Fellows' AI Interest Group

Shaping the RSA's AI Agenda Summary Document
7th July 2023, 11:00 – 14:00, RSA House (The Steps)

In attendance

Fourteen Fellows of the Royal Society for the Encouragement of the Arts, Manufactures and Commerce (RSA), including Marc Goblot, Andy Gordon, Ismael Kherroubi Garcia, Michaela Moher, Silviu Pirvu, Charlie Pownall and Ian Stobie.

Acknowledgements

Author of this report: Ismael Kherroubi Garcia

Brief Overview

This document summarises the Fellows' Artificial Intelligence (AI) Interest Group's event held on 7th July 2023 at RSA House. The session had four distinct parts, with two breaks in between. Ideas are attributed to the group and smaller groups formed in part three of the session. Acronyms used are "AI" (short for "artificial intelligence") and "FRSAs" (short for "Fellows of the RSA").

Part one involved introductions from everybody in the room and their responses to the question "What three big areas should we work on?" The session was designed to allow the responses to this initial question to be iterated and improve through the group's dialogue. Part two was an open discussion to understand the conditions that have enabled past projects to succeed, the reason attendees found the RSA's potential in the AI space exciting, and the foreseeable stumbling blocks to our project.

After a short break, we formed three groups for part three. Each group focused on potential efforts as framed by either of the three pillars of the Design for Life mission: Building Capabilities, Growing Hubs, or Developing Infrastructure. After a second break, the discussion turned to resourcing of our efforts.

Part One: Intros and Ideas

Everybody at the event was asked to share their name and response to the question: **What three big areas should we work on?**

Anonymised responses are offered in the appendix. Some answers were given by people before the event but then couldn't attend, and some didn't provide answers (hence some inconsistency with the number of attendees at the top of this document). Responses in italics have been written according to Ismael's understanding of what was said during the event, rather than attendees themselves.

The ideas shared during part one can be roughly categorised into five themes of questions:

- **Creativity:** What creative methods are there to identify novel applications of AI tools? How can creativity – as a human capacity – help guide future AI research and development to meet real social needs?
- **Regulation:** What do AI tools mean for intellectual property and copyright? How can we bridge the gap between fast-paced technological advancements and much slower regulatory developments?
- **Education:** How can we best inform the general public about both the potential and limitations of AI technologies? What does tech education look like for children, workers of all ages, non-workers, and retirees?
- **Misinformation:** What are the societal impacts of convincing texts, images and sounds resulting from generative AI tools? How can we combat the spread of misinformation and the crumbling of our socio-epistemic institutions?
- **Economy:** How do we enable a more equitable distribution of the economic growth enabled by technological advancements? How can this counterbalance the current widespread violation of intellectual property rights, and ultimately inform better practices?

Part Two: Past and Future Steps

After a brief overview of past efforts relating with AI and led by the RSA, as well as the idea that funding, skills mapping and tangible outputs by December 2023 are to follow, attendees were asked three questions:

- What made past efforts succeed?
- What is exciting about the future steps?
- What foreseeable stumbling blocks are there?

Whilst setting December 2023 as a deadline for an initial output seemed reasonable, the concern was raised that funding efforts might best be conducted once a clear project had been identified. To this effect, it seems worthwhile identifying sources of funding and shaping projects

in tandem. **The goal here is to strike a balance between what funders are effectively happy to support, and what Fellows are happy to work on.**

The concern was also raised that the output would be disappointing if it turned out to be a "static report"; a document stored on a website and which was only occasionally referenced. The impact the group wants to make is far greater than the production of white papers. **Whilst documents *can* serve as clear and measurable outputs, we must work on promoting them and potentially developing and delivering services in light of them.** For example, interactive toolkits could serve our purpose better than white papers. Regardless of the format, the documents we produce must inform action as, only then, can they make an impact.

Part Three: Design for Life

This session was the most successful, and the task was to join one group within which to discuss potential next steps for the AI interest group with either the ethos of *Building Capabilities*, *Growing Hubs*, or *Developing Infrastructure* in mind.

Building Capabilities

Change starts with supporting people to learn and lead. At the heart of Design for Life is the commitment to build the skills, connections and confidence people need to fulfil their potential.

This group drew the most interest and was the largest. One idea that emerged was the potential to create learning modules *for RSA Fellows* who – as most of the public – generally lack knowledge about AI. However, these modules are not intended to be "static" but to inform further educational initiatives promoted through our vast Fellowship.

Growing Hubs

People effect change by coming together. Design for Life grows the connections that organisations, communities, industries and regions need in order to flourish.

The key notion this group shared related with the complexity and methodology for effecting change through effective community building. The theoretical notion of *impact networks* was described as one method for fostering cross-community communications and collaboration. As well as the idea of mapping the skills and interests held across the AI interest group, it was

suggested that we have a visual representation of the different communities we are all a part of. This can help discover overlaps and ways we can broaden our impact.

Developing Infrastructure

We are aiming for systemic, long-term change. Design for Life develops and influences the conditions needed to enable people, communities and organisations to regenerate our world.

This was the smallest group, and, whilst government were seen as the agent of change that the RSA should influence, two definitions of *infrastructure* were discussed: technological and regulatory. By technological infrastructure, the group discussed how government could make hardware and software available to the broader public so that the sort of compute power needed for AI developments not be concentrated in big tech, or not academic affiliation. By regulatory infrastructure, the group suggested the potential FRSAs have in informing policy, which often lacks in technical knowledge. Crossing the technological and regulatory definitions of *infrastructure*, the group identified the potential for public services to adequately incorporate AI systems – this requires providers to have clear understandings of the services to be provided, and that government agencies have capacity to scrutinise and select only the most effective tools.

Part Four: Resourcing Our Efforts

A presentation suggested three potential streams of funding to ensure that our work is adequately resourced, and makes the greatest impact:

- **Commercial partnerships:** By working with potential funders, we can identify efforts they are keen to support, like the Forum for Ethical AI achieved in 2017;
- **Employee benefits:** Fellow-led initiatives could be funded through a new programme of fundraising whereby employers donate to the RSA, and these donations are ring-fenced for such initiatives; and
- **Public grants:** Whilst we cannot rely solely on grant money, there are numerous opportunities already out there.

Each option came with caveats. *Commercial partnerships* raise the question of our *independence*, and the degree to which the funder sees themselves as shaping projects. The suggested employee benefit programme may not be the quickest to become a reality, as we learned that the RSA's current efforts are focussed on growing membership. Finally, public grants often require specific status – as a charitable, academic, or even for-profit entity. Whilst the goal is to engage as RSA Fellows and bring the RSA into the picture, two workarounds were suggested:

- FRSAs who have academic affiliations or the right types of entities in their networks can bring the needed legal status in to grant applications; and
- Consortia can be formed with other entities who qualify for funding – the key here is to demonstrate that the RSA and Fellows are open and effective in AI-related projects.

In trying to identify the Fellowship's potential for such consortia, it was suggested that a "citizens' jury" be formed by Fellows. We can then be seen as a go-to group of "lay" people who represent diverse views and can inform policy and developments concerning AI.

It was also discovered during this session that the RSA can provider certain infrastructure:

- **Circle:** As we know, this is a platform for sharing ideas and, as this event shows, is effective in bringing diverse Fellows together;
- **Zoom:** To facilitate an online version of this event, the RSA will provide a premium Zoom link to ensure all features are available during the call; and
- **Miro:** A collaborative and visually engaging whiteboard will be made available to this group to continue the conversation.

Reflections and Next Steps

The event was the first in-person, Fellow-led event for everybody in the room – RSA's Community Manager's efforts on Circle were key to making this possible. Whilst The Steps as a venue proved somewhat awkward, the group clearly shared a will to make the event a success. The shared drive and passion amongst FRSAs to promote a more harmonious society becomes almost tangible when brought together. Of course, this was only the first AI-focused, Fellow-led event, and its success is yet to be shown.

The group cannot become merely a place for exhilarating discussion, but a space that fosters new ideas and effective actions. Identified action items include:

- Creating a Miro board for the AI Interest Group
- Ascertaining a method for listing skills and interests amongst the group that can inspire projects
- Investigating how Miro can be used to map skills, but also other networks of which group members are part
- Considering coming to grips with the Levelling-Up and Regeneration Bill to produce a position statement
- Returning to your own three big areas of interest and reflecting on how you might change them in light of our event.

Appendix

Three big areas we should work on
Authenticity: *We must stand to the challenge of deepfakes and misinformation raised by generative AI tools*
Understanding the relationship between AI and IP/monetisation for creatives
How can AI help diversity of all types & inequality?
AI's place in Environment and Climate Change issues.
Ways for disabled and other people outside the mean to realise the potential of AI and equitably redistribute benefits - especially to enhance alternative models of learning, communication and expression.
Personal profiles that machine learn from you. Data sovereignty and reclaiming self-determination.
Open Source and transparent approaches to ensure access for all and not only determined by Big Tech and how they want you to use it. How we make the public and government literate to balance the power relations.
Bridges to the Future: How AI might be used as a bridge to Regenerative Futures & Life centred Design
Responsibility: How AI can be leveraged creatively for impact in a responsible, safe manner
Creatively for Good: How and where AI might be used, imaginatively, for social innovation and systems change
Future of work: Impact on employment and wealth concentration (e.g.: à la Luddite)
Lifelong learning: Employers' role in developing staff's AI knowledge
Education: Enriching the tech syllabus with critical thinking
Security implications: *There is concern for the impact of novel generative AI tools on how we navigate and understand the social world*
Regulatory gap: *How can we narrow the gap between increasingly fast technological advancements, and comparatively slow regulation efforts?*
Cultural capital - can AI provide opportunities for individuals to gain cultural capital thus enabling social mobility?

Three big areas we should work on
Lifelong learning - how can we develop a 'relationship' with AI so that it supports our lifelong learning journey (which includes our formal education, career/s, interests and 'retirement')?
Creativity - what is the potential for AI to help us access and nurture our creativity and develop our creative thinking e.g. finding solutions to combat climate change?
Data standards for purposeful uses. - Which sectors are the highest priority to adopt data standards. Are there data standards in place (e.g. RICS for surveying) or new ones are needed?
Progressing the Art & Science of AI that solves real world issues. Refocusing the public conversations on the opportunities to create value and save lives, not only recurring concerns and perspectives of ethics (especially ones that are independent of the tech, and mostly dependent on human uses).
Licensing and regulatory mechanisms for data used for AI training - Where the regulations cannot do anything about it?
Run an RSA-wide 'Citizen's Assembly' on how AI can best be used to regenerate society.
Partner with relevant third parties to develop measured and balanced content/campaigns to **improve AI education/literacy**
Primary education: *How early on can we begin instilling children with knowledge and skills related to AI?*
Waking up the people who have not realised how powerful AI could be as a force for good (and evil).
Enable culture change at charities who often perceive negative stories about AI tools but are oblivious to their potential for good.
Learning from/with children - the next generation already have it in their hands (snapchat AI) and will expect the world of work to have changed by the time they get there…less than five years.
The consequences of developing tech with no consideration of the impact on those it will affect
Is it too late for regulation?
As humans what life do we want? How much work do we want in our lives?

Three big areas we should work on
Create RSA **Foundational AI knowledge learning module** - MVP for Fellows Networks use - feedback - then formalise as RSA tool for wider use. Possible commercial product to RSA corporate commercial & NFP orgs usage.
AI Usage Framework Assessment Tool - MVP Fellows/RSA use in DfL initiatives. (will include best practice and compliance / regulation). Possible commercial product. May need Fellow partnering services to deliver follow-up
Economy & Business Network Tech4Good - Competition - What good use of AI looks like' - DfL initiatives with use of AI - challenge to Fellows & all Networks for project initiatives/ideas. Who tech assesses & votes on these? Sources of funding and partner tech co support & Fellows time / expertise.
Engaging "non-techies"- How do we demistify AI/ML for non-technical experts? How do we help people to see the opportunities rather than the threat, and tap into their ideas about how it can engage their endeavours
Alternative funding models-GenAI will dramatically reduce the costs of production, but mainly in a derivative way. How do we remunerate people who are feeding in novel ideas? Do we need overall productivity to increase, or do we need to redistribute value created and head to the beach?
IP-can the internet remain free? Do genAI models need to pay for access to high quality content, to avoid a cultural/knowledge cul-de-sac?