

# SPRITE+'s strategy and focal area definition 2024 – 2027

---

## 1 Introduction

This is a brief overview of the work done to define SPRITE+'s focus for the period 2024-27. Focus is different from our vision, principles and objectives (see appendix for reference) in that they will define the substantive topics which will form our research agenda for the next three years. This will affect the topics of the events that we run (e.g. innovation fora and sandpits), the engagement that we do and the projects that we fund.

Feeding into this process are our own internal discussions and deliberations and engagements with stakeholders and outcome of two key events: the futures summer camp held in July 2023 (and subsequent collaborative report) and the Expert fellows meeting in May 2024 (and subsequent processing of outputs).

These two events are described briefly in section 2.

## 2 SPRITE+ events

### 2.1 The Futures summer camp

The 2023 Futures Summer Camp was funded by SPRITE+ and organised jointly with the Research Institute for Sociotechnical Cyber Security (RISCS) and facilitated by the School of International Futures (SOIF). We invited 28 experts from academia, government, and industry to explore what the future holds for the complicated intersection of technology and democracy.

The multi-disciplinary group was asked to consider:

- the threats and challenges to democratic institutions that will be faced by society in 5, 10 or 15 years,
- how the collision of technology and democracy might lead to positive impacts, and
- what actions we should take now to mitigate negative outcomes and maximise positive ones.

Across two days, workshop participants developed five distinctive themes at the intersection of technology and democracy that might characterise a near-to-midterm future (taking a five-to-fifteen-year perspective). The themes that emerged are summarised in the report on the summer camp but briefly they were:

1. **An Inauthentic World:** A future of increasingly perfect 'fakes' and anonymous activity, in which nobody is sure what is real and who they are interacting with—whether human or machine.

2. **Swarms to Flocks:** Digital technologies have enabled humans and machines to ‘swarm’ in novel ways, resulting in unprecedented behaviours and opportunities for group formation and action.
3. **Simulating the Future:** Ever more sophisticated simulations of the future come to underpin political decision-making, with democratic institutions taking new forms as a result.
4. **Beyond the Nation State:** The failure of traditional democratic institutions to respond to new technologies, and the democratisation of those technologies, pushes people to take matters into their own hands.
5. **New Models of Dissent:** A future in which technology has enabled and necessitated the emergence of new forms of dissent, while rendering some existing forms ineffective or obsolete.

The research questions that emerged from these discussions are summarised in Appendix B. They cover these areas:

- Inauthenticity
- Privacy, agency and safety
- Groups and communities
- New ways of sustaining democracy
- Strengthening critical responses to technology

## 2.2 The spring 2024 Expert fellows meeting.

The 2024 expert fellows meeting was held in London at Scale Space on the 30<sup>th</sup> April and 1<sup>st</sup> May. The meeting was framed around four provocation talks by experts in an emerging technology:

- 1) Cognitive Robotics
- 2) Neurotechnology
- 3) Biotechnology
- 4) Digital twins

The choice of the four topics arose from various discussions arising during our engagements with stakeholders over the previous six months. We could undoubtedly have chosen different topics quantum computing is one obvious choice– and the absence of AI is noteworthy. On the later this has been excluded for two reasons. Firstly, it is already the focus of widespread work and therefore it was not clear that SPRITE+ would be able to add significant value. Secondly AI is cross cutting. It is critical component of all four of the technologies on our lists and it therefore will feature heavily in any case, and it is more in keeping with the 5 to 15-year horizon that we are aiming to cover.

Each of the talks was followed by deliberative breakout groups in which the fellows discussed the relationship between the emerging technology and the topics of interest to SPRITE+ i.e. Trust, Identity Privacy and Security.

During the deliberative sessions, the participants were encouraged to write thoughts on post it notes and the post these on the wall. 225 were posted by the end of the event. These were gathered by the organisers and transcribed and then uploaded to the platform well sorted.

Nine of the participants volunteered to take part in the well sorted exercise, which involves each person sorting the posts into clusters.

Before the exercise SPRITE's director went through the posts and removed 35 of them which were illegible, too general or obscure to meaningful or duplicates which left us 190. Each of the group completed the task in about 1 hour.

Having completed the sorting process, the well sorted platform then generates clusters using a hierarchical clustering algorithm. We selected the maximum of ten clusters (on account of the large number of posts).

The ten clusters that emerged could be characterised as follows (with each followed by example posts).

- 1) Trust and trustworthiness
  - a. *Trust - how is it created and lost? Unplanned inferences and Personal Data*
  - b. *Human trust is reciprocal and mutual - when/how will this be possible with AI*
- 2) Data
  - a. *Datafication (Biotech/transparency/identity/trust/safety/privacy/accuracy/fairness)*
  - b. *Incentives for data transparency need to be introduced.*
- 3) Legislation, governance, policy
  - a. *Regulation is lagging behind.*
  - b. *Accountability - Who's responsible if things go wrong?*
- 4) Human-Robot interactions
  - a. *If we are relying on robots, what will this mean for future human interactions?*
  - b. *Do human-like robots make users more or less privacy aware?*
- 5) Human Augmentation
  - a. *How will people manage their own neurosecurity?*
  - b. *Biohacking and the boundaries of human identity*
- 6) Errors and learning
  - a. *Learning is also about making mistakes so should we tolerate it?*
  - b. *Social learning - learning from errors in different contexts*
- 7) Digital twins
  - a. *Is a twin a replacement for data? If so, data privacy needs to change.*
  - b. *Threat models and digital twins*
- 8) Ontology and meaning
  - a. *Are we constrained by our vocabulary and how it relates to concepts we're trying to discuss?*
  - b. *A shared vocabulary? With definitions?*
- 9) Something of a bucket category of threat frameworks with half the posts focused on AI.
  - a. *What are the security implications? - with gen AI - content sharing by children and others?*
  - b. *Critical infra-structures. New attack surface.*
- 10) A single post: Being human is not a protected characteristic.

See the well sorted report for full details. Categories 4 and 7 and to a lesser extent 5 were in a sense unsurprising as they were driven by the choice of provocation talks. However, even for these the overall themes of the posts did not just echo back the topic but rather nuanced. So, in cluster 4 the focus was on human robot interactions rather than robots per se and for category 5 the algorithm had grouped together Biotech and Neurotech and the focus of the post was on various facets of human augmentation. With digital twin's cluster this was perhaps the most vanilla.

During the final part of the expert fellows meeting the participants engaged in group deliberation this time to answer five more operational questions:

- 1) Given all that we have discussed have any ideas for research foci emerged? (e.g. deep dive projects, sandpit topics, innovation fora etc)
- 2) In the light of these future tech development are there any organisations that SPRITE should be engaging with?
- 3) What are they key areas where policy needs to be developed?
- 4) Are there any areas of focus where SPRITE+ might produce a white paper (or several)?
- 5) Are there any other areas that might be important for us to focus on?

The full set of the responses can be found in appendix C.

### 3 Observations

The combination of the two meetings and follow up exercises outlined in section 2 suggest that there may be two broad threads that have emerged as candidate organising themes for SPRITE+.

- 1) Human technology interaction.
- 2) Cyber-physical systems and infrastructure.

Crosscutting those two themes are a set of interrelated lenses:

- GERP: Governance, ethics regulation and policy
- Language, meaning, education and communication.
- Trust and trustworthiness.
- (new) Threats and risks

The responses/outputs from both the fellows and the futures summer camp largely fall within one of those categories/lenses.

However, these are a little too granular to drive a short term (three year) strategy and so we must dig a little deeper.

A particularly strong signal seems to have come through that a significant gap exists in the issues arising from Neurotechnology. Neurotech maps onto both the two broad themes above and throws up issues in all or the cross-cutting lenses. We note that there are existing network pluses and other UKRI investments working on neurotech, but these have tended to be on particular application areas rather than the externalities of the tech (including TIPS).

A second potential focal area is an amalgam of simulation and digital twins. This is less well defined because the tech is embryonic.



Security, Privacy, Identity, Trust,  
Engagement, NetworkPlus

Another area which has emerged from our engagement function is the intersection of resilience and security. This ties directly into a priority UKRI programme *Building a secure and resilient world*. – the SPRITE director is also a co-I of SALIENT, UKRI’s Research and Coordination Hub for resilience research and so has taken this opportunity to define with SALIENT’s director a joined-up piece of work on the topic, so this emerges as a third area.

These three then will be added to the three live areas that have formed part of the fifth years developmental work: Trustworthy Digital identities, Digital Cities and Connected places and Living in an inauthentic world.

A final area which has emerged less strongly in the discussion, but which was deliberately de-emphasised (to stop it overwhelming the agenda with immediate concerns) is human-AI interaction. We are now over a year into the latest breaking wave of AI and its impacts are starting to crystallise. It therefore seems appropriate to take stock and consider how AI development will intersect with all these other focal areas.

In sum then our focus for 2024-27 is the six areas:

- 1) Living in an inauthentic world.
- 2) Trustworthy Digital identities.
- 3) Digital Cities and Connected places.
- 4) Resilience and TIPS.
- 5) Neurotechnology and its impact on TIPS.
- 6) Digital Twins and Simulative technologies.

For the first three we are not anticipating any further immediate development work as the projects that are associated with these areas are live. We may conclude downstream that further work may be necessary, but this is likely to be in the form of follow-on grant applications rather than work going on within the (current) SPRITE envelope. The resilience theme is in development now and will almost certainly be the theme for the sandpit in 2025 followed by a secondary fund resourced by SALIENT

For Neurotech and digital twins the tentative pathway is as follows:

- 1) Formation of a SLWG to identify indicative lines of enquiry and key stakeholders.
- 2) An innovation forum (bringing in some of the key stakeholders) to map out the topic area and possible research questions and perhaps producing a white paper.
- 3) Either a one-year deep dive project funded via a call to the expert fellows and/or the 2026 sandpit topic.

We should aim for step 1 to be complete by February 2025 and the innovation forum to be held in the Summer of 2025 with the call going for a deep dive project going out in September with successful projects to start January 2026.

This leaves us with the following proposed schedule of substantial projects.

Topic	SLWG	Innovation Fora	Sandpit	Deep Dive project
Living in an inauthentic world.		2024		



Security, Privacy, Identity, Trust,  
Engagement, NetworkPlus

Trustworthy Digital identities.				2024-2026
Digital Cities and Connected places.				2024-2026
Resilience and TIPS.			2025	
Neurotechnology and its impact on TIPS.	2024-25	2025	2025-2027	
Digital Twins and Simulative technologies.	2024-25	2025	2025-2027	

It is worth noting that all of these areas are intersectional. Neurotechnology would be supported and accelerated by human Digital twins and vice versa. Trustworthy digital identity necessary underpin ethically grounded connected places and so on. In terms of Sprite+'s principle of *making magic in the joins*, we can and should be looking out for these emerging intersections as well as considering each area on its own.



Security, Privacy, Identity, Trust,  
Engagement, NetworkPlus

## Appendix A

### SPRITE+'s Mission and Vision

SPRITE+ is a vehicle for communication, engagement, and collaboration for people involved in research, practice, and policy relevant to TIPS in digital contexts. Since launching in 2019, we have established ourselves as the go-to point of contact to engage with the broadest UK network of interdisciplinary, cross-sector digital TIPS experts and vehicle for the development of emerging TIPS talent.

Core to our mission and vision is engagement both within and beyond the TIPS research community, and with all sectors of society. In this second phase, we aim to deliver a set of activities that will continue to build a sustainable network of engaged stakeholders to match the ambitions, scope and scale of SPRITE+. By leveraging our existing networks, reaching out to new stakeholders in new ways and expanding our use of digital channels, we will increase awareness of opportunities to engage with us that is inclusive and accessible to all.

Furthermore, our activities will align with the EPSRC's strategy and priorities, and its commitment to build local, regional and national partnerships *built on trust and openness*.

### SPRITE+'s Principles

SPRITE+ is founded on four basic principles with engagement sitting at the heart of each one.

**Interconnection: Making magic in the joins.** Digital TIPS challenges and opportunities are best addressed with an interdisciplinary 'quadruple helix' model of academic-government-industry-citizen engagement.

**Diversity: Great innovation from diverse voices.** Our commitment to equality, diversity, and inclusion (EDI) is woven through all activities and we take practical steps to encourage and facilitate participation.

**Challenge: Encouraging challenges to traditional perspectives.** We embrace 'traditional' cybersecurity perspectives, but we actively seek to broaden and challenge our understandings of digital TIPS for the benefit of society.

**Collaboration: Support; Communicate; Network.** We work closely with cognate organisations and networks with an ethos of collaboration, and seeking opportunities for leverage and support, not duplication.

## Appendix B: Research questions generated at the futures summer camp

### Inauthenticity

Participants asked how societies could manage growing inauthenticity, whether collaboration (or even democracy) can still happen effectively in the context of it, and whether inauthenticity could, in some contexts, be beneficial. Specifically, participants suggested exploring:

- The impact of the accelerating democratisation of ‘tradedcraft’, making it easier for people to hide their identities and operate anonymously.
- The mechanisms that might facilitate collaboration between humans and machines, when it is difficult or impossible to ascertain what, if anything, is authentic.
- The strategies that citizens might be able to use to authenticate things they see, hear, read, watch or otherwise interact with in a future world of perfect fakes. (This is critical to being able to make decisions in a democratic society).
- The contexts in which inauthenticity might be beneficial for democracy.

There was some related discussion about what it means to have more done by machines, with a question about the contexts in which *not* having a human involved in an action or decision might lead to positive outcomes, and another around whether it’s possible to do innovation without human communities.

### Privacy, agency, and safety

While on the one hand technology might facilitate greater anonymity, several groups had also touched on the opposite problem, with new challenges to privacy, agency and safety in technologically enabled democratic societies. There appeared to be concern about how at-risk individuals and communities might be able to avoid harm, or even opt out of some future technological developments:

- How people who might need to ‘hide’ (such as victims of domestic abuse) can be anonymous and safe in a world of voluminous corporate and open data.
- Whether, and how, people who so chose might be able to live in future society without technology platforms or digitally based services.
- How, or if, the nation state might be able to protect its citizens in an age of trans-national, virtual communities.

These issues were touched on in other areas; for instance, underpinning work that might be done to apply technology in a beneficial way to democracy (more on this below). One of the main concerns around creating credible simulations of the future to guide decision making was the need to do so at the same time as preserving privacy and agency, for example.

#### 1. Groups and communities

Another set of questions centred on the emergence and development of groups and communities in the context of digital technology, and on how these groups might relate to each other and to democratic institutions. Participants expressed interest in working on:

- How technology might interface swarm or flock communities with democratic governments.
- How to maintain productive tensions in and between groups, challenging institutions without leaning towards polarisation.
- How technology-enabled social groups of democratic process can get along.
- How, and whether, multiple swarms or flocks of dissenting humans and machines can co-exist and achieve productive outcomes.

More generally, the discussion touched on questions relating to how technology might be used to strengthen democracy in the round, and on how democratic societies can improve their responses to new technologies.

## New ways of sustaining democracy

Participants felt there were opportunities to use technology to sustain democracy and make it more robust, but there was little clarity on what, precisely, these opportunities were.

Accordingly, there was interest in exploring:

- The ways of maintaining and sustaining democracy in digital, technological societies. Addressing dependence on systems such as PNT [Positional and Navigational Technologies, that underpin GPS] was part of this— participants were uncertain whether society as it works at present could function effectively in a future in which PNT is untrustworthy.
- How digital technologies might be used to help surface the values of democratic societies.
- How (digital) anonymity might enhance judicial and democratic processes.
- How to create credible simulations of the future to improve democratic decision-making (while preserving privacy and agency).

## Strengthening critical responses to technology

Underpinning much of the discussion was a sense that democratic societies and institutions are poorly equipped to manage the emergence of new technologies, resulting in lots of the problems we see today at the intersection of the two. Further work was proposed on:

- What 'critical technology education' might look like, for policy makers and for citizens.
- How far the electoral cycle hinders the possibility to make long term plans to address technological change, and how this might be addressed.
- How new social architectures and incentive structures might be engineered so that technologies are developed not just for profit, but instead in response to social challenges with the involvement of the wider community ('no innovation without representation').

There was also an eagerness to re-emphasise the sociotechnical and find ways of doing this effectively: for instance, by designing new social architectures around technological challenges that are not solely business-centric ('no innovation without representation') and finding ways of



Security, Privacy, Identity, Trust,  
Engagement, NetworkPlus

motivating businesses to explore technology development for reasons other than profit, leading to more equitable outcomes.

## Appendix C: Responses to the concluding exercise at the expert fellows meeting.

- 1) **Given all that we have discussed have any ideas for research foci emerged? (e.g. deep dive projects, sandpit topics, innovation fora etc)**
  - a. Impact on humanity/future societies
  - b. Establishing responsible uses
  - c. Promoting proactive legislation
  - d. Enhancing educational needs
  - e. AI and security
  - f. Gen AI and Cybercrime
  - g. Regulations
  - h. Definitions for terms
  - i. Data ownership and consent to use: Regulation, Policy and Law
    - i. Look at Technologies, Legal, social elements.
    - ii. A set of people convened by SPRITE would then deliver a bid for funding that would create a set of policy papers/research papers cohesive across areas.
  - j. Interoperable digital identities in a cloud AI ecosystem:
    - i. A full tech/legal/social analysis of options, investigating bottom-up identity solutions.
  - k. Citizen data accounts:
    - i. how this would change trust, cement privacy, optimise security and enable identity.
  - l. Language for this collective discipline
  - m. Trust always comes after
  - n. Interdisciplinarity in trust
  - o. Digital inclusion and increased complexity towards security
  - p. Build a library of EmTech use cases (and scenarios)
    - i. that can be used to provoke/interrogate implications (for policy research etc)
    - ii. to examine from multiple perspectives: TIPS, Disciplines, Publics
  - q. Trust particularly in human tech pairings (AI/robot/neuro)
  - r. How much is our tech pairing part of identity?
    - i. Which bit is the person?
  - s. International collaboration on norm setting and standards.
  - t. Deceptive options in digital twins for security/defence
  - u. Sandpit around human machine pairing with tech village.
  - v. What is the future of Money?
  - w. Neurodata/Biodata forensics.
  - x. Cybersecurity of sociotechnical systems of socio-technical systems
  - y. Something around new threats where software controls the physical world.
  - z. Cybersecurity of neurodata and engineering biology

**2) In the light of these future tech development are there any organisations that SPRITE should be engaging with?**

- a. Cabinet office
- b. Central Digital and data office
- c. UKRI
- d. DSIT
- e. Industry stakeholder
- f. Charities
- g. DARE
- h. International models – Singapore
- i. AI Veirfy Foundation
- j. NCA
- k. META
- l. Google
- m. X
- n. Tiktok
- o. Amazon
- p. Digital Regulation Co-operation forum
- q. Institute of Science and Technology
- r. Is there a European SPRITE+ equivalent
- s. Digital Futures at Bristol
- t. Regulatory horizon council
- u. Which? Consumer protection
- v. NICE
- w. MHRA
- x. Alan Turing Institute
- y. Arts organisation
- z. BBC ( I would like to suggest a documentary)
- aa. INCOSE
- bb. Royal Academy of engineering
- cc. ENISA

**3) What are they key areas where policy needs to be developed?**

- a. Assessing current regs/laws for compatibility
- b. International collaboration is regulation/policy setting
- c. Anticipatory Governance initiative to identify new policy needs
- d. Establishing independent oversight
- e. AI
- f. Deepfake – online safety
- g. Neuro act
- h. Tech
- i. See 1 i- l
- j. Neuro-regulation
- k. Biosecurity
- l. How are policies communicated to the user? Accessibility
- m. Neurotech and biotech
- n. Identification of strategic/restricted technology of future

- o. Digital twins for CNI and their vulnerabilities
- p. Requirements for deep fakes to preserve uncanny valley.
- q. Expansion of GDPR for areas omitted
- r. AI regulation that balances innovation
- s. Neurotech

**4) Are there any areas of focus where SPRITE+ might produce a white paper (or several)?**

- a. Intersectionality between the tech instead of siloed thinking.
- b. Macro overview
- c. Establishing key stakeholders (experts and public)
- d. How to promote social inclusion
- e. Threat modelling of Gen AI
- f. See 1 | ->|
- g. Neuroregulation
- h. Definition of digital twins
- i. Meaning and the main issues around deep fakes
- j. Neurotech its privacy and policy implications
- k. Misuse case and a risk oriented red team framework (per tech after framework)
- l. Using humans as components in solution architecture
- m. Threat Landscape of emerging tech (discussed in these two days)

**5) Are there any other areas that might be important for us to focus on?**

- a. Establishing who should be overseeing the technologies.
- b. Public engagement/festival/open days
- c. Blockchain
- d. Critical Infrastructure
- e. Cyber warfare
- f. Data Exploitation
- g. Power inequalities in the digital economy
- h. Novel attacks
- i. Assisting early career researchers to access grants.
- j. Development of training/outreach
- k. SPRITE+ YouTube channel
- l. Metaverse?
- m. Filters as social mediators (harms and offenses)
- n. Time for tech luddites to preserve resilience Trust but verify.
- o. Not quantum!
- p. Consider exclusion/polarisation as a threat.