance
purchasing judgment, professional depth, complex care, relationships, escalation, supervision, safety, and accountability

↓

Three-Layer Service Architecture

Layer 1 — Always-Available Virtual Safe Space
24/7 foundational support • multilingual text/audio • reflection • non-clinical coaching • psychoeducation • life skills • navigation

↓

Layer 2 — Human-Supervised Support
quality assurance • safety review • escalation • intervention • professional boundaries • culturally responsive review • navigation • governance

↓

Layer 3 — Professional and Specialized Care
psychotherapy • professional assessment • diagnosis where authorized • complex risk evaluation • trauma treatment • therapeutic relationships • specialized intervention

↓

Core Activities

Provide immediate foundational support; automate or augment suitable structured functions; supervise AI; connect users to professional care; escalate where required; train the workforce; evaluate safety, accessibility, privacy, and cultural responsiveness; test resource allocations; and conduct appropriately governed aggregate system learning.

↓

Immediate Outputs

Faster initial access; multilingual use; text/audio availability; referrals; human interventions; supervision; accessibility improvements; privacy and safety findings; professional training; governance recommendations; and evidence on operating and supervision costs.

↓

Short-Term Outcomes

greater continuity • fewer unsupported gaps • better navigation • clearer boundaries • stronger trust • more appropriate use of professional attention • improved human-care connections • earlier identification of systemic barriers

↓

Medium-Term Outcomes

greater complex-care capacity • sustainable professional roles • stronger fiscal productivity • more equitable access • stronger safety and privacy governance • clearer legal frameworks • better policy-learning loops • reduced founder dependency

↓

Long-Term Outcome

A DURABLE CANADIAN HUMAN-SUPERVISED PUBLIC SUPPORT CAPABILITY

Free at the point of appropriate use • continuous • multilingual • accessible • trauma-informed • privacy-preserving • text and audio-enabled • professionally connected • technology-neutral

↓

Two National Aspirations

CARE GUARANTEE

Appropriate professional care should not be denied primarily because of inability to pay.

CONTINUITY GUARANTEE

Canadians waiting for, searching for, or moving between appropriate care should not be left with absolutely nowhere appropriate and safe to turn.

↓

Potential Long-Term Public Impact

Earlier access • fewer complete support gaps • better use of professional expertise • stronger institutional trust • sustainable social innovation • durable national capability • transferable international lessons in responsible AI-enabled social infrastructure

The Logic in One Sentence

Public resources fund scalable availability and protected human expertise; those capacities create a three-layer continuum of support; the continuum is tested for improvements in access, continuity, professional capacity, safety, governance, and fiscal productivity; and, only if the evidence supports the model, those improvements may mature into durable infrastructure capable of advancing the Care Guarantee and Continuity Guarantee.

Logic Model Through the Lens of Resource Reallocation

The economic theory of change can be expressed more narrowly:

Public funding is finite

↓

Professional human attention is finite

↓

Some structured, repetitive, informational, navigational, and text-based functions may be suitable for responsible AI assistance

↓

AI augments or automates appropriate functions within clear safeguards

↓

Professional human attention is potentially released

↓

Released attention is deliberately redirected toward complexity, relationships, professional judgment, safety, intervention, supervision, and therapeutic depth

↓

AI-enabled infrastructure increases appropriate foundational availability

↓

More people can access a first layer of support without paying for every interaction

↓

Fewer people experience complete gaps before, between, and after professional services

↓

A comparable funding envelope may generate greater overall public value

↓

Evidence determines whether the Continuity Guarantee can become a credible and durable public capability

The word **may** is important throughout this chain. Each transition is a hypothesis requiring empirical validation.

That is the core theory of change.

Appendix C: Relevant Policy, Practice, and Innovation Background

This proposal emerges from a longer trajectory of advocacy, frontline service, social innovation, and experimentation. It did not begin with artificial intelligence. It developed through successive attempts to understand different dimensions of mental-health accessibility: affordability, continuity, provider fit, human-capacity limitations, digital availability, resource allocation, governance, and the institutional responsibilities that arise when a social innovation begins addressing a persistent public need.

Public-Coverage Advocacy: Identifying the Affordability Problem

In 2021, I initiated a public petition calling for mental-health services provided by mental-health professionals to be included within Ontario's public health-insurance framework.

The principle underlying that advocacy was straightforward: **access to appropriate professional mental-health care should not depend primarily on a person's ability to pay.**

That work highlighted the affordability dimension of mental-health access and later developed into the principle expressed in this proposal as the **Care Guarantee**.

The proposal does not treat AI-enabled support as an alternative to that objective. It treats expanded professional access and foundational continuity as complementary goals.

For A Safer Space (FASS): Testing Free Human Continuity

For A Safer Space (FASS) represented my first practical effort to address gaps that affordability reform alone could not resolve.

FASS was developed as what I believe was **the world's first organization designed to provide unlimited, free mental-health counselling, psychotherapy, emotional support, and coaching sessions globally**. Because this is a potentially verifiable superlative, it should remain qualified unless independent evidence is available to substantiate a stronger formulation.

Through a model involving clinical interns and trainees from universities in different jurisdictions, appropriate supervision, and connections to licensed or specialized professional services where necessary, FASS sought to create an accessible human bridge between: **"I need help"** and **"I have found the right professional."** Its intended role was complementary.

Help-seekers were encouraged to use appropriate licensed or specialized professional care alongside FASS whenever such care was available and required. The experience reinforced the value of having somewhere to turn while people waited, searched for an appropriate provider, encountered poor therapeutic fit, moved between services, or required support between professional appointments.

It also revealed a structural constraint that shaped the next stage of innovation: **Human attention is finite.** Free access can remove the charge imposed on the user. It does not eliminate the underlying scarcity of counsellors, supervisors, languages, operating hours, and professional capacity.

FASSLING.AI: Exploring Continuous Availability and the Limits of Founder-Led Infrastructure

FASSLING™ emerged directly from the capacity limitations identified through FASS. The underlying accessibility mission remained substantially the same: people should not be left entirely without an appropriate place to turn because professional care is temporarily unavailable, unaffordable, difficult to navigate, or not yet the right fit. What changed was the delivery architecture.

Where FASS explored whether free human support could provide a bridge before, between, and alongside professional care, FASSLING™ explored whether artificial intelligence could make part of that bridge **continuously available**. It was developed as a free, multilingual, 24/7 environment for non-medical emotional support, guided reflection, non-clinical life coaching, service navigation, and holistic life-skills assistance. Its purpose has consistently been complementary rather than substitutive: FASSLING™ was not designed to replace counselling, psychotherapy, psychology, psychiatry, social work, crisis intervention, emergency services, or other forms of appropriate professional care.

The underlying problem FASSLING™ sought to address was practical. People who recognize that they need support may spend considerable time searching for an appropriate therapist or other helping professional. They may encounter cost barriers, waitlists, limited operating hours, language barriers, cultural mismatch, or an unsuccessful provider fit. Some may already be receiving professional care while still needing an appropriate place to reflect between appointments. Others may not yet know whether psychotherapy, counselling, coaching, social services, peer support, medical care, or another form of assistance is most appropriate.

FASSLING™ was developed as a potential **buffer during those gaps**. Its premise is not that AI is equivalent to professional care. It is that the temporary absence of the right professional should not necessarily require the complete absence of support. A person may still benefit from an opportunity to organize their thoughts, reflect on what they are experiencing, receive appropriate non-clinical emotional support, practise life skills, explore possible next steps, or obtain basic navigation while continuing to seek suitable human care.

Artificial intelligence introduced a fundamentally different form of availability. Unlike one-to-one human service delivery, an AI-enabled system can operate across time zones, remain available outside conventional office hours, support text and audio-enabled interaction, and communicate across many languages without requiring an additional human provider to become personally available for every additional interaction. FASSLING™ has been developed with multilingual capability across more than 95 languages, although the precise number of supported languages and the quality of performance across those languages should be independently documented before formal submission.

The significance of this capability is not that human service has become unnecessary. Human care and AI-enabled support possess different strengths. Professionals provide judgment, therapeutic depth, relational accountability, complex assessment, intervention, and genuine human presence. AI can potentially provide immediacy, repeatability, multilingual reach, structured support, and continuity. FASSLING™ was therefore an early practical exploration of the same division of capacity proposed in this submission: **technology for appropriate availability and scale; human beings for judgment, complexity, safety, accountability, and connection**.

From Emotional Support to a Holistic Life-Skills Ecosystem

The development of FASSLING™ also revealed that emotional support alone is insufficient to create a psychologically safer life. People frequently need practical capacities that help them navigate everyday challenges before, during, and after periods of emotional difficulty. These may include communication, boundary-setting, self-awareness, emotional regulation, stress management, resilience, decision-making, relationship skills, conflict navigation, goal clarification, and other foundational life competencies.

For that reason, I developed a broader family of FASSLING™ AI-supported programs focused on holistic and essential life skills. The design logic is complementary: a person may use one component to practise a particular skill and then use the emotional-support and coaching environment to reflect on the feelings, uncertainty, resistance, or interpersonal challenges that emerge during that process.

The resulting concept is therefore broader than a single conversational tool. It is an exploratory **ecosystem of bounded support functions** intended to address different dimensions of everyday well-being while maintaining the distinction between non-clinical support and professional mental-health care.

This distinction is important from a policy perspective. Not every difficult human experience is a clinical disorder. Not every supportive conversation requires psychotherapy. Not every life challenge should be medicalized. At the same time, some circumstances clearly do require professional assessment, treatment, crisis intervention, or other specialized human care. A mature support architecture should therefore be capable of distinguishing among levels of need and connecting people to the appropriate form of assistance.

A Prototype of Continuous Foundational Support

FASSLING™ should be understood within this proposal principally as a **working social-innovation prototype**, not as a predetermined national platform.

Its relevance lies in what its development has made possible to explore: whether free, continuously available, multilingual, AI-enabled foundational support can reduce some of the gaps created by the unavoidable scarcity of human availability.

In the policy journey underlying this submission:

- **FASS explored the need for an accessible human bridge.**
- **FASS revealed the capacity limits of a human-only bridge.**
- **FASSLING™ explored whether an appropriate part of that bridge could become continuously available through technology.**

The next question is no longer primarily technological or entrepreneurial. It is institutional: Can the underlying capability be made safe enough, professionally governed enough, private enough, accessible enough, culturally responsive enough, and economically sustainable enough to justify evaluation as part of public infrastructure?

That question should not be answered by FASSLING™ alone, nor should government simply assume that one founder’s prototype should become a national service. The appropriate next step is a controlled

demonstration in which the **underlying capability is independently tested under public-interest standards**.

The Limits of Maintaining a Public-Interest AI Safety Net Through a Nonprofit Alone

The development of FASSLING™ also exposed a second problem that is directly relevant to this proposal: the increasing mismatch between the scale of the public responsibility involved and the institutional capacity of a founder-led nonprofit to carry that responsibility indefinitely.

An emotionally sensitive AI-enabled service requires far more than the operation of a conversational model. At meaningful scale, responsible delivery may require continuous investment in model and inference capacity, secure technical infrastructure, privacy and cybersecurity, accessibility, multilingual quality assurance, professional-boundary governance, trauma-informed design, safety systems, escalation protocols, legal review, insurance, quality assurance, human supervision, regulatory compliance, independent testing, and ongoing evaluation.

As regulation and public expectations surrounding artificial intelligence appropriately become more rigorous, these responsibilities also become more complex. In my experience developing FASSLING™ without a conventional external funding base, this creates a structural problem. A nonprofit can demonstrate a public need and build an early prototype, but there are limits to how much continuing financial, technical, professional, ethical, and legal responsibility one organization—and particularly one or two founders—can reasonably absorb while attempting to provide a service without charging users.

The more widely a free public-interest service is used, the greater those responsibilities may become. This creates the same paradox identified elsewhere in this proposal: **The more public value a social innovation creates, the greater the private burden its founder may be required to carry.** That is not a sustainable foundation for essential social infrastructure.

This is not an argument that social entrepreneurs should receive immunity from accountability. FASSLING™, like any emotionally sensitive service, should remain subject to appropriate privacy, safety, professional, ethical, and legal expectations.

The issue is instead one of **institutional allocation of responsibility**.

A single nonprofit should not be expected to independently define national standards for AI-supported emotional services, resolve complex questions of liability and professional scope, finance indefinite computational availability, oversee increasingly sophisticated safety requirements, and carry the long-term responsibility for ensuring that an essential support layer remains available to a growing population.

If the underlying capability serves a genuine and enduring public need, responsibility must eventually become broader than the founder.

From Prototype to Public-Policy Question

FASSLING™ has therefore brought this policy journey to an important transition point. A prototype exists. The accessibility problem it seeks to address has been identified through frontline experience.

The technological possibility of continuous, multilingual, AI-enabled foundational support can now be explored in practice. However, a founder-led nonprofit cannot by itself determine whether that capability is sufficiently safe, effective, equitable, trusted, professionally responsible, and economically sustainable to become part of Canada's long-term support architecture. That determination requires government, professionals, researchers, people with lived experience, privacy and accessibility experts, cultural and linguistic expertise, responsible-AI specialists, and independent evaluators.

My request to the Government of Canada is therefore **not simply to fund or nationally adopt FASSLING™ as it presently exists**. It is more institutionally significant: **Use the prototype and the frontline learning behind it as a starting point for testing whether Canada should establish a human-supervised, technology-neutral public capability that ensures people have an appropriate first place to turn when professional care is not immediately available.**

FASSLING™ can demonstrate what may be technologically possible. It should not be required to carry the entire public responsibility alone.

If evidence ultimately shows that continuous foundational emotional-support infrastructure produces meaningful public value, then Canada should determine how to make that capability accessible under stronger governance, professional oversight, privacy protection, sustainable financing, and institutional accountability than any individual nonprofit could provide independently.

The long-term objective is therefore larger than FASSLING™ itself. It is not to guarantee that one bot, one nonprofit, one founder, or one AI platform operates forever. It is to determine whether Canada can make the underlying promise durable: **Free-at-the-point-of-appropriate-use, continuously available, multilingual, accessible, trauma-informed, human-supervised foundational support connected to appropriate professional care.**

In that sense, FASSLING™ represents a stage in the innovation journey rather than its endpoint.

- **The social entrepreneur can build the prototype.**
- **The demonstration can determine whether the capability works.**
- **Professional and public governance can establish the safeguards.**
- **And, if the evidence warrants it, government can help ensure that the capability becomes available beyond the limits of one founder or organization.**

That is the transition this proposal asks Canada to evaluate: **from a founder-built proof of possibility to a publicly accountable capability capable of serving Canadians at scale.**

Frontline Experience in Government-Funded Service Delivery

My perspective is also informed by previous frontline experience as a **Government Services Counsellor and Integrated Health Solution Specialist in a government-funded service environment operated by Morneau Shepell/LifeWorks, now TELUS Health**. The precise job title, funding relationship, organizational description, and successor-company wording should be independently verified before formal submission.

That experience provided exposure to structured counselling, short-term interventions, service boundaries, standardized processes, referral pathways, and the operational realities of finite professional

labour. It also contributed to a frontline observation central to this proposal: many structured mental-health and supportive-service environments contain significant amounts of work that may be structured, repetitive, protocol-guided, informational, or increasingly text-based.

Historically, trained professionals delivered nearly all of these functions because there was no technically viable alternative. Contemporary AI creates the possibility that some bounded functions may now be augmented or automated responsibly. That observation produced the resource-allocation question at the centre of the proposal:

> **If professional human attention is scarce, which functions truly require that scarce expertise, and which can increasingly be supported by another form of capacity?**

This remains a hypothesis to be tested rather than a settled conclusion.

Trauma-Informed Virtual Safe Space Advocacy

My Forbes Nonprofit Council article, *Why Every Nonprofit Needs a Trauma-Informed Virtual Safe Space*, developed the principle that a virtual safe space should be understood as more than a digital communications tool. Its public value depends on psychological safety, dignity, accessibility, privacy, inclusion, non-judgment, and connection to appropriate human support. That work informs the recommendation in this proposal that any public virtual safe-space capability should be evaluated not merely for technical functionality, but for its potential role as **social and digital infrastructure governed around human dignity**.

The specific publication title, date, URL, and relevant claims should be cited in the final submission.

Legal Reform and Culturally Responsive Social Innovation

My Forbes Nonprofit Council article, *Why Culturally Sensitive Social Entrepreneurship Demands a Legal Rethink*, examined the governance challenges that arise when social innovation evolves more quickly than the legal and professional categories through which traditional services are regulated.

Social entrepreneurs may develop models spanning education, coaching, peer support, community care, mental-health support, technology, culture, navigation, and professional services. These models can create legitimate questions concerning scope of practice, supervision, privacy, duty of care, consent, data governance, AI accountability, cross-jurisdictional responsibility, and liability.

The work also highlighted the importance of cultural context, relational ethics, trust, family structures, communication norms, spirituality, and different understandings of support. It informs this proposal's recommendation that government develop clearer and more coherent public frameworks for emerging emotionally sensitive AI-enabled services.

The objective is not deregulation or immunity from responsibility. It is **proportionate accountability within clearer public standards**.

Social Entrepreneurship, Service, and Founder Burden

My Forbes Nonprofit Council article, *The Nonprofit as Dojo: How Unconditional Love Fuels Social Entrepreneurship*, examined service, resilience, responsibility, and personal sacrifice within nonprofit leadership.

Those experiences also raise a wider public-policy question. Social entrepreneurs often enter neglected areas before institutional systems are prepared to act. They contribute time, expertise, personal resources, and reputational capital while accepting significant operational uncertainty. That willingness to experiment has public value. However, compassion and founder resilience should not become permanent substitutes for institutional responsibility.

Where a social innovation begins serving an enduring public need, policymakers should consider when responsibility must move beyond the founder.

The relevant principle is: **Social entrepreneurs can be catalysts. They should not have to become permanent shock absorbers for systemic gaps.**

Global Advocacy for Legal Reform

I have also initiated a global petition calling attention to legal and governance reforms intended to protect responsible social innovators while preserving accountability and public safety.

The underlying principle is directly relevant here: **society should not ask individuals to address systemic public problems while leaving them alone to absorb all of the legal and regulatory uncertainty created by responsible innovation.**

Clear public standards can simultaneously protect users and provide responsible innovators with more predictable obligations regarding professional boundaries, privacy, supervision, liability, AI governance, and public-interest experimentation.

The existence, title, date, petition text, signature count if referenced, and URL should be independently documented in the final submission.

The Policy Journey at a Glance

The progression behind this proposal can be summarized as follows:

1. Public-coverage advocacy identified the affordability problem.

The starting principle was that appropriate professional mental-health care should not be inaccessible primarily because of cost.

↓

2. FASS tested free human support as an accessible bridge.

It explored continuity before, between, and alongside professional care.

↓

3. FASS revealed the capacity limits of human-only delivery.

Removing the price charged to the user did not eliminate the scarcity of human labour.

↓

4. FASSLING™ explored continuous and multilingual availability.

AI introduced a possible way to expand bounded foundational availability beyond the operating constraints of one-to-one human delivery.

↓

5. Frontline experience within government-funded service delivery raised a public-finance question.

Governments already purchase substantial amounts of professional and structured supportive capacity.

↓

6. Frontline observation suggested that some structured functions are increasingly suitable for AI assistance.

Not every valuable function necessarily requires the same level of professional attention.

↓

7. AI created the possibility of distinguishing the cost of availability from the cost of expertise.

Computational infrastructure may purchase scale and continuity differently from professional human labour.

↓

8. The approximate 50/50 demonstration model emerged as a testable resource-allocation hypothesis.

It asks whether a comparable public envelope can be divided between scalable digital capacity and protected professional human capacity.

↓

9. The Care Guarantee and Continuity Guarantee emerged as complementary long-term public aspirations.

One addresses affordability of expertise; the other addresses gaps in availability.

↓

10. The remaining question became institutional.

If evidence demonstrates that the underlying capability serves a persistent public need, should successful social innovation mature into durable, technology-neutral public infrastructure?

The innovation therefore did not begin with AI. **It began with listening.**

The Central Conclusion of the Policy and Practice Trajectory

Taken together, these experiences inform the central institutional proposition of this submission.

- **Social entrepreneurs can identify needs, remain close to frontline experience, and demonstrate possibilities.**
- **Artificial intelligence can create scale where scale is appropriate and safely governed.**
- **Helping professionals can provide judgment, therapeutic depth, intervention, safety, accountability, and genuine human connection.**
- **Government is uniquely positioned to establish public standards, protect equitable access, clarify responsibilities, fund and evaluate infrastructure, support workforce transition, and ensure that socially necessary capabilities can endure.**

The proposal therefore asks government to evaluate more than a technology.

It asks government to determine whether a new **public capability** has become technologically and institutionally possible.

That capability is: **Free-at-the-point-of-appropriate-use, continuous, multilingual, trauma-informed, accessible, privacy-preserving, human-supervised foundational emotional-support infrastructure that protects professional care and reduces the likelihood that people are left with absolutely nowhere appropriate and safe to turn.**

If evidence supports such a model, the long-term objective should not be to preserve one founder's project indefinitely. It should be to preserve the underlying public capability the project was created to explore.

- **Innovation can reveal the possibility.**
- **Professional expertise can safeguard it.**
- **Technology can help scale it.**
- **Government can determine whether it should endure.**