AI Agents in Human Systems and Material Sciences—A Holistic Framework for Threat / Hazard Analysis, Vulnerability Assessments, and Risk Analysis

Chate Asvanonda
Neuro Nodal
Chate@NeuroNodal.com

Bruce Redinger
Neuro Nodal
Bruce@NeuroNodal.com

Abstract

AI agents are at the forefront of transforming how both material sciences and human systems are managed, ensuring real-time vulnerability assessments, threat analysis, and predictive risk management. These AI agents dynamically adjust to their environment, specializing as needed to suit the complexities of the system they oversee. Whether embedded in a Digital Twin (DT) of a power grid or a human body, AI agents continuously monitor the system's health, anticipate threats, and offer optimized solutions for system resilience. This white paper explores how AI agents, by leveraging multi-dimensional data and real-time simulations, ensure system efficiency and resilience across both material science (e.g., Critical Infrastructure/Key Resources [CI/KRs]) and human systems (anatomical, physiological, psychological…), driving predictive intelligence and proactive solutions.

Executive Summary

This white paper titled "AI Agents in Human Systems and Material Sciences—A Holistic Framework for Threat / Hazard Analysis, Vulnerability Assessments, and Risk Analysis" explores how AI agents are revolutionizing the management of complex systems across material sciences (such as the 16 Critical Infrastructure/Key Resources [CI/KRs]) and human systems (such as healthcare and physiology). These AI agents are dynamic, adaptive entities that morph into specialized expert agents as needed, allowing them to continuously monitor, assess, and optimize the systems they oversee. By integrating real-time data and performing multi-dimensional risk assessments, AI agents ensure that systems are resilient, vulnerable areas are identified, and threats are mitigated before they become critical. This framework presents a novel approach to system management, positioning AI agents as the central actors within Digital Twins (DTs)—virtual models of physical or human systems that evolve continuously in response to real-world conditions. Unlike traditional AI models that often operate within predefined silos, the AI agents proposed in this white paper can adapt to any system, whether it's monitoring the health of a power grid or managing a patient’s medical condition. These agents interact with one another, share insights, and collaboratively develop solutions tailored to the unique needs of each system.

Key aspects of this framework include:

AI Agents as Morphing Experts: The ability of AI agents to morph into expert roles based on the system’s vulnerabilities and needs allows them to seamlessly transition between domains. For example, in material systems, they may act as engineers or cybersecurity specialists, while in human systems, they might function as cardiologists or nutritionists.

Multi-Dimensional Risk and Vulnerability Assessments: AI agents operate across the physical, abstract, and computational dimensions to provide holistic risk analysis. They continuously monitor real-time data, simulate potential threats, and conduct concentric threat analysis within defined areas of interest, ensuring that vulnerabilities are identified and mitigated.

Collaboration and Scalability: AI agents collaborate within a multi-agent system, interacting with other agents to develop optimal solutions. They can scale to handle increasingly complex systems by spawning additional specialized agents as needed, making this approach highly adaptable across industries.

Cross-Domain Application: The proposed framework is domain-agnostic and can be applied across a wide range of sectors, from healthcare and public infrastructure to cybersecurity and finance. This makes it a flexible, scalable solution that can optimize both human and material systems in real-time.

Proactive Risk Management: AI agents enable a shift from reactive to proactive system management by continuously monitoring and predicting system vulnerabilities and threats.

Improved Resilience: This approach significantly enhances the resilience of critical infrastructure and healthcare systems by preemptively addressing weaknesses and optimizing performance.

Real-Time Decision Making: AI agents provide real-time insights that enable better decision-making across sectors, leading to fewer system failures, better healthcare outcomes, and more efficient infrastructure management.

Scalability Across Industries: The ability of AI agents to specialize and scale ensures that this framework can be applied across multiple domains, making it highly flexible and effective in a variety of settings.

This white paper presents a unique and forward-thinking framework for the application of AI agents in managing Digital Twins across both material and human systems. It addresses critical gaps in existing AI models by enabling real-time, multi-domain intelligence and decision-making, making it an essential approach for scholars and professionals in the AI community to consider. If implemented, this framework could revolutionize how industries manage complex systems, leading to improved efficiency, resilience, and risk management across sectors.

Introduction: The Expanding Role of AI Agents in System Management

AI agents are fundamentally transforming how we manage complex systems across both material sciences and human systems. Unlike traditional models that operate in static, predefined roles, AI agents exhibit a dynamic flexibility that allows them to morph into specialized experts as needed. These agents continuously monitor and assess the systems they oversee, whether in Digital Twins (DTs) that model critical infrastructure, such as power grids and transportation networks, or in models representing human physiological systems. AI agents play a crucial role in ensuring that vulnerabilities are swiftly identified, threats are mitigated, and performance is optimized in real time. This capability is made possible by their ability to harness real-time data and integrate information from multiple dimensions—including physical, abstract, and computational dimensions. This multi-dimensional approach enables AI agents to adapt their actions to the specific needs of the system, offering tailored solutions for risk management and resilience.

In this white paper, we explore the pivotal role AI agents play in both material and human systems. Their unique capacity to dynamically adjust, specialize, and collaborate across domains allows them to ensure the smooth and secure operation of diverse systems. Whether the task is protecting critical infrastructure from cyberattacks or monitoring a patient's vital signs in real-time, these agents excel at proactively managing risks and enhancing system performance through continuous, adaptive learning and decision-making.

AI Agents in Material Science and Human Systems: A Unified Approach

AI Agents in Material Science (CI/KRs)

In material sciences, AI agents are deployed to monitor and manage Digital Twins that represent critical infrastructure, such as energy grids, water systems, transportation networks, and healthcare facilities. These agents continuously assess real-time conditions, ensuring that vulnerabilities are identified, and threats are anticipated.

For example, an AssetMonitor AI may track the wear and tear of key components in a power grid, while a CyberDefense AI analyzes the grid for digital intrusions.

Through their dynamic capabilities, these agents morph into specialized roles, depending on the real-time needs of the system:

GridGuardian AI could oversee energy distribution to ensure the grid meets demand during a peak usage period.

WeatherWatch AI monitors environmental conditions like storms or extreme heat, alerting other agents to adjust operations and mitigate potential infrastructure failures.

By using real-time data across multiple dimensions, AI agents make critical decisions that optimize system resilience and efficiency in material systems.

AI Agents in Human Systems

AI agents are equally essential in human systems, where they manage Digital Twins that represent a real-time virtual model of an individual's physiological state. In this context, AI agents continuously monitor medical data, vital signs, and environmental factors, transforming their roles based on the patient's needs.

For instance:

A CardioExpert AI may monitor a patient's cardiovascular system for abnormalities, while a Nutritionist AI analyzes dietary data to optimize health outcomes.

Similarly, a StressManagement AI may evaluate stress levels, collaborating with other AI agents to recommend lifestyle adjustments or therapeutic interventions.

This real-time adaptability allows AI agents in human systems to move from reactive healthcare to a more predictive and personalized approach, improving patient outcomes by preemptively addressing risks and optimizing health interventions.

Collaborative Capabilities of AI Agents and LLMs: A Multi-Model Synergy for Digital Twins

AI agents and Large Language Models (LLMs) exhibit a substantial ability to collaborate, significantly enhancing their performance when they reference outputs from other models or agents. This collaborative behavior enables the generation of more accurate, refined responses by synthesizing insights from multiple sources. By leveraging outputs from diverse models or perspectives, AI systems can produce higher-quality assessments, similar to how LLMs improve their results by integrating outputs from different models or layers.

Key Components of Collaboration in AI Agents and LLMs

This collaborative capability can be structured in role-based interactions, which mirror structures found in both LLMs and AI agent ecosystems. In such systems, models or agents often take on one of two primary roles:

Proposers and Aggregators:

Proposers are responsible for generating initial outputs or reference responses, providing raw insights based on their specific data or task focus. For example, an AI agent might propose an initial risk assessment based on the energy consumption data in an industrial environment.

Aggregators synthesize the proposals from multiple sources into a single, high-quality output. In LLM architectures, aggregators function similarly, combining outputs from multiple components or models to produce a cohesive result. This mirrors the process of integrating various insights into a comprehensive risk assessment in AI-driven systems.

Application of Collaborative Roles in Digital Twins

In Digital Twin environments, this structure of proposers and aggregators is reflected in how various AI agents collaborate to manage different dimensions of a system. For instance:

Proposer agents are tasked with monitoring specific data streams or system vulnerabilities, such as detecting anomalies in energy consumption or identifying irregularities in equipment performance.

Aggregator agents compile the information from these proposers, creating a comprehensive output, such as an integrated risk assessment or maintenance strategy.

This collaboration allows Digital Twins to operate with real-time, multi-dimensional awareness. The agents responsible for proposing insights can specialize in different data dimensions—be it physical sensor data, abstract risk models, or computational simulations—while the aggregators integrate this information into actionable outcomes.

Enhancing Collaboration through Knowledge Transfer

In AI systems, knowledge transfer between models enhances collaborative behavior. The knowledge or insights generated by one model can be distilled and utilized by another model or agent to improve performance. This transfer of knowledge mirrors the concept of an ensemble of models, where insights from several models are synthesized into a single, more deployable system.

In Digital Twin systems, AI agents collaborate by sharing knowledge in a similar manner. Proposer agents can deliver their specialized insights to an aggregator agent, which distills and integrates the inputs from multiple sources. This method of collaboration ensures that the final output is both efficient and informed by multiple perspectives, improving decision quality and system performance.

Example of Knowledge Transfer in Digital Twins:

In a smart factory environment, proposer agents could monitor various aspects of the production line, such as machine temperature, output speed, and energy usage. Each proposer delivers its specialized analysis, and an aggregator agent synthesizes this data into an overall production health assessment. This collaborative process enables the Digital Twin to predict potential failures, optimize performance, and schedule maintenance based on comprehensive insights.

Collaborative Decision-Making for Multi-Dimensional Systems

In many systems, especially those involving Digital Twins, decision-making requires integrating data from multiple dimensions—physical, abstract, and computational. Collaborative AI agents can adopt proposer and aggregator roles to ensure that data from all these dimensions is captured, processed, and synthesized.

- *Physical Dimension*: Proposer agents monitor real-time sensor data such as temperature, pressure, or speed from physical components of a system.
- *Abstract Dimension*: Agents focus on non-physical factors like risk perceptions, operational thresholds, or economic considerations.
- *Computational Dimension*: Proposers analyze simulations and predictive models, projecting future system states based on current and historical data.

Aggregators then combine this multi-dimensional information into a single, integrated decision output. For example, in healthcare, an aggregator might combine physical health data, predictive risk models, and patient history to generate an accurate prognosis and treatment plan.

Benefits of Collaborative AI Agent Systems

Collaborative AI agents offer significant benefits in Digital Twin environments and beyond by enabling systems to generate richer, more accurate insights through distributed processing and multi-model integration. The key benefits include:

Improved Decision Accuracy: By integrating insights from multiple proposer agents, collaborative AI systems make better-informed decisions based on diverse sources of data, leading to greater accuracy.

Scalability: This collaborative approach scales well with increasing system complexity, as tasks can be distributed across proposer agents and aggregated for efficient real-time decision-making.

Efficiency through Knowledge Sharing: The process of sharing knowledge between agents allows for faster and more efficient decision-making while preserving accuracy. This method enables real-time, large-scale decision-making without requiring excessive computational power.

Multi-Dimensional Awareness: Collaboration ensures that data from all relevant dimensions is considered in decision-making, resulting in more comprehensive and informed outputs. For instance, insights from physical, abstract, and computational dimensions are synthesized to address complex system behaviors.

The collaborative capabilities of AI agents and LLMs play a pivotal role in improving the performance of systems, particularly in environments like Digital Twins. By adopting roles such as proposers and aggregators, AI agents can generate real-time, multi-dimensional insights that enhance decision-making quality. The ability to collaborate across models, share knowledge, and synthesize inputs from diverse perspectives ensures that AI-driven systems are equipped to handle the complexity of real-world environments. This collaboration not only improves accuracy and scalability but also provides a highly efficient and robust approach to managing complex systems.

Mixture-of-Agents (MoA) Methodology in AI Agent Collaboration

The Mixture-of-Agents (MoA) methodology further enhances this collaborative capability by allowing AI agents to operate across multiple layers of decision-making, similar to how LLMs collaborate across layers of model outputs. The MoA framework enables multiple AI agents to process inputs concurrently, generating diverse responses and continuously refining them through iterative collaboration. This iterative refinement maximizes the potential of AI agents in both human and material systems, as agents continuously adjust their outputs based on new data and insights from other agents.

In a material system, such as a power grid, MoA could involve multiple layers of AI agents analyzing different facets of grid performance—some agents monitoring physical infrastructure while others focus on cybersecurity threats. In a human system, MoA might involve AI agents specialized in different medical domains, collectively synthesizing data to optimize treatment plans. Each layer of agents contributes to a progressively more refined output, ensuring that every possible vulnerability is monitored, and that threats are continuously mitigated.

AI Agents Driving System Resilience and Optimization

This paper highlights the transformative role of AI agents in managing complex systems across both material sciences and human systems. By dynamically morphing into expert roles based on the real-time needs of the system, AI agents ensure that vulnerabilities are identified, threats are mitigated, and system performance is optimized. These agents leverage collaborative frameworks and methodologies like Mixture-of-Agents, enabling them to adapt, specialize, and provide continuously improving solutions across multiple domains. Through their capacity to integrate multi-dimensional data and work collaboratively, AI agents represent the future of proactive system management, ensuring resilience, safety, and optimal performance in a wide range of applications—from critical infrastructure to personalized healthcare.

AI Agents in Material Science and Human Systems: A Unified Approach

AI Agents in Material Science (CI/KRs)

In material sciences, AI agents oversee the operation and maintenance of Digital Twins that represent critical infrastructure, such as energy grids, water systems, transportation networks, and healthcare facilities. These agents continuously monitor real-time conditions, ensuring that vulnerabilities are detected, and threats are anticipated. For example, an AssetMonitor AI may continuously track the wear and tear of key components in a power grid, while a CyberDefense AI monitors for digital intrusions that could compromise the system's integrity.

AI agents collaborate, analyze incoming data, and morph into specialized roles as necessary. For instance:

- GridGuardian AI might oversee energy distribution, making sure supply meets demand during peak usage.
- WeatherWatch AI would monitor environmental conditions like extreme weather events, alerting other agents about potential disruptions that may affect the infrastructure.

This flexibility allows AI agents to assess the multi-dimensional risks affecting material infrastructure, ensuring resilience and efficiency in operations.

AI Agents in Human Systems

In human systems, AI agents manage Digital Twins that represent a virtual model of an individual's physiological state. These AI agents adapt based on the patient's specific needs, analyzing data in real-time from medical devices, wearables, and health records. For instance, a CardioExpert AI might monitor a patient's cardiovascular system, while a Nutritionist AI may focus on optimizing diet based on the individual's health profile.

AI agents in human systems collaborate to ensure that health risks are preemptively managed:

- EndoHealth AI may focus on hormone balance in diabetic patients, collaborating with ImmunoDefense AI to ensure that immune system function remains stable.
- StressManagement AI could analyze stress responses in real-time and suggest lifestyle adjustments, working alongside CardioExpert AI to reduce cardiovascular risk.

By morphing into these specialized expert roles, AI agents offer a proactive, real-time management approach that allows healthcare providers to move from reactive care to a highly personalized, predictive model.

AI Agents as the Central Operators of System Resilience

AI agents are the central actors in ensuring system resilience, regardless of whether they are deployed in material systems or human systems. Their ability to adapt to the specific needs of the system makes them uniquely capable of monitoring vulnerabilities, conducting threat analysis, and providing real-time solutions for optimizing system performance.

Morphing into Expert Agents

AI agents have the inherent ability to morph into different specialized roles, becoming expert agents for the domain they oversee. In a power grid, this could mean AI agents transforming into GridHealth AI or CyberWatch AI, focusing on infrastructure vulnerabilities and cybersecurity risks. In a hospital setting, AI agents could become InfectionControl AI, tracking disease spread, or PatientFlow AI, optimizing hospital resources during peak demand.

This adaptability allows the AI agents to provide highly tailored, domain-specific insights:

- In material systems, they may simulate structural wear and tear, assess energy distribution patterns, or model the impact of severe weather events.
- In human systems, they may predict disease progression, adjust medication protocols, or recommend lifestyle interventions based on real-time health data.

AI agents' ability to specialize means that they provide not just broad insights but detailed, highly relevant intelligence that can drive real-time decision-making.

AI Agent Collaboration and Multi-Agent Systems

AI agents do not work in isolation; they form collaborative multi-agent systems where they exchange information and expertise to arrive at the best solutions for a given system. Multiple agents can be spawned to handle complex problems, each taking on a specialized task and communicating with other agents to ensure a holistic approach.

For instance:

- In a power grid, a GridGuardian AI might collaborate with WeatherWatch AI to redistribute energy in anticipation of a severe weather event, while CyberWatch AI ensures that the grid's SCADA systems are secure from cyber threats.
- In a human system, a CardioExpert AI might work with Nutritionist AI to develop a treatment plan for a patient with heart disease, taking into account dietary factors, lifestyle habits, and genetic predispositions.

These collaborative systems ensure that every angle of the system is monitored and optimized, allowing the AI agents to provide the best possible recommendations for resilience and performance.

Vulnerability Assessments and Threat Analysis in the Dimensional Space

AI agents are tasked with continuously monitoring systems for weaknesses, performing vulnerability assessments and conducting threat analysis within the dimensional space. The dimensional space comprises the physical, abstract, and computational dimensions, which together provide a complete understanding of the system's vulnerabilities and threats.

Vulnerability Assessments by AI Agents

AI agents are equipped with the ability to identify and assess vulnerabilities based on real-time data from the Digital Twin. These vulnerabilities could be physical, such as a crack in a bridge or an abnormality in a patient's vital signs, or they could be abstract, such as a potential market shift or regulatory change that could impact operations.

- In material systems, agents like AssetMonitor AI might detect aging infrastructure that requires repair or replacement before a critical failure occurs.
- In human systems, agents like CardioExpert AI might detect early signs of heart disease, triggering a series of interventions designed to prevent a major health event.

Threats in the Dimensional Space

Threats arise in the dimensional space, where AI agents evaluate both internal and external factors that could impact the system. AI agents use concentric threat analysis, assessing how threats in a specific Area of Interest (AoI) could exploit vulnerabilities. For example, a WeatherWatch AI might predict how an incoming storm will affect an energy grid, working with GridGuardian AI to prepare mitigation strategies.

- ***Physical dimension***: AI agents analyze sensor data from physical components, such as temperature, pressure, or energy flow.
- ***Abstract dimension***: AI agents consider non-physical factors, such as regulatory changes, market conditions, and human behaviors.
- ***Computational dimension***: AI agents run simulations and predictive models to anticipate how vulnerabilities and threats will evolve over time.

This multi-dimensional analysis allows AI agents to develop a complete picture of the system's current and future risks.

Risk Analysis and Solution Optimization by AI Agents

Once vulnerabilities and threats are identified, AI agents collaborate to conduct a comprehensive risk analysis. By combining their specialized knowledge, these agents provide insights into how best to mitigate risks and optimize system performance.

- ***In material systems***, AI agents might simulate how increased energy demand during a heatwave could lead to transformer failures, recommending that load be redistributed to prevent outages.
- ***In human systems***, AI agents might predict that a patient is at high risk for a stroke, recommending changes in medication and lifestyle to reduce the risk.

The AI agents then propose the most effective solutions, ensuring that the Digital Twin is optimized and that the real-world system—whether it's a human body or an infrastructure network—operates with maximum resilience.

AI Agents Driving Resilience Across Material Science and Human Systems

AI agents are the driving force behind system resilience and optimization, enabling both material science systems (like the 16 CI/KRs) and human systems to be managed proactively and effectively. By morphing into specialized expert roles, these AI agents continuously monitor systems, conduct vulnerability assessments, and perform risk analysis in real time. Their ability to work collaboratively across multiple dimensions—physical, abstract, and computational—allows them to anticipate and mitigate threats, ensuring that systems remain resilient and efficient. As we look toward the future, the role of AI agents in managing complex, interconnected systems will become even more critical. Their adaptability, scalability, and intelligence will drive the next generation of system optimization, from healthcare to infrastructure, ensuring that we can predict, prevent, and respond to risks in real-time across any domain.

Unique Aspects of this White Paper

- ***AI Agents as Dynamic Experts***: AI agents typically operate within a specific, predefined domain (e.g., cybersecurity, predictive analytics). This proposal pushes the boundaries by suggesting that AI agents can morph into specialized expert agents depending on the real-time needs of the system. They adapt and take on new roles, whether it's analyzing energy consumption for a power grid or monitoring a patient's heart rate for a healthcare Digital Twin. This kind of agentic flexibility is not widely explored in current AI models, which tend to remain more static in their functionality.
- ***Integration of Multi-Dimensional Threat and Risk Analysis***: The integration of physical, abstract, and computational dimensions into a cohesive system for risk analysis is comprehensive and forward-looking. The ability of AI agents to operate across multiple dimensions and synthesize complex, multi-domain data is a step forward from current siloed approaches to AI. It allows for real-time risk mitigation that factors in everything from physical asset conditions to regulatory environments and human behaviors.
- ***Cross-Domain Applicability***: This proposal highlights the versatility of AI agents. Instead of limiting their application to one domain (e.g., healthcare or infrastructure), the agents can operate across both human systems and material systems, morphing into whatever expert roles are needed. This cross-domain adaptability is rare in current AI research, which tends to compartmentalize AI solutions.

Gaps Addressed in the AI Community

This white paper addresses several key gaps in current AI research and application:

- ***Siloed AI Models***: Many existing AI models operate within rigidly defined boundaries, focusing on a single domain (e.g., healthcare diagnostics or industrial operations). These models struggle to integrate real-time, multi-domain intelligence that includes spatial, temporal, and abstract factors. This white paper proposes AI agents that can transcend these silos, operating seamlessly across domains and providing holistic risk assessments that include diverse data streams.
- ***Real-Time Multi-Domain Decision-Making***: While AI can process large datasets, the ability to synthesize real-time data from multiple dimensions—such as physical infrastructure, human behavior, and predictive simulations—is currently underdeveloped. AI systems often miss out on the opportunity to leverage real-time updates for dynamic decision-making. This paper fills that gap by promoting AI agents that can make real-time decisions by constantly evolving with the system they manage.

- ***Vulnerability and Threat Assessment in Complex Systems***: Existing AI systems focus heavily on pattern recognition and predictive models but may lack the ability to assess vulnerabilities dynamically in real time, especially within complex systems like critical infrastructure and healthcare. The proposal's emphasis on concentric threat analysis and multi-dimensional risk analysis provides a novel framework for real-time monitoring and prediction across interconnected systems.

Positive Ramifications if Implemented

If the concepts outlined in this white paper are implemented, the positive ramifications could be profound, both in the AI science community and across various industries:

- ***Proactive, Predictive System Management***: Instead of reacting to failures or emergencies, AI agents would enable proactive management of both human and material systems. In healthcare, this could mean preventing major health crises by continuously monitoring patient vitals and adapting treatment plans in real time. In critical infrastructure, AI agents could anticipate system failures (e.g., in power grids or transportation networks) and take preemptive action to mitigate risks.
- ***Increased System Resilience***: By having AI agents continually monitor and adapt to real-time conditions, critical systems (whether material or human) will become more resilient. AI agents can collaborate, simulate potential threats, and adapt quickly to new challenges, ensuring that infrastructure and healthcare systems can withstand crises like natural disasters, pandemics, or cyberattacks.
- ***Scalable AI Solutions Across Domains***: This proposal offers a scalable AI solution that is domain-agnostic, meaning it can apply to virtually any system—healthcare, infrastructure, finance, agriculture, and beyond. The ability to spawn multiple specialized agents allows the system to scale as needed, making it flexible and adaptable for small systems like personal healthcare or large-scale operations like national power grids.
- ***More Informed Decision-Making in Real Time***: The continuous assessment of vulnerabilities and threats across multiple dimensions means that decision-makers (whether doctors, engineers, or policymakers) will have more accurate, real-time data to base their decisions on. This could lead to better resource allocation, fewer emergency interventions, and improved outcomes in both human health and infrastructure management.
- ***Reduction of System Failures and Risks***: Because AI agents will constantly monitor and predict vulnerabilities, we can expect a significant reduction in system failures across various sectors. In healthcare, this could result in fewer hospital readmissions or severe

health events, while in infrastructure, it might prevent costly breakdowns in energy distribution or transportation systems.

- ***Cross-Industry Collaboration***: Implementing this framework will likely encourage greater cross-industry collaboration, as AI agents will need to access a wide variety of data streams. Healthcare, infrastructure, IT, and even finance could start sharing critical data with AI agents, enhancing the overall predictive capabilities and fostering interdisciplinary innovation.

This white paper presents a distinct and valuable proposal for AI scholars and practitioners to consider, as it suggests a framework that combines flexibility, real-time adaptation, and multi-domain applicability. By proposing AI agents that morph into expert roles based on the system's needs and operate within multi-dimensional frameworks, this approach could significantly enhance system resilience and predictive intelligence across industries. It addresses critical gaps in current AI systems, particularly the siloed nature of many models and their inability to operate fluidly across domains. If implemented, this model could transform how we manage complex systems, leading to better decision-making, reduced risk, and optimized performance across a wide range of human and material systems. This proposal could serve as a foundation for new research, cross-sector partnerships, and technological advancements that push the boundaries of what AI can achieve in both human and material domains.

Formula

$$R(t) = f\left[(V(t) \times T(t)) \times (P(t) + A(t) + C(t))\right] + \sum_{i=1}^{n} (w_i \times \mathrm{AI}_i(t) \times \mathrm{DT}(t))$$

Where:

R(t): Risk score at time ***t***, indicating the overall risk level based on AI agent analysis.

f: A function that integrates and weights the components of the formula based on the system's complexity and the agents' predictions.

V(t): Vulnerabilities identified by AI agents at time ***t*** within the Digital Twin.

T(t): Threats identified by AI agents at time ***t*** that could exploit the system's vulnerabilities.

P(t): The ***Physical dimension*** at time ***t***, encompassing real-time sensor data and physical attributes of the system (e.g., health metrics in human systems or structural integrity in material systems).

A(t): The ***Abstract dimension*** at time ***t***, including non-physical variables such as regulations, market trends, human behavior, and policies affecting the system.

C(t): The ***Computational dimension*** at time ***t***, representing predictive models, simulations, and future projections that AI agents use to forecast possible outcomes.

$\sum_{i=1}^{n}(w_i \times \mathrm{AI}_i(t) \times \mathrm{DT}(t))$: The contribution of ***n*** AI agents, each with a weighting factor ***wi***, acting within the Digital Twin. Each AI agent ***AI*** monitors specific aspects of the Digital Twin and collaborates to provide insights.

AIi(t): The AI agent responsible for a particular dimension or aspect of the Digital Twin at time ***t***, such as a healthcare specialist or an infrastructure expert.

DT(t): The Digital Twin model representing the system at time ***t***, continuously updated with real-time data reflecting the current state of the system.

wi: Weighting factor that adjusts the influence of each AI agent on the final risk score, depending on the agent's domain expertise and the system's requirements.

Step-by-Step Explanation:

1. Risk Score R(t):

The final risk score ***R(t)*** represents the overall risk level for the system (whether a human or material system) at a specific time. This score is determined based on the inputs provided by the AI agents and their analysis of vulnerabilities, threats, and multi-dimensional data.

- A higher risk score indicates that the system is more vulnerable or facing greater threats.
- A lower risk score suggests that the system is stable, and its vulnerabilities are well managed.

2. Vulnerabilities V(t) and Threats T(t):

- ***V(t) (Vulnerabilities)***: AI agents continuously monitor the Digital Twin and identify vulnerabilities at time ***t***. These vulnerabilities represent weaknesses within the system that could lead to failures or issues, such as an overworked transformer in a power grid or a patient's weakened immune system.
- ***T(t) (Threats)***: AI agents simultaneously assess threats that could exploit these vulnerabilities. These threats may be external, such as extreme weather or cybersecurity

risks in material systems, or internal, such as health complications or disease progression in human systems.

Multiplying vulnerabilities by threats emphasizes that vulnerabilities alone may not be as critical unless they are exacerbated by threats. For example, a minor vulnerability in a transformer becomes critical if a heatwave (external threat) occurs.

Dimensional Integration P(t) + A(t) + C(t):

The next part of the formula reflects how AI agents incorporate data from three key dimensions to assess the system's vulnerabilities and threats in a holistic way:

- ***P(t)*** (Physical dimension): This dimension includes real-time sensor data and physical measurements. In material systems, this could be the temperature of infrastructure components or the energy load on a power grid. In human systems, it could be patient vitals like heart rate or blood pressure.
- ***A(t)*** (Abstract dimension): This captures non-physical variables, such as human behavior, regulatory frameworks, and operational constraints. For example, a financial regulatory change might affect the stability of critical manufacturing, or a patient's lifestyle choices might influence their health outcomes.
- ***C(t)*** (Computational dimension): This includes predictive models and simulations that AI agents use to forecast potential future risks. AI agents run these models based on historical data and real-time conditions to predict how vulnerabilities may evolve.

Adding these dimensions shows that each dimension provides independent, critical insights into the overall risk level. Combining them creates a more complete picture of the system's health.

AI Agents $\sum_{i=1}^{n}(w_i \times \mathrm{AI}_i(t) \times \mathrm{DT}(t))$**:** This part of the formula accounts for the various AI agents operating within the Digital Twin:

- ***AIi(t)*** represents each specific AI agent operating within the system. For example, in a healthcare system, one agent may be focused on cardiology while another monitors mental health. In a material system, different agents might be responsible for energy distribution, cybersecurity, or structural integrity.
- ***wi*** is the weighting factor that adjusts how much influence each agent has on the overall risk score. Certain agents, depending on their expertise or the system's immediate needs,

might carry more weight in determining the final score. For instance, during a heatwave, an AI agent responsible for cooling systems may have a higher weighting factor than one responsible for routine maintenance.

- ***DT(t)*** represents the Digital Twin model itself, which is continuously updated with real-time data. The Digital Twin reflects the actual system (whether a power grid or a human body) and is the core model on which AI agents base their predictions and assessments.

Summing over ***n*** AI agents shows that multiple agents may operate simultaneously, each contributing to the overall risk analysis. As more agents monitor different aspects of the system, they collaborate to provide a comprehensive view of the system's health.

Function f:

The function ***f*** integrates the different components of the formula, applying weighting and adjustments based on the complexity of the system, the severity of vulnerabilities and threats, and the precision of the AI agents' predictions. This function ensures that the final risk score is accurately reflective of the system's current status.

For example, ***f*** may emphasize vulnerabilities and threats during times of crisis (e.g., a pandemic or natural disaster) but may give greater weight to computational predictions when preparing for long-term risk management or infrastructure upgrades.

Example Application:

In a hospital's IT system, AI agents monitor cybersecurity (CyberDefense AI), patient flow (PatientFlow AI), and medical equipment status (AssetMonitor AI):

1. ***V(t)***: Cyber vulnerabilities are identified due to outdated firewalls.
2. ***T(t)***: Threats are identified from increased cyberattacks during a pandemic.
3. ***P(t)***: Physical dimension monitors the number of connected medical devices.
4. ***A(t)***: Abstract dimension includes hospital policies on data security and staff compliance with cybersecurity protocols.
5. ***C(t)***: Computational dimension models potential impacts of a data breach on patient care, based on historical patterns of attacks.

Each AI agent collaborates, and the formula produces a risk score ***R(t)*** that informs the hospital's IT team about which areas to address (e.g., firewall updates, training staff) to mitigate the cybersecurity risks in real time.

This formula captures the holistic role of AI agents in continuously assessing vulnerabilities, identifying threats, and generating real-time risk scores for systems represented by Digital Twins. The flexibility of the formula allows it to be applied across both material sciences and human systems, offering a comprehensive approach to managing system resilience, risk, and optimization.

References

1. Agrawal, P., Nair, A., Abbeel, P., Malik, J., and Levine, S. (2016). Learning to poke by poking: Experiential learning of intuitive physics. CoRR, abs/1606.07419.
2. Asvanonda, C., and Redinger, B., Digital Twins in Human Systems and Material Sciences - A Framework for AI-Driven System Resilience VAs and Risk Analysis, 2024.
3. Asvanonda, C., and Redinger, B., Dimensional Integration in AI - A Multi-Dimensional Framework for Risk Management Predictive Analytics and Resilience in Critical Infrastructure Systems, 2024.
4. Asvanonda, C., and Redinger, B., Enhancing Large Language Models for Digital Twins - A Deep Learning Approach with Domain-Specific Fine-Tuning, 2024.
5. Babaeizadeh, M., Finn, C., Erhan, D., Campbell, R. H., and Levine, S. (2017). Stochastic variational video prediction. CoRR, abs/1710.11252.
6. Baevski, A., Zhou, Y., Mohamed, A., and Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, Advances in Neural Information Processing Systems, volume 33, pages 12449–12460. Curran Associates, Inc.
7. Bardes, A., Ponce, J., and LeCun, Y. (2021). Vicreg: Variance-invariance-covariance regularization for self-supervised learning. In International Conference on Learning Representations (ICLR 2022). arXiv preprint arXiv:2105.04906.
8. Battaglia, P., Pascanu, R., Lai, M., Jimenez Rezende, D., et al. (2016). Interaction networks for learning about objects, relations and physics. Advances in neural information processing systems, 29.
9. Becker, S. and Hinton, G. E. (1992). Self-organizing neural network that discovers surfaces in random-dot stereograms. Nature, 355(6356):161–163.
10. Bertsekas, D. (2019). Reinforcement learning and optimal control. Athena Scientific.
11. Bordes, A., Usunier, N., Chopra, S., and Weston, J. (2015). Large-scale simple question answering with memory networks. arXiv:1506.02075.
12. Bromley, J., Guyon, I., LeCun, Y., Sackinger, E., and Shah, R. (1994). Signature verification using a "siamese" time delay neural network. In NeurIPS.
13. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T.,

Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are fewshot learners. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, Advances in Neural Information Processing Systems, volume 33, pages 1877–1901. Curran Associates, Inc.

14. Bryson, A. and Ho, Y. (1969). Applied optimal control. Blaisdell, Waltham, MA. Carey, S. (2009). The Origin of Concepts. Oxford University Press, New York, New York, USA.
15. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S. (2020). End-to-end object detection with transformers. In 16th European Conference, Glasgow, UK (ECCV 2020), page 213–229.
16. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., and Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. In Advances in Neural Information Processing Systems.
17. Carreira-Perpi˜n´an, M. A. and Hinton, G. (2005). On contrastive divergence learning. In Cowell, R. G. and Ghahramani, Z., editors, Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics, volume R5 of Proceedings of Machine Learning Research, pages 33–40. PMLR. Reissued by PMLR on 30 March 2021.
18. Chen, T., Kornblith, S., Swersky, K., Norouzi, M., and Hinton, G. (2020a). Big self-supervised models are strong semi-supervised learners. In NeurIPS.
19. Chen, X., Fan, H., Girshick, R., and He, K. (2020b). Improved baselines with momentum contrastive learning. arXiv preprint arXiv:2003.04297.
20. Chopra, S., Hadsell, R., and LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), volume 1, pages 539–546. IEEE.
21. Chua, K., Calandra, R., McAllister, R., and Levine, S. (2018). Deep reinforcement learning in a handful of trials using probabilistic dynamics models. CoRR, abs/1805.12114.
22. Craik, K. J. W. (1943). The nature of explanation. University Press, Macmillan.
23. Cyber Security and Infrastructure Security Agency, Critical Infrastructure Sectors, https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors
24. Dai, X., Tong, S., Li, M., Wu, Z., Psenka, M., Chan, K. H. R., Zhai, P., Yu, Y., Yuan, X., Shum, H.-Y., and Ma, Y. (2022). Ctrl: Closed-loop transcription to an ldr via minimaxing rate reduction. Entropy, 24(4):456.
25. Dehaene, S., Lau, H., and Kouider, S. (2021). What is consciousness, and could machines have it? Robotics, AI, and Humanity, pages 43–56.
26. Denton, E. and Fergus, R. (2018). Stochastic video generation with a learned prior. arXiv preprint arXiv 1802.07687.

27. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
28. Doi, E., Balcan, D. C., and Lewicki, M. S. (2007). Robust coding over noisy overcomplete channels. IEEE Transactions on Image Processing, 16(2):442–452.
29. Ermolov, A., Siarohin, A., Sangineto, E., and Sebe, N. (2021). Whitening for self-supervised representation learning.
30. Evtimova, K. and LeCun, Y. (2022). Sparse coding with multi-layer decoders using variance regularization. arXiv:2112.09214.
31. Finn, C., Goodfellow, I. J., and Levine, S. (2016). Unsupervised learning for physical interaction through video prediction. CoRR, abs/1605.07157.
32. Finn, C. and Levine, S. (2017). Deep visual foresight for planning robot motion. In 2017 IEEE International Conference on Robotics and Automation (ICRA), pages 2786–2793. IEEE.
33. Fragkiadaki, K., Agrawal, P., Levine, S., and Malik, J. (2015). Learning visual predictive models of physics for playing billiards. CoRR, abs/1511.07404.
34. Frye, C., Mijolla, D., Kunesh, M. Mansir, J and Feige, I., Human-interpretable model explainability on high-dimensional data, 2020 https://arxiv.org/abs/2010.07384
35. Gehring, J., Synnaeve, G., Krause, A., and Usunier, N. (2021). Hierarchical skills for efficient exploration. Advances in Neural Information Processing Systems, 34:11553–11564.
36. Goldberger, J., S.Roweis, Hinton, G., and Salakhutdinov, R. (2005). Neighbourhood components analysis. In Saul, L. K., Weiss, Y., and Bottou, L., editors, Advances in Neural Information Processing Systems 17, pages 513–520. MIT Press, Cambridge, MA.
37. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems, pages 2672–2680.
38. Gopnik, A. and Meltzoff, A. N. (1997). Words, Thoughts, and Theories. MIT Press, Cambridge, MA. Gopnik, A., Meltzoff, A. N., and Kuhl, P. K. (2001). The Scientist in the Crib: What Early Learning Tells Us About the Mind. Perennial, New York, NY.
39. Goroshin, R., Bruna, J., Tompson, J., Eigen, D., and LeCun, Y. (2015a). Unsupervised feature learning from temporal data. In International Conference on Computer Vision (ICCV 2015).
40. Goroshin, R., Mathieu, M., and LeCun, Y. (2015b). Learning to linearize under uncertainty. In Advances in Neural Information Processing Systems (NIPS 2015), volume 28.
41. Gottlieb, J., Oudeyer, P. Y., Lopes, M., and Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. Trends in Cognitive Sciences, 17:585–593.

42. Gregor, K. and LeCun, Y. (2010a). Emergence of complex-like cells in a temporal product network with local receptive fields. arXiv preprint arXiv:1006.0448.
43. Gregor, K. and LeCun, Y. (2010b). Learning fast approximations of sparse coding. In Proc. International Conference on Machine learning (ICML'10).
44. Grill, J.-B., Strub, F., Altch´e, F., Tallec, C., Richemond, P. H., Buchatskaya, E., Doersch, C., Pires, B. A., Guo, Z. D., Azar, M. G., Piot, B., Kavukcuoglu, K., Munos, R., and Valko, M. (2020). Bootstrap your own latent: A new approach to self-supervised learning. In NeurIPS.
45. Ha, D. and Schmidhuber, J. (2018a). Recurrent world models facilitate policy evolution. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, Advances in Neural Information Processing Systems, volume 31.
46. Ha, D. and Schmidhuber, J. (2018b). World models. arXiv preprint arXiv:1803.10122.
47. Hadsell, R., Chopra, S., and LeCun, Y. (2006). Dimensionality reduction by learning an invariant mapping. In CVPR.
48. Hafner, D., Lee, K.-H., Fischer, I., and Abbeel, P. (2022). Deep hierarchical planning from pixels. arXiv preprint arXiv:2206.04114.
49. Hafner, D., Lillicrap, T., Fischer, I., Villegas, R., Ha, D., Lee, H., and Davidson, J. (2018). Learning latent dynamics for planning from pixels. arXiv 1811.04551.
50. Hafner, D., Lillicrap, T., Norouzi, M., and Ba, J. (2020). Mastering atari with discrete world models. arXiv preprint arXiv:2010.02193.
51. He, K., Fan, H., Wu, Y., Xie, S., and Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. In CVPR.
52. Henaff, M., Canziani, A., and LeCun, Y. (2019). Model-predictive policy learning with uncertainty regularization for driving in dense traffic. In ICLR-19. arXiv:1901.02705.
53. Henaff, M., Weston, J., Szlam, A., Bordes, A., and LeCun, Y. (2017). Tracking the world state with recurrent entity networks. In International Conference on Learning Representations (ICLR 2017).
54. Hinton, G. and Sejnowski, T. (1983). Optimal perceptual inference. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 448–453, Washington 1983. IEEE, New York.
55. Hinton, G., Vinyals, O., Dean, J., Distilling the Knowledge in a Neural Network, arXiv preprint arXiv:1503.02531v1, 2014.
56. Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., and Griffiths, T. L. (2022). People construct simplified mental representations to plan. Nature, 606(7912):129–136.
57. H´enaff, O. J., Srinivas, A., De Fauw, J., Razavi, A., Doersch, C., Eslami, S. M. A., and van den Oord, A. (2019). Data-efficient image recognition with contrastive predictive coding. In ICML.

58. Janner, M., Li, Q., and Levine, S. (2021). Offline reinforcement learning as one big sequence modeling problem. In Advances in Neural Information Processing Systems.
59. Jing, L., Zbontar, J., et al. (2020). Implicit rank-minimizing autoencoder. Advances in Neural Information Processing Systems, 33:14736–14746.
60. Jordan, M. I. and Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. Cognitive science, 16(3):307–354.
61. Kahneman, D. (2011). Thinking, fast and slow. Macmillan. Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.
62. Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017a). Building machines that learn and think like people. Behavioral and brain sciences, 40.
63. Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017b). Building machines that learn and think like people. Behavioral and Brain Sciences, 40:E253.
64. LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324.
65. LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M., and Huang, F. (2006). A tutorial on energy-based learning. In Bakir, G., Hofman, T., Sch¨olkopf, B., Smola, A., and Taskar, B., editors, Predicting Structured Data. MIT Press.
66. Lerer, A., Gross, S., and Fergus, R. (2016). Learning physical intuition of block towers by example. In Proceedings of the 33nd International Conference on Machine Learning,
67. ICML 2016, New York City, NY, USA, June 19-24, 2016, pages 430–438.
68. Levine, S. (2021). Understanding the world through action. arXiv:2110.12543. https://arxiv.org/abs/2110.12543.
69. Luc, P., Clark, A., Dieleman, S., Casas, D. d. L., Doron, Y., Cassirer, A., and Simonyan, K. (2020). Transformation-based adversarial video prediction on large-scale data. arXiv preprint arXiv:2003.04035.
70. Luc, P., Couprie, C., Lecun, Y., and Verbeek, J. (2018). Predicting future instance segmentation by forecasting convolutional features. In Proceedings of the european conference on computer vision (ECCV), pages 584–599.
71. Luc, P., Neverova, N., Couprie, C., Verbeek, J., and LeCun, Y. (2017). Predicting deeper into the future of semantic segmentation. In Proceedings of the IEEE international conference on computer vision, pages 648–657.
72. Marcus, G. and Davis, E. (2019). Rebooting AI: Building artificial intelligence we can trust. Vintage.
73. Mathieu, M., Couprie, C., and LeCun, Y. (2015). Deep multi-scale video prediction beyond mean square error. In ICLR 16. arXiv preprint arXiv:1511.05440.
74. Mattar, M. G. and Lengyel, M. (2022). Planning in the brain. Neuron, 110(6):914–934. Mercat, J., Gilles, T., El Zoghby, N., Sandou, G., Beauvois, D., and Gil, G. P. (2020).

Multihead attention for multi-modal joint vehicle motion forecasting. In 2020 IEEE International Conference on Robotics and Automation (ICRA), pages 9638–9644. IEEE.
75. Miller, A. H., Fisch, A., Dodge, J., Karimi, A.-H., Bordes, A., and Weston, J. (2016). Key-value memory networks for directly reading documents. In EMNLP-16.
76. Miller, W. T., Sutton, R. S., and Werbos, P. J. (1995). Neural networks for control. MIT press.
77. Misra, I. and Maaten, L. v. d. (2020). Self-supervised learning of pretext-invariant representations. In CVPR.
78. Moerland, Thomas, M., Broekens, J., and Jonker, Catholijn, M. (2020). Model-based reinforcement learning: A survey. arXiv:2006.16712. https://arxiv.org/abs/2006.
79. Mohamed, A., Lee, H.-y., Borgholt, L., Havtorn, J. D., Edin, J., Igel, C., Kirchhoff, K., Li, S.-W., Livescu, K., Maaløe, L., et al. (2022). Self-supervised speech representation learning: A review. arXiv preprint arXiv:2205.10643.
80. Morari, M. and Lee, J. H. (1997). Model predictive control: Past, present and future. Computers and Chemical Engineering, 23:667–682.
81. Murphy, G. L. (2002). The Big Book of Concepts. MIT Press, Cambridge, MA. Nagabandi, A., Kahn, G., Fearing, R. S., and Levine, S. (2017). Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning. CoRR, abs/1708.02596.
82. Narendra, K. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. IEEE Transactions on neural networks, 1(1):4–27.
83. National Geospatial Intelligence Agency (NGA), https://www.nga.mil/
84. Oh, J., Guo, X., Lee, H., Lewis, R. L., and Singh, S. (2015). Action-conditional video prediction using deep networks in atari games. Advances in neural information processing systems, 28.
85. Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583):607–609.
86. Orhan, E., Gupta, V., and Lake, B. M. (2020). Self-supervised learning through the eyes of a child. Advances in Neural Information Processing Systems, 33:9960–9971.
87. Pari, J., Shafiullah, N. M., Arunachalam, S. P., and Pinto, L. (2021). The surprising effectiveness of representation learning for visual imitation. In Robotics Science and Systems 2022. arXiv preprint arXiv:2112.01511.
88. Richalet, J., Rault, A., Testud, J. L., and Papon, J. (1978). Model predictive heuristic control: Applications to industrial processes. Automatica, 14(5):413–428.
89. Riochet, R., Castro, M. Y., Bernard, M., Lerer, A., Fergus, R., Izard, V., and Dupoux, E. (2019). Intphys: A benchmark for visual intuitive physics reasoning. arXiv:1803.07616.
90. Silver, D., Singh, S., Precup, D., and Sutton, R. S. (2021). Reward is enough. Artificial Intelligence, 299:103535.

91. Sobal, V., Canziani, A., Carion, N., Cho, K., and LeCun, Y. (2022). Separating the world and ego models for self-driving. arXiv:2204.07184.
92. Spelke, E. S. and Kinzler, K. D. (2007). Core knowledge. Developmental Science, 10:89–96.
93. Srinivas, A., Jabri, A., Abbeel, P., Levine, S., and Finn, C. (2018). Universal planning networks. CoRR, abs/1804.00645.
94. Srivastava, N., Mansimov, E., and Salakhudinov, R. (2015). Unsupervised learning of video representations using lstms. In Bach, F. and Blei, D., editors, Proceedings of the 32nd International Conference on Machine Learning, volume 37 of Proceedings of Machine Learning Research, pages 843–852, Lille, France. PMLR.
95. Sukhbaatar, S., Weston, J., Fergus, R., et al. (2015). End-to-end memory networks. Advances in neural information processing systems, 28.
96. Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting. ACM Sigart Bulletin, 2(4):160–163.
97. Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1701–1708. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of machine learning research, 11(12).
98. The Federal Emergency Management Agency, National Risk and Capability Assessment, Threat and Hazard Identification and Risk Assessment (THIRA), https://www.fema.gov/emergency-managers/national-preparedness/goal/risk-capability-assessment
99. The United State Geological Survey (USGS), https://www.usgs.gov/
100. Walker, J., Razavi, A., and Oord, A. v. d. (2021). Predicting video with vqvae. arXiv preprint arXiv:2103.01950.
101. Wayne, G. and Abbott, L. (2014). Hierarchical control using networks trained with higherlevel forward models. Neural Computation, 26(10):2163–2193.
102. Wiskott, L. and Sejnowski, T. J. (2002). Slow feature analysis: Unsupervised learning of invariances. Neural computation, 14(4):715–770.
103. Yarats, D., Kostrikov, I., and Fergus, R. (2021). Image augmentation is all you need: Regularizing deep reinforcement learning from pixels. In ICLR.
104. Yu, T., Thomas, G., Yu, L., Ermon, S., Zou, J., Levine, S., Finn, C., and Ma, T. (2020).Mopo: Model-based offline policy optimization. arXiv preprint arXiv:2005.13239.
105. Zaadnoordijk, L., Besold, T., and Cusack, R. (2022). Lessons from infant learning for unsupervised machine learning. Nature Machine Intelligence, 4:510–520.

106. Zbontar, J., Jing, L., Misra, I., LeCun, Y., and Deny, S. (2021). Barlow twins: Selfsupervised learning via redundancy reduction. In International Conference on Machine Learning, pages 12310–12320. PMLR.