



# AI Openness Update From Agentic to Public Good in 2025



## Index

<b>1. Overview</b>	<b>3</b>
1.1 Executive Summary: TLDR	3
1.2 Introduction: AI Openness and Public Good in 2025	4
<b>2. Global AI Law and Policy Update</b>	<b>7</b>
2.1 Fireside Chat: Global AI Policy in 2025	7
2.2 2025 Policy Landscape	10
2.2.1 Open Weights Definition from the Open Weights Alliance	10
2.2.2 Global AI Summits	11
2.2.3 Thought Leadership: Current AI, A global partnership for open and Public Interest AI	12
2.2.4 UK Policy Roundup	14
<b>3. The 2025 Data to 31 May</b>	<b>16</b>
3.1 The UK Landscape	16
3.1.1 UK Adoption of AI openness	16
3.1.2 LinkedIn Data	16
3.1.3 UK Public Sector Data	18
3.1.4. Download Data	18
3.1.5 Repository Data from GitHub	19
3.2 Global Landscape	20
3.2.1 Hugging Face Data	21
3.2.2 Global Rankings by Country	22
3.2.3 Global Growth	23
3.3 Europe - the EU and UK	24
3.3.1 AI Repositories in Europe	24
3.4 Agentic AI Data	25
3.4.1 The global Agentic Landscape	25
3.5 Cybersecurity Data	26
<b>4. Agentic AI</b>	<b>27</b>
4.1 Thought Leadership: Agentic AI	28
4.2 DSIT, Identity, and AI Agent Security	30
4.3. Thought Leadership: Identity	34
4.4 Thought Leadership: Securing Agentic AI	36
4.5 Thought Leadership: Quantum Readiness, Securing AI and Public Trust	38
<b>5. AI Global Literature Review</b>	<b>40</b>
5.1 Key themes from the 2025 reporting landscape	40
5.2 The UK's AI Landscape	40
5.3 Trends Shaping the AI Landscape	41
5.4 AI Adoption & Maturity	41
5.5 The Public Sector	42
5.6 Generative and Agentic AI	42
5.7 Security	43
5.8 Scaling Challenges	43
5.9 The Rise of AI Openness and Public Good AI	43
5.10 Challenges of AI Openness and its commercial adoption	44
5.11 Future Outlook for Open Source AI	44
<b>6. Conclusion</b>	<b>45</b>
<b>7. Formalities</b>	<b>47</b>
7.1 Contributors	47
7.2 About the Creators of this Report	49
7.3 Acknowledgements	50
7.4 References	50
7.5 Sponsors	52



## 1. Overview

### 1.1 Executive Summary: TLDR

This report updates OpenUK's last AI Update Report as at December 2024 and focuses on the current challenges in AI Openness, which is increasingly being referred to as Digital Public Interest or Public Good AI.

The Report is broken into 4 sections as well as an introduction:

**(i) Policy update:** which includes a hard hitting Fireside Chat with Harvard Fellow, Ben Brooks who explains the removal of the US Presidential Executive Order and the impact of almost 1000 Bills in the US, and his view that the EU may well not enforce the AI Act despite current work on a Code of Conduct. In the UK, the AI Bill announced a year ago has yet to be tabled while the Copyright consultation has caused friction between the creative sectors and tech.

The policy update also looks to the AI Summits instigated by the UK in 2023, through the 2025 Paris AI Action Summit and Indian Summit likely to take place in early 2026. The French Summit saw the launch of 2 important Public Good or Public Interest initiatives, ROOST and Current AI, with the Founder of Current AI, Martin Tisne sharing an overview.

**(ii) Data Analysis:** The data includes repository information but also other sources, to show the global European and UK positioning. We see France as the fastest growing in Europe, while the UK remains number one by number of repositories. India has grown by 38% whilst the US retains its number one global position with a whopping 1,080 repositories exceeding 1,000 stars.

Skills data from LinkedIn shows that more is needed in the UK and the Prime Minister has responded to this gap with a skills program and £1bn of funding for compute.

**(iii) Literature Review:** As the depth of reporting increases, OpenUK has analysed 30 reports in a literature review. Key conclusions are set out before the detailed analysis at 5.1 and include: The UK leads in AI market size but faces structural barriers, Generative AI (GAI) is democratising access and use of AI, Agentic AI is the next wave of transformation, open source AI is now a strategic priority, AI infrastructure and governance remain bottlenecks to scaling and progression, security, trust and ethics remain major concerns and, finally, a national focus on AI openness could be the UK's leading edge.

**(iv) Agentic AI Thought Leadership:** Toran Bruce Richards founder of AutoGPT one of the most influential general purpose agents in the world which is open (and which sits under the Lovable tool) explains how restricting agents' access, which many are calling for from a security perspective removes the fundamental value that the agents bring. Sal Kimmich explores identity as a solution and Matt Barker builds on this whilst Isabelle Mouray and Richard Steel explore Quantum. Andrew Martin of ControlPlane shares the outputs of the recent DSIT hackathon with the Hyperscalers

**(v) Conclusion:** This 2025 AI Update Report reflects a fast-evolving global landscape in which AI openness is no longer a fringe topic, but a strategic, economic, and ethical imperative - increasingly framed as a digital public good. Against a backdrop of geopolitical divergence, regulatory unknowns, and intensifying innovation, the UK remains well-positioned in global rankings but faces structural hurdles in skills, infrastructure, and scale. Crucially, the literature and data converge on one point: openness - in systems, governance, and collaboration - could be the UK's defining competitive edge. With leading voices in agentic AI, active public sector experimentation, and growing engagement in open source ecosystems, the UK has a unique opportunity to lead the global conversation on responsible, transparent, and public-interest-driven AI.

This report not only updates the data but reframes the narrative - from competition to stewardship - urging bold actions, inclusive investment, and renewed public trust.

## 1.2 Introduction: AI Openness and Public Good in 2025

Amanda Brock,  
CEO, OpenUK



### Shifts in the landscape and language of AI Openness

2025 has seen a number of shifts in the landscape, in particular a change in the language from open source AI to AI Openness and Public Good or Public Interest AI. In that light, we have changed our report title.

2025 began with a Chinese AI New Year celebration and the “R1” model, trained via large-scale reinforcement learning (RL) from China’s DeepSeek just in time for China’s Year. R1’s weights are distributed on the open source MIT licence and it is best described as an “open model”. A reasoning model, R1 was apparently built at a cost of circa \$5million in comparison to past LLMs’ \$100m price tags. This was done by “distilling” other LLMs including Qwen and Meta’s open Llama LLM. [DeepSeek’s repository states](#) that they “demonstrate that the reasoning patterns of larger models can be distilled into smaller models” in R1. It has fared very well in industry tests against larger models and closed models. Nvidia, which acknowledged the importance of open models in its development at its first GTC in Paris this year has also indicated that [Small Language Models](#) will be the future of AI.

The value of openness in AI, as with past open innovation, is a combination of iterative development; democratisation through access for all; and trust through transparency.

In terms of iterative development, each innovator builds “on the shoulders of giants”, recycling and reusing that innovation, then sharing their outputs as open innovation enabling the next generation to iterate on top. This iterative pattern of innovations offered by openness is at the heart of the value of open source software and how open source software solved past digital challenges. AI that is open will follow this pattern.

The release of R1 led Sam Altman to share his personal view that OpenAI is on the “wrong side of history” when it comes to opening its models. OpenAI has also committed to releasing an open language model.

### R1 and openness

The reduced cost of creating R1 made model creation more accessible, but its diminutive stature and the potential for reduced compute also sent shock waves across the stock market and in particular, Nvidia’s stock price. The stock price quickly recovered, with many believing the application of [Parkinson’s law](#) is relevant to compute. Even as we reduce the compute required by models, so too we will increase the functions to which they apply and our use is unlikely to diminish.

Like gas in a vessel, our growing need for GPUS will likely fill the capacity smaller models create. There are however legitimate and valid questions over data centre need. Are we living in a data centre bubble? Will the huge current demand on the compute infrastructure, processing and memory in data centres potentially shift over time? It is clear that a critical inhibitor to AI evolution today remains data centre capacity, access to power and compute.

How long the challenge will continue at the current scale remains to be seen. As technology evolves the need for compute and memory to execute specific functions may decrease, but so too the growth of our usage of the technology may broaden.

There were of course other questions raised by R1 and the principal one was geopolitics.



## Geopolitics

The geopolitical alarm bells rang with the release of R1, both because it was created in China and its being an open model. These alarm bells rang loudest in the US but could also be heard across a number of countries, with countries such as Italy flagging data privacy concerns.

Geopolitics has been top of the policy and digital agenda in 2025. As a consequence of China producing a viable market challenger in its open reasoning model R1- despite 5 years of open development which could easily have led any careful observer to the view that this was an inevitability - its use of US-shared AI to build upon, caused concerns in some quarters. Whilst the President caused concerns in other parts of the world which have led to increased demands for digital and AI sovereignty the President too has been concerned by geopolitics and focused on threats to the US's world leading AI position - across the "entire stack".

We also see geopolitics having a profound impact on policy and regulation of AI in 2025, adeptly explained in a Fireside Chat with Harvard Fellow, Ben Brooks.

## Reluctance in adoption of AI that is open

With an in-depth analysis of 30 reports as part of our literature review, Dr Barth's findings include a reluctance in adoption of open models. I wonder whether this reluctance in commercial adoption of AI that is open is reminiscent of the risk aversion and reluctance in adoption of open source software in its early days?

The risk averse gatekeepers - legal, finance and procurement - may be delaying successful adoption of AI that is open today as they did the adoption of open source software 15 years ago. Is this risk aversion stifling innovation and the use of innovation in enterprise as it did the growth of open source software in the UK in the past?

I'd suggest that those early adopters and nations which enable their ecosystems of AI openness, like those who developed leading open source software in the recent past, will be the winners in the AI race in the long term. The UK must learn from this recent history in software in both its own shores and globally. The more successful nations have seen greater adoption and proliferation of open source in their software ecosystems.

To be successful today the UK must engage more with its open source software ecosystem, its open data communities, and the growing ecosystem around AI openness. By taking appropriate actions it will underpin future success in AI through openness.

## Global AI Action Summits

The UK instigated global summits which saw a shift in 2025 as the French summit moved from safety to action and placed openness at the centre of the discussion.

Both China and the US delivered the same clear message to summit attendees - that they would not be held back in their innovation by other countries' legislation.

JD Vance delivered a speech clarifying that the US felt nothing good came out of collaboration with certain countries and that the US would not tolerate behaviour from countries which sought to diminish US tech companies. The lines of a geopolitical battle around sovereignty were clearly drawn.

India co-hosted the Paris Summit and will host the next global AI summit, likely in early 2026. Prime Minister Modi made clear in his opening address in Paris that "access for all" through "open source" will be critical. We can expect to see the public good and public interest AI conversation instigated by President Macron strengthen through the remainder of 2025 and into the 2026 summit, and outputs of the French summit like ROOST's open source tools and the work of the Current AI foundation growing.

### **The UK's position**

Using repository data from Hugging Face for the 1st time, along with GitHub's repositories, the UK is seen as globally 3rd in AI that is open. This is a 19% increase from 2024. But that growth, while respectable, is being outpaced. India grew 38% in the same period, France by 26% and the US by 23%. We also analysed LinkedIn and Scarf data to broaden our research into the data beyond the repository. We will continue to evolve this approach in further reporting, enabling broader topics to be covered.

### **The Year of the Agent**

2024 will be noted in history as the year of the LLM but 2025 is shaping up to be the year of agentic AI. Scotland's AutoGPT still reigns supreme in agents that are open. But the data leaves open questions. OpenUK has been tracking agents and agentic AI and with the help of GitHub and Runa Capital we delve further into the repository data of the agentic landscape. Our Thought Leadership extracts cover the debate around agentic security versus opportunity, in cutting edge pieces from AutoGPT's Toran Bruce Richards, Sal Kim-mich, Andrew Martin, Matt Barker, Richard Steel and Isabelle Mauny. This conversation is globally leading on agentic trust and quantum readiness.

### **Conclusion**

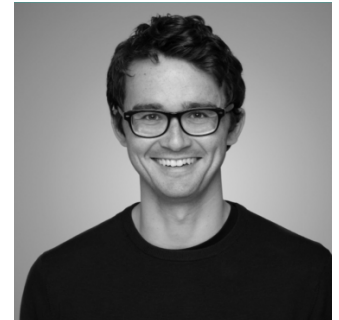
2025 has been defined as the year of the small model and agentic AI. But it is also notable as the year when the AI openness conversation has exploded onto the global stage. We now need to see global and UK trust in that openness evolve.

Whether the UK now takes the lifeline in its AI policy is a question for our government and policy makers. Our growth is falling behind other countries, but our communities of practice are strong and with that support translated from policy into practical actions the UK's open communities are ready to support its growth and leadership in AI and in particular AI openness.

## 2. Global AI Law and Policy Update

### 2.1 Fireside Chat: Global AI Policy in 2025

**Ben Brooks**  
Fellow,  
Berkman Klein Center,  
Harvard University



#### **Your current role in Harvard**

I'm a fellow at the Berkman Klein Center, Harvard where I focus on the regulatory and legislative response to models and specifically open weight models. I help Federal and State legislators in the US and elsewhere to help them think about the future of model layer regulation.

We're in the middle of a legislative tsunami. There are over a thousand state and federal bills in the US. The EU is undecided about how, if or when it will implement the EU AI Act and the UK has been flip-flopping between different regulatory proposals. A lot of my work whilst at Harvard, has been engaging with policy makers, and helping them to understand how to think about problems in the context of a very complex AI supply chain.

#### **What's happening in the US?**

The Trump administration is a broad church with a diverse set of voices. You've some who come out of the "effective accelerationist" world and want to remove barriers to development, move fast and secure US leadership in AI. And then there are others who are deep within the national security community and who still share many of the concerns and questions that the Biden administration had about frontier models. My view is the administration is still deciding how it's going to regulate and respond to capable frontier models over the coming months and years.

In the first few days of the Trump Administration, they repealed the Biden Executive Order on AI. It was the longest Executive Order in US history. The Trump Administration launched a consultation on their AI strategy and then more recently they've decided to rescind the "AI diffusion rule" which had imposed export controls on closed source frontier models. It's not entirely clear currently what they might replace these instruments with.

They've also just announced that the US AI Safety Institute will be rebranded the Center for AI Standards and Innovation. It's yet to be seen what the long-term strategy looks like, but certainly they've emphasised deregulation, rapid development and preserving US leadership in every part of the tech stack.

#### **Will there be a Trump Executive Order as a consequence of the consultation?**

I think everyone's expecting further Executive Orders on AI.

These will likely have a focus on compute and power generation for data centres and removing red tape barriers to infrastructure development. They are unlikely to have a regulatory flavour in the way that some of the Biden-era instruments did. But, it could go in a number of directions. We know that there are voices within the Republican party who favour for example more intensive export controls, not just on chips but on intangible technology like models, weights, data and research.

#### **The split between Federal and State law?**

In the past 5 or 6 months, we've seen over a thousand state bills touching on AI, which is more than the total number of State Bills on AI for the preceding 2 years. There's over 120 federal bills on AI and a flurry of Executive Orders introducing policy and strategy or repealing Biden-era instruments. In the US certainly there's been a tidal wave of legislative activity.

I think people tend to get hung up on the volume of that activity. If you actually dig into it, a significant proportion of those bills are not regulatory. They are appropriations bills, they're bills setting up studies and task forces. A smaller fraction of them have a regulatory effect, and of those bills, hundreds focus specifically on deep fakes and hundreds focus specifically on high-risk AI systems – conducting or automating deci-



sion-making in certain critical domains. Very few of them go to the question of model development specifically and certainly very few of them directly touch on openness.

### **California Senate Bill 1047**

Senate Bill (SB) 1047, which was vetoed by California's governor, and new variations of that Bill that we're seeing this year would introduce or codify developer duties for the developers of powerful AI models. The challenge is that these Bills would require all developers to mitigate certain risks to some level. The standard for mitigation is vague or undefined and that Big Tech companies with closed source products can absorb a lot of that litigation and regulatory risk. It becomes much more difficult for small developers who choose to release open models where they don't necessarily have visibility and control over the downstream deployment and application of those models.

I've always been sceptical of that approach to regulating AI. But as I say, we continue to see some of those ideas percolating through post-election into this year.

### **National Institute of Standards in Technology (NIST)**

**In August 2024 we saw NIST saying that open models were okay for now at least. Is that still in place or has that also been moved along?**

NIST under the Biden Administration was unambiguous that at this point in time they did not support further intervention to inhibit or restrict the release of open models. The challenge is that with the release of DeepSeek, the US establishment shifted from a conversation about catastrophic risk to a conversation about China risk.

Within a few days of DeepSeek's release, we saw a bill coming out of the Republican Party that would impose export controls on intangible technology. Not just model weights, but data and research, making it difficult or impossible for researchers to release not just weights, but potentially papers that discuss architectures or training pipelines.

### **Are we looking at models which need that kind of documentation shared?**

Yes. The best case scenario is you've released not just the model weights but you've got the training and inference code. You've got information about the architecture and ideally about the data sets used for pre-training and post-training.

The more barriers we put in place, the more friction we put in place to the release of those artifacts, the harder it is for developers to release that stuff to the public and the harder it is for third party developers and deployers to either replicate the model or inspect it or modify it and integrate it.

So, geopolitics has had a massive impact on the US's approach. We've gone from relatively technocratic concerns with online safety risk and catastrophic risk to now with the conversation heavily dominated by geopolitical competition between the US and China.

**In the EU the Commission acknowledges it needs to invest in developing and championing its ecosystem whilst trying to implement this sprawling AI Act. Could regulation work at this point in time while AI is still evolving and at the pace at which it's evolving?**

I've been very clear with policy makers in the US, the EU, the UK and elsewhere that the first thing we need to do is a gap analysis and understand where our existing regulatory and legislative frameworks fall short. The reality is there probably are some gaps and there's low hanging fruit for governments to take action. For example, until a few weeks, there was no federal criminal framework governing non-consensual abusive or intimate imagery in the US. There was civil liability, but not criminal liability. There are obvious forms of AI misuse where existing law may fall short and the government can in a very targeted, very proportionate way take action. In other cases, our existing regulatory agencies may be adequate and may be fit for purpose.

The previous UK Government was one of the few to actually do a systematic survey of its existing regulators. And of the 13 or 14 surveyed, all but one of them said they have sufficient mandate and sufficient resources to oversee the use of AI in these different domains.

So I think we need to do that gap analysis and to identify low hanging fruit interventions.

### **Do you have any thoughts on any shift in regulatory needs as a consequence of Agentic AI?**

Targeting high-risk sensitive applications of AI makes a lot of sense. If legislators and regulators agree that there are gaps in the law and existing frameworks don't adequately account for increasingly agentic systems then it may be worthwhile to introduce specific legislation or specific rulemaking targeting, for example automated decision-making or systems that have a high level of delegation or that can undertake complex tasks in sensitive environments.

A lot of this comes back to this concept of a landscaping of AI, what it does and then landscaping the regulation that's going to apply to it in the existing environment.

### **The AI safety institute in the US rebranded as the Center for AI Standards and Innovation. Do we know what that means yet?**

I think it's part of the administration's broader push to shift from a focus on the downside risk to a focus on how we realise the upside opportunity. I think it's a very good outcome that they've retained the Institute in some form. I've certainly encouraged the administration and the Institute to focus not just on evaluation and mitigation but to focus on how we help downstream developers and deployers utilise AI and apply AI safely.

And to ensure that isn't just a conversation involving half a dozen Bay Area base model developers. It's a conversation that involves thousands of downstream developers and potentially millions of deployers across the economy. I think that change in mandate and focus is promising and it's yet to be seen what that actually means in practice.

### **Where do we see the description or definition of open source going?**

I care about this definitional debate specifically as it relates to regulation of AI because what we're seeing in the EU and the US is that governments and policy makers for the most part don't want to get in the way of open source and they don't intentionally mean to stifle open source. One of the things they do to support this will be to introduce exemptions or partial exemptions for open weight or open source models.

We've seen this in the AI Act in the EU and cropping up in State proposals across the US. The challenge is that if open source is defined too narrowly, there will be very few models and very few ecosystems that benefit from these exemptions. If, for example, policy makers took the view that open source means the Open Source AI definition (OSAID) from the Open Source Initiative, then that won't apply to many Llama and Mistral models. It wouldn't even apply to DeepSeek R1. We don't have information about the DeepSeek data sets. What that means is that it potentially has significant implications for the developer ecosystem either because it might inhibit access to capable models or because they will actually have to comply if they fine-tune and re-release those models.

I think definitions matter in different contexts. In a regulatory and a legal context we need a definition that is inclusive and incentivises the release of capable open models. So it may well be that one size doesn't fit all. I just encourage that from a regulatory and legal perspective, we take a fairly inclusive lens on what is an open model. I do think OSI and OpenUK and other organisations have rightly drawn attention to the problem of "open washing" and I think it is the case that we need fair labeling and if something is not as open as it claims to be we need clear signals that that is the case.

A bad outcome would be a world in which downstream developers and deployers grow attached to particular models or particular model developers and there's a rugpool and they can't switch. They're kind of trapped with unfavorable non-open terms as they build out their businesses and as they build out their applications. I think one way to address that is by having a competitive ecosystem with good, capable, open alternatives you can plug and play and are reasonably flexible: you can move from one to the other so if you don't like the terms that you get with Llama you can move to an alternative and those alternatives ought to exist.

### **Do you have any final thoughts on the environment in the US?**

I'd just encourage people to keep an ear to the ground. This administration is a broadchurch. It has different voices on this issue. And what might work for US big tech might not work for US small tech. It's going to be interesting over the coming months as they start to flesh out that long-term strategy, how much of that strategy accounts not just for the OpenAIs, the Googles and the anthropics of the world, but also accounts for that long tail of smaller developers, independent researchers and the big grassroots community that is building on top of open weight models.

## The EU

The EU is in a difficult position. It's convened a more than thousand person working group to develop the EU AI Act Code of Practice. The Code of Practice - at least for now - is the only known means of compliance with the model provisions of the AI Act. For the most part, the Code of Practice is relatively light touch and not unexpected. But there are parts of the Code that I think are going to be challenging for certain developers.

The EU is now under a lot of political pressure to delay the enforcement of the Act and Code, at least in relation to models. So, it's yet to be seen how that develops, but I think everyone can expect a potentially fairly messy debate over the Code of Practice and whether it has been too watered down or whether it is too onerous.

## The UK

The UK is an interesting one. The UK AI Safety Institute has rebranded as the UK [AI Security Institute](#). There is a tightening of the mandate which I think is generally very positive. But the Labour government, like the Conservative government before it, has tended to oscillate between a number of different positions on how it wants to regulate or not regulate model development, particularly at the frontier.

We've seen a number of times this year the suggestion that the UK is going to regulate model development, and require that models are evaluated by the institute before they're released. That is going to face some resistance from the Trump administration and from US industry.

**There's a group of companies that have signed up to participate in that AISI evaluation in the UK using the open sourced [Inspect evaluation platform](#). What does that mean? What's the process going to be? Will it require you to engage with the institute? Will you be able to self-certify? Do we know?**

No, I think at least publicly those details are uncertain.

There are a number of Private Members' bills being put together that would codify some of this more formally. And of course some of the Bay Area companies have existing relationships, voluntary relationships with the Institute. Several major firms signed up to voluntary testing in '23 after the first AI Summit at Bletchley Park. A number certainly made undertakings as they did with the Biden Administration. What a compulsory version of that framework looks like is a bit unclear.

Moreover, it is one thing to evaluate these models but the question is what do you do with those evaluation results and what's your decision framework for determining acceptable risk? I don't think at this point the government can or should get into the business of prescribing that risk threshold. I don't think the UK's AISI or the US's Center for AI Standards and Innovation should be in the business of setting that threshold. But I do think they can build up the science of evaluation and they can standardise the process of evaluation and I think that's a net positive for the ecosystem.

## 2.2 2025 Policy Landscape

In addition to the shifts explained in Ben Brooks' Fireside chat there has been further work on defining openness in the AI context. The outputs of the AI Action Summit in Paris demonstrated a shift in the focus of the conversation, moving away from open source in AI to the use of terms like open weights, open models and digital public good or public interest in the context of AI that is open.

### 2.2.1 Open Weights Definition from the Open Weights Alliance

The approach of defining weights sits between the definition and disaggregation approach. [The Open Source Alliance](#) is a new organisation set up in 2025, and focusing on the official AI convergence challenge of the French AI Action Summit. Its purpose is to usher in global governance of open source for the AI era. By identifying national standard bearers, the OSA aims to give countries a voice in matters relating to their digital sovereignty. The OSA has been formed in response to the controversial release of the OSI definition of AI openness.



The Open Weights Definition was announced at the AI Action Summit as its first action. The purpose of the Open Weight Definition is to take the pressure off open source by giving vendors who aspire to call their products open source an alternative that only guarantees protection of 2 of the Four Freedoms - i.e. to use and share - with limited opportunity to study or modify. It's a complement to open source software not a competitor to it.

### 2.2.2 Global AI Summits French Action Summit, February 2025

OpenUK produced an [AI Action Summit Report](#) following its attendance at the Summit, represented by CEO, Amanda Brock.



Building on the 2 AI Summits organised by the UK and Korea, in 2023 and 2024 the Paris summit saw a shift from a focus on Safety to Action. The UK and US were notable in declining to sign up to the Summit Declaration.

Typical of President Macron's support of openness and public good in AI, there was a much greater focus on openness than at past Global AI Summits. This was accompanied by an obvious shift in the conversation around open source as a term and the conversations focused on software that is licensed as open source software as tooling for AI whilst open models were referred to as such.

There was also significant discussion on the need for access to data on which models are trained to build transparency and trust. Trust has become a significant theme across AI in 2025.

When discussing the value of openness in AI, there was much talk of Digital Public Goods and AI openness as a public good. 2 major initiatives in openness or public interest/ good launched as part of the Summit's output - ROOST and Current AI.

**ROOST Tools** introduces both a [foundation](#) with \$30m of funding and donations of open source software used to manage AI donated by Big Tech to ROOST. Describing itself with a view to its becoming de facto standards in good practice around AI.

**Current AI** is best explained by its founder Martin Tisné, see below.



### 2.2.3 Thought Leadership: Current AI, A global partnership for open and Public Interest AI

Martin Tisne,  
Founder  
CurrentAI



#### **The right moment: a window of opportunity in a shifting world**

The current geopolitical landscape presents a critical window of opportunity to shape the future of AI. While the United States grapples with proprietary models yet to find a stable business model amidst a shifting political environment, actors in Europe and Asia have been paving the way towards open AI, exemplified recently by models like Deepseek driven by resource access constraints.

In Europe especially, while there are strong calls for greater digital sovereignty, there's a growing awareness that regulation alone won't achieve this. This creates an opening for more autonomy in the field of AI through an open and public approach, distinct from both proprietary and nationalistic models, which Current AI aims to champion and build.

#### **Our governance: from an interim committee to a permanent dual structure**

Current AI is a new global partnership launched at the 2025 AI Action Summit, backed by an initial \$400 million (now \$420 million) from governments, philanthropics, and companies including the French government, AI Collaborative, The John D. and Catherine T. MacArthur Foundation, The Patrick J. McGovern Foundation, the Ford Foundation, Salesforce and Google Deep Mind.

Current AI is seeking association status under French law, with the subsequent goal of securing the privilege and immunity of an international organisation.

Current AI is currently run by an Interim Committee representing its core partners to ensure the organization is set up before transitioning to a permanent governance structure. A 9-member Board of Directors, comprising independent experts, will set strategy, oversee impact, and ensure independent funding decisions. A broader Council, with 12-17 delegates from participating countries, funders, and civil society, will safeguard our mandate and provide fiscal oversight. A lean Secretariat, led by a CEO, will manage day-to-day operations and execute our strategy.

#### **Scaling for impact: from 3 pilots to 10 program areas, from \$420M to \$2.5 bn**

Our 18-month launch plan prioritises building strong operational and governance foundations while piloting programmatic activities.

In 2025, we will focus on three initial pilot programs on linguistic diversity, health, and accountability. We will use these early learnings to expand to 10 key program areas. These programs will focus on expanding access to critical datasets, promoting open AI tools and standards, and establishing robust accountability frameworks, across and beyond our eleven country partners, ranging from Germany to Nigeria and India.

We aim to grow our funding from \$420 million to \$2.5 billion over the next 5 years, enabling us to support large-scale initiatives that serve public interest.

### Indian AI Summit, 2026

India will host the next global AI Summit. This is believed to be taking place in February 2026, following on from India co-hosting the French Action Summit in 2025.

Prime Minister Modi of India was very clear that there ought to be a focus on “action for all” and referenced “open source” as a means of achieving that in his [opening speech](#) at the Paris Summit.

It is therefore likely that open source or AI openness, more likely referred to as public good AI will be a centre piece of the next summit.

OpenUK will be co-hosting a series of events across India with local open source and policy partners in August as part of The Road to the Indian AI Summit, which will support the local communities in framing the open source questions for the Summit.



## 2.2.4 UK Policy Roundup

### AI Bill not forthcoming

Whilst an AI Bill was announced in the King’s Speech in July 2024, but despite rumours of a Bill it has not been shared.

### Copyright Consultation

DSIT shared a dense Copyright consultation with a primary focus on the legitimacy of use of content in AI training. A 10 week consultation followed with over 10,000 responses being received by the April deadline. The consultation received a level of criticism for trying to “appease both sides of the argument” and failing to please either the tech and AI sector or content creators, largely represented by the creative sector.

On April 23 there was a debate in Parliament instigated by James Frith MP, a supporter of the creative sector’s anti AI usage proposals. The arguments across the debate have become very emotive. The outcome of the consultation is now awaited.

### AI Action Plan

The [UK AI Opportunities Action Plan](#) was announced in February with 50 Actions and was to be “[mainlined into the UK’s veins](#)” according to PM Keir Starmer. The plan was adopted without amendment.



The plan missed the opportunity to “turbo charge” the economy using the UK’s vibrant open technology community. It also failed to shape actions around openness in AI. With an open source first policy now almost 15 years old the UK needs action not further policy, to drive the agenda forward. It is time the UK leveraged the talent in the open tech submarine under its AI economy. The plan’s author, Matt Clifford stood down in June 2025. His replacement has not yet been appointed.

### **AISI Rebrand**

The UK AI Safety Institute rebranded to the UK AI Security Institute in February 2025 following the release of an [AI Safety Report](#) and the AI Action Summit in Paris.

### **AI Action Summit**

The UK along with the US did not sign up to the Summit Declaration.

### **London Tech Week**

Significant announcements at London Tech Week in June included training for 1million young people with Big Tech.

Additionally an investment of up to £1bn is to be made in compute with up to £750m being invested in Edinburgh University.

### **Security Update, Sovereign AI and Defence**

As a consequence of geopolitical shift in 2025 there is increased discussion of Sovereign AI, although exactly what that is has yet to be defined.

### **Lord Holmes’ Bill**

Lord Holmes of Richmond instigated a private member’s Bill - The Artificial Intelligence (Regulation) Bill which began its life in the House of Lords and had reached the House of Commons before Parliament was prorogued. However, as the Bill had not passed when Parliament was dissolved, it did not progress. Lord Holmes re-introduced the [Bill as Bill 76 in March 2025](#).

### **Regulators**

#### **Bank of England**

The Bank of England published [financial stability in focus](#).

#### **Information Commissioner**

The ICO published [‘How should we assess security and data minimisation in AI?’](#).

### **The EU AI Act and Northern Ireland**

Post-Brexit the applicability of EU laws in Northern Ireland creates a complex landscape.

“The United Kingdom (UK) has so far taken a light-touch regulatory approach to its strong artificial intelligence (AI) sector. While the lack of horizontal laws may create a complicated legislative patchwork, the government argues that this is conducive to innovation and agile technology. With its new broad AI Act, the EU has taken a different path, but despite diverging approaches there are some promising signs for future cooperation.”

The AI Act applies to ‘a high-risk AI system’ released under a ‘free and open-source licence’ but includes an additional dependency that the AI system in question is placed on the market or put into service in the EU and Article 3(9) defines placing on the market as the ‘first making available of an AI system or a general-purpose AI model on the Union market’. There is then a question as to Northern Ireland being a place on which an AI System may be put on a market.

### 3. The 2025 Data to 31 May

#### Intro

This section digs into the data on AI, breaking it down initially by geography looking at the UK, the global landscape and Europe. We then consider 2 of the 2025 hot topics, Agents and Cybersecurity. This unique data from GitHub - supported by Runa Capital - updates repository-focused data 6 months on from our last AI data report. This is complemented by download data from Scarf pulled specifically for this report. To broaden the picture we also use other publicly available data from Hugging Face and LinkedIn.

#### 3.1 The UK Landscape

##### 3.1.1 UK Adoption of AI openness

The UK continues to position itself as a significant player in the global AI ecosystem, with increasing adoption of AI across sectors, strategic government initiatives, and growing engagement with AI openness. However, key structural challenges persist - particularly in talent development, data readiness, and scaling AI into production to see the elusive return on investment.

AI is already visibly reshaping employment across the UK. According to LinkedIn's May 2025 [AI Skills Trends in the UK report](#), job roles across every sector are being transformed by AI. The technology sector is leading in adoption, but AI's influence is expanding into traditionally non-digital domains. Still, growth in AI talent remains comparatively slow, prompting policy intervention.

In the public sector, the UK is emerging as a leader in applied AI. Key use cases already being deployed include: an NHS predictive tool that forecasts patient falls with 97% accuracy - expected to prevent up to 2,000 falls daily and save £2 billion annually; The AI-powered Connect tool for energy infrastructure and the Consult tool, which processes public consultation feedback to support better policymaking. These examples underscore a trend of mission-driven, high-impact AI deployments. Yet, challenges remain.

The UK is seeing rapid growth in its engagement with AI openness. Scarf data indicates that 20,000 UK companies used AI that is open in the past year. This signals rising experimentation and adoption within the private sector.

On GitHub, the UK now hosts 135 AI repositories with over 1k+ stars (as of May 2025), a 19% increase since 2024.

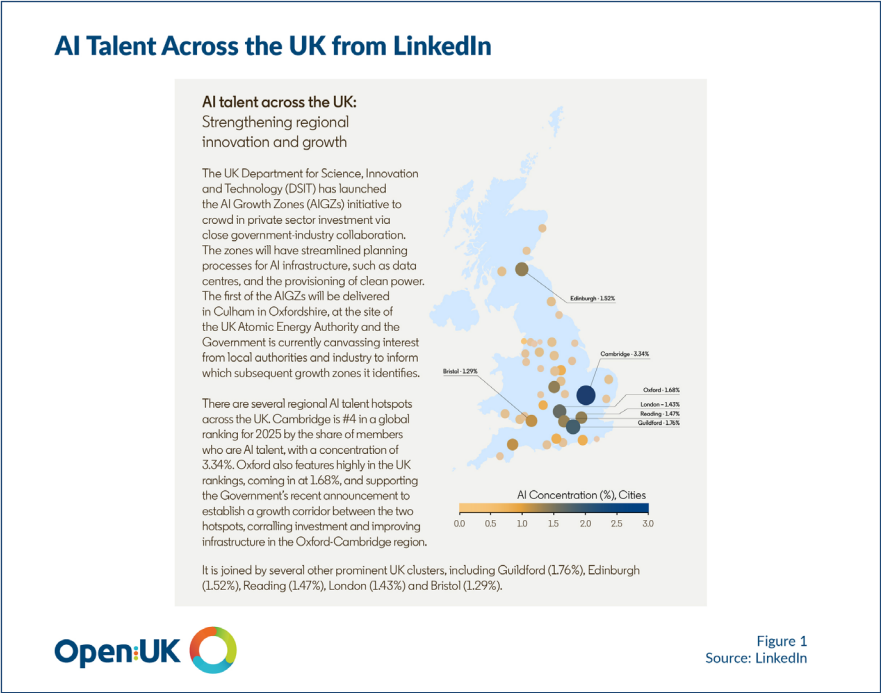
While there was a minor dip in new repository creation, the overall trend suggests continued activity and maturing innovation.

The UK remains third globally for top-starred repositories, behind only the US and China.

##### 3.1.2 LinkedIn Data

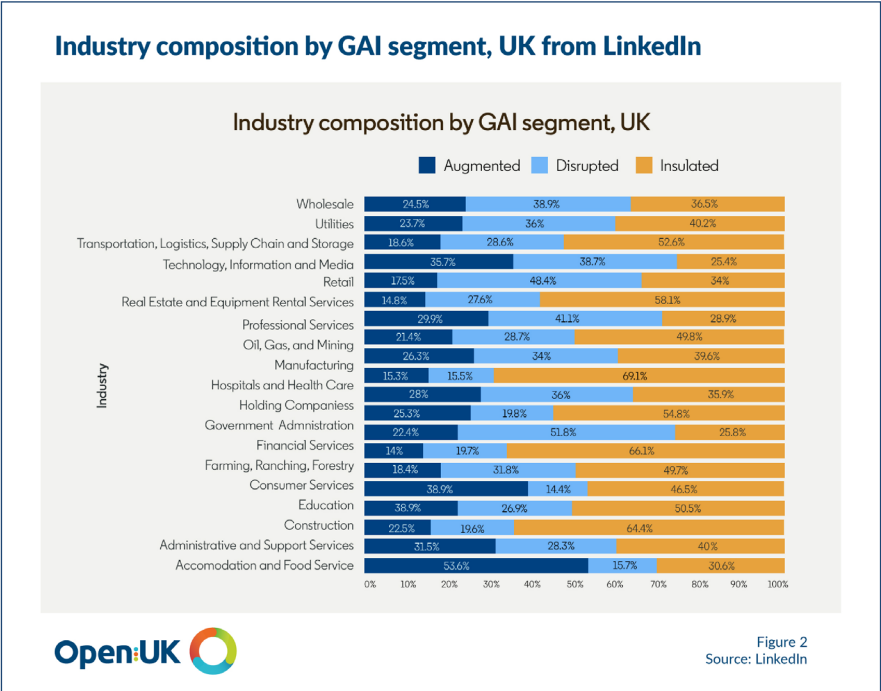
In an analysis of UK talent across the UK, LinkedIn's May 2025 Report, [AI Skills Trends in the UK](#) suggests that AI is shaping the world of work, creating a demand for new jobs and skills. Using anonymised and aggregated data from the LinkedIn platform in 2024, which is used by more than a billion members worldwide and more than 40 million in the UK it explores how AI is impacting the UK economy and workforce. The report indicates slow growth in AI talent in the UK, particularly when considered in relation to other countries.

DSIT has launched the AIGZ project to nurture and support AI talent development in the UK and Prime Minister Keir Starmer announced a [Skills Boost](#) as part of a series of AI announcements at London Tech Week in June 2025, including training of 1,000,000 young people in AI by Microsoft.



Source: [LinkedIn](#)

From the same LinkedIn report, the figure below sets out the extent to which jobs across all sectors are likely to change in response to GAI. As GAI evolves and adoption increases, jobs across all sectors are likely to change. While the Technology sector has moved more quickly than others, GAI's impact will extend well beyond the tech world in being augmented, disrupted or insulated to the progress of GAI.



Source: [LinkedIn](#)

### 3.1.3 UK Public Sector Data

According to the recent CapGemini 2025 report, [Data Foundations for Government](#), suggests that, across all regions surveyed, there are “9 in 10 public sector organisations to focus on agentic AI in the next 2-3 years, but data readiness is still a challenge and 75 percent of public sector organisations are either exploring or actively working on gen AI initiatives.

### 3.1.4.Download Data

Scarf, a company that creates open source usage analytics for sales and marketing intelligence, tracks usage of AI openness across a number of inputs. These include: Artifact downloads through Scarf Gateway, Telemetry directly from open source code, Tracking pixels in open source documentation, and Artifact downloads from partner registries.

When any given event is processed, Scarf matches the IP addresses to a company based on enrichment data from multiple metadata providers, then selects the most likely company based on the metadata provider selections. Scarf does not retain any PII associated with the events.

Through this method, Scarf has seen 20,000 companies in the UK have used AI that is open innovation AI in the past year and 8500 companies just in the past month. This represents a 3x year/year growth according to this data - seeing 8500 unique companies engaging with AI that is open in the last month. That figure was about 2700 / month last year.

Scarf also notes that they have seen about 530k companies globally using AI in the last year, 144k just in the last month. Globally that growth rate is 2.4x y/y, so by this metric, the UK is adopting OSS AI faster than the global average in our dataset. ‘Using’ in this data generally means downloaded, but also could mean invoking code that sent telemetry directly as well.

#### Global Downloads of AI that is Open

In association with  SCARF®

**2.4x year on year growth globally**



Open:UK 

Figure 3  
Source: Scarf, May 2025

#### UK Downloads of AI that is Open

In association with  SCARF®

**3x year on year growth in the UK**



Open:UK 

Figure 4  
Source: Scarf, May 2025




3.1.5 Repository Data from GitHub

Open source activity on GitHub offers a valuable lens into the evolution of AI, providing insights into where innovation is happening, how it’s being shared, and who is leading it. In the UK, the growth of AI-related repositories reflects a dynamic and expanding developer ecosystem, with openness playing a central role in making advanced AI tools more accessible and adaptable. There’s a noticeable rise in interest around agentic AI - signalling a shift toward more autonomous and interactive systems. This section explores GitHub data to better understand the UK’s role in global AI development and its commitment to openness.

Top UK AI Repositories, as at 31 May 2025

Top 12 AI Repositories

In association with 

Repo Name	Repo Description	Stars
Significant-Gravitas/AutoGPT	AutoGPT is the vision of accessible AI for everyone, to use and to build on. Our mission is to provide the tools, so that you can focus on what matters.	175846
mlabonne/llm-course	Course to get into Large Language Models (LLMs) with roadmaps and Colab notebooks.	54012
openai/openai-python	The official Python library for the OpenAI API	26929
EthicalML/awesome-production-machine-learning	A curated list of awesome open source libraries to deploy, monitor, version and scale your machine learning	18544
arc53/DocsGPT	DocsGPT is an open-source genAI tool that helps users get reliable answers from knowledge source, while avoiding hallucinations. It enables private and reliable information retrieval, with tooling and agentic system capability built in.	15680
alysxau/screenity	The free and privacy-friendly screen recorder with no limits	14567
ivy-llc/ivy	Convert Machine Learning Code Between Frameworks	14207
owainlewis/awesome-artificial-intelligence	A curated list of Artificial Intelligence (AI) courses, books, video lectures and papers.	11930
libvips/libvips	A fast image processing library with low memory needs.	10352
sashabaranov/go-openai	OpenAI ChatGPT, GPT-3, GPT-4, DALL-E, Whisper API wrapper for Go	10026
google-deepmind/sonnet	TensorFlow-based neural network library	9854
OpenMined/PySyft	Perform data science on data that remains in someone else's server	9704




Fig 5  
Source: GitHub, May 2025

AutoGPT vs DeepSeek Comparison Table

An initial experimental comparison of AutoGPT and DeepSeekV3 and DeepSeekR1 based on Github data analysed by Runa Capital suggests that AutoGPT is older and has had more time to grow, but its scale of contribution - with over 700 contributors and thousands of commits - signals a truly open, community-driven project.

In contrast, DeepSeek’s models are developed by small, centralised teams with limited community input. Despite being technically impressive, they follow a closed-core model, typical of foundation model development. The comparison highlights a broader tension in AI: agentic, collaborative innovation (AutoGPT) versus high-performance, centrally managed releases (DeepSeek).

AutoGPT vs DeepSeek Comparison Table

In association with 

Repo name	Significant-Gravitas/AutoGPT	DeepSeek V3	DeepSeek R1
Total Contributors	745	21	12
Total Commits	6744	66	36



Figure 6  
Source: GitHub, May 2025

### AI Repositories created in the UK by year created

Eight months on, we see a slight decrease on December 2024's figure, but such a small decrease in new repository creation, that it would be viewed as consistent and a stabilisation.

#### UK AI Repositories by Creation Year With 1k+ Stars

In association with  Runa Capital

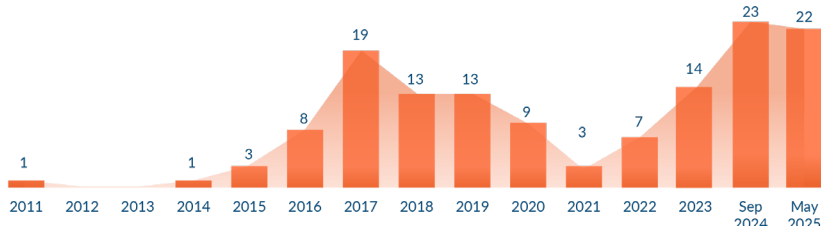


Figure 7  
Source: GitHub, May 2025

### The Number of UK AI Repositories

The total number of AI repositories with 1,000 GitHub stars in the UK increased to 135 as at 31 May, being a 19% increase on the number as at December 2024.

#### Number of UK AI Repositories With 1k+ Stars

In association with  Runa Capital

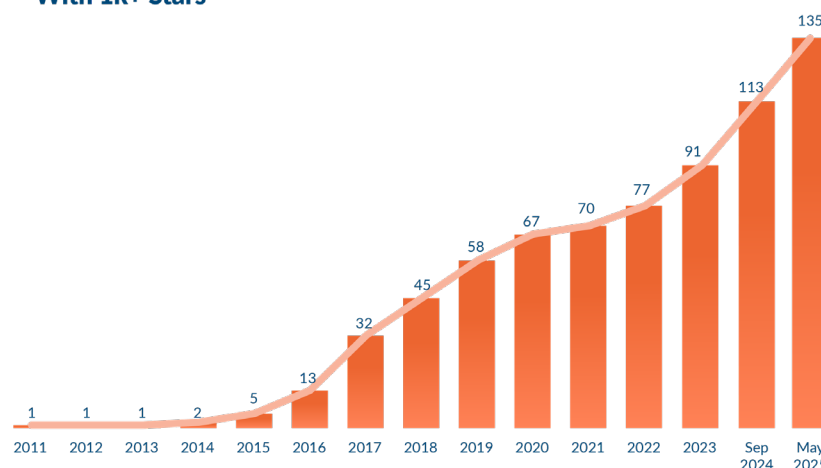


Figure 8  
Source: GitHub, May 2025

## 3.2 Global Landscape

Global AI development is progressing rapidly, driven by open ecosystems, national strategies, and commercial incentives - but scaling, infrastructure, and security remain key challenges. Overall, AI adoption is accelerating but uneven across geographies.

AI that is open continues its steady growth, as seen in Hugging Face's cumulative and monthly graphs showing a consistent rise in models, datasets, and tools. The platform reflects broadening global engagement and innovation. In the Github landscape data, provided by Runa Capital, the US leads with 1,080 1k+ star AI repositories, followed by China (613) and the UK (140).

While growth is strong across all regions, India (38%) and France (26%) are growing faster than the UK and Germany. The US also dominates in the Agentic AI realm, with 114 notable repositories, followed by China

at 36. Most other countries, including the UK and France, are in single digits - indicating early-stage exploration in this emerging AI frontier.

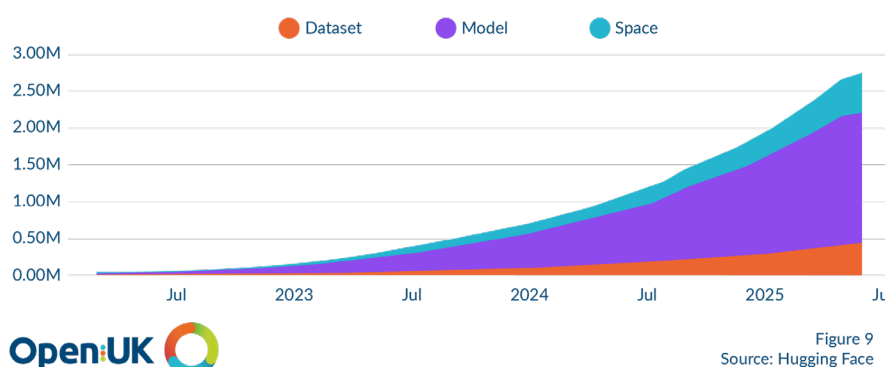
Globally, 62% of developers express cybersecurity concerns around AI that is open - slightly lower than the UK's 64%. There is a preference for proprietary models to mitigate these perceived risks.

### 3.2.1 Hugging Face Data

Drawing on Hugging Face's global data [Hub Stats](#), the charts below provide both detailed and high-level insight into how the Hugging Face Hub is evolving in terms of content and engagement. The cumulative view gives a big-picture trajectory: you can see momentum is maintaining steady upward growth. The monthly view helps you spot short-term trends aligning with other key moments in the ecosystem.

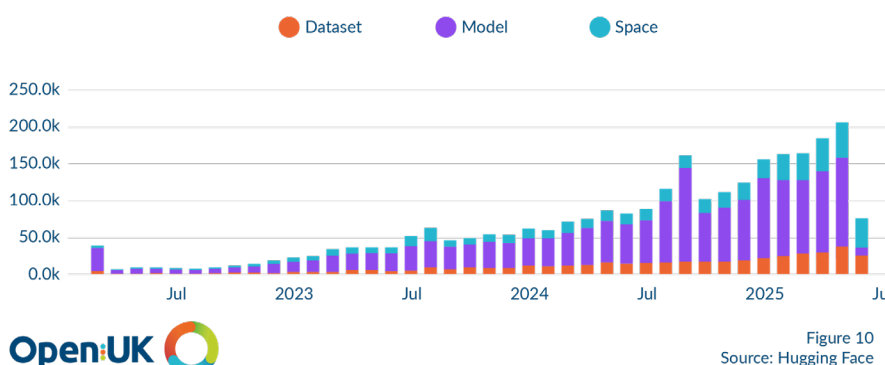
The first graph found on [Hugging Face's Hub Stats](#), shows a cumulative count over time. It stacks the monthly figures to display the total number of models, datasets, and spaces published up to each point in time. This reveals overall adoption trends and highlights when growth accelerated or stabilised.

**Hugging Face Cumulative Hub Growth**



This second graph found on [Hugging Face's Hub Stats](#) plots new creations per month on the Hugging Face Hub - categorised into models, datasets, and spaces. It shows how many of each type were added month-over-month. You can easily identify growth spurts, plateaus, or slowdowns in each content category on the platform but follow the overall growth over time.

**Hugging Face Models, Databases and Spaces Created by Month**



3.2.2 Global Rankings by Country

When we look at country-holding of repositories with 1k+ GitHub stars on a global ranking basis, we see the US leading with China trailing in second place with 57% of the repositories that the US has. The UK is third globally but with only 13% of the repositories that the US has and then Germany sits a close 4th to the UK. France is in a surprising 6th position despite France’s accolade as number one in European “open source AI” from [Tortoise Media](#) but has been growing at a faster pace than the rest of Europe.

GitHub AI Repositories with 1k+ stars by country In association with  Runa Capital

Country	Repositories with 1K+ Stars
United States	14222
China	6727
Germany	2176
United Kingdom	1894
France	1317
India	852



Figure 11  
Source: GitHub, May 2025

Looking at the top countries globally, there are increases across the board but the top countries remain the same.

GitHub AI Repositories with 1k+ stars  
labelled AI by country In association with  Runa Capital

Country	Repositories with 1K+ Stars
United States	1068
China	601
United Kingdom	135
Germany	134
France	73
India	55



Figure 12  
Source: GitHub, May 2025



### 3.2.3 Global Growth

When it comes to a consideration of growth, the UK's 19% growth in AI repositories with 1k+ stars can be seen to reflect 4th position globally.

#### Growth in AI repositories with 1K+ stars

In association with



Figure 13  
Source: GitHub, May 2025

Whilst the UK's growth compared to the EU and Germany is good, it is significantly slower than France, the US and India, with India's 38% growth being the global fastest.

#### Global Growth in Numbers of AI Repositories with 1k+ Stars in 2025

In association with

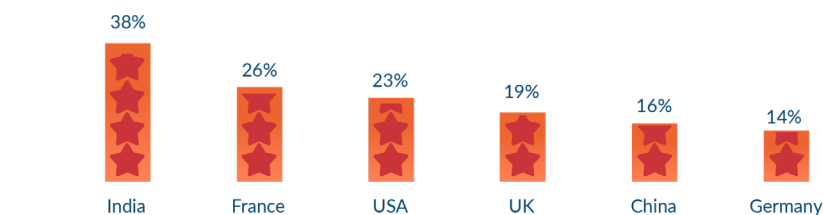


Figure 14  
Source: GitHub, May 2025

Overall, global AI development is accelerating, with open ecosystems, national strategies, and commercial momentum driving growth.

Hugging Face and GitHub data show steady expansion in models, tools, and high-impact repositories, with the US leading, China second, and the UK in third.

However, growth is uneven: India (38%) and France (26%) are outpacing the UK (19%) and Germany (14%), signalling shifting global dynamics. The challenge now is turning this global momentum into secure, scalable, and inclusive AI systems.

### 3.3 Europe - the EU and UK

The figures in Europe see a very significant shift on 2024's data across Europe and in the global landscape.

#### 3.3.1 AI Repositories in Europe

A comparison of the number of AI repositories across Europe with 1k+ GitHub stars broken down by those hosted from the UK and those hosted from the EU sees the EU's increase to 556 whilst the UK increases to 135.

##### AI Repositories across Europe

In association with  Runa Capital



Number of AI repos with  
1k+ stars in EU



Number of AI repos with  
1k+ stars in UK



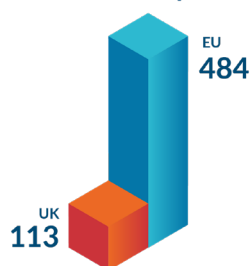
Figure 15  
Source: GitHub, May 2025

#### Growth figures

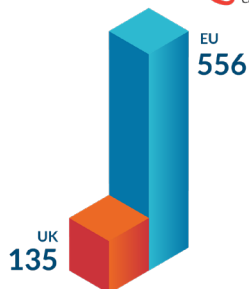
In September 2024 we saw 484 repositories with 1k+ GitHub stars in the EU and 113 in the UK. By May 2025, we saw 556 such repositories in the EU against 135 in the UK.

##### Growth in Number of AI Repositories with 1K+ stars in Europe in 2025

In association with  Runa Capital



Increase in AI Repositories  
in Europe as at September 2024



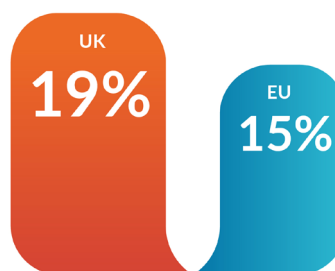
Increase in AI Repositories  
in Europe as at May 2025



Figure 16  
Source: GitHub, May 2025

##### AI Repository Growth % in past 6 months

In association with  Runa Capital



Growth in number of AI repositories  
by percentage to May 2025



Figure 17  
Source: GitHub, May 2025

EU AI repositories have grown by 15% whilst the UK sees a 19% increase over the past 6 months.

#### Top 3 European Countries

Delving further into this data we see the positioning of the top 3 countries in Europe:

UK was at 113 in 2024 and is now at 135

Germany was at 118 in 2024 and is now 134

France was at 58 in 2024 and is now at 73

Top European Countries  
by Number of AI Repositories

In association with  Runa Capital



Figure 18  
Source: GitHub, May 2025

The data shows that the UK is pushing ahead of Germany and France in terms of GitHub AI repositories with 1k+ stars. This raises the question as to whether there is a shift in the UK’s positioning in AI openness. The data however sees France as the fastest growing in Europe.

3.4 Agentic AI Data

3.4.1 The global Agentic Landscape

To get a sense of how many AI repositories are currently working on Agentic AI, we filtered AI repositories with 1K+ stars by those with the word ‘agent’ in the tag.

Top Countries by Agentic Repositories  
with 1K+ GitHub Stars

In association with  Runa Capital

Country	Agentic Repositories with 1K+ Stars
United States	114
China	36
United Kingdom	8
Singapore	6
France	5
Germany	4
Italy	3
India	3
Hong Kong	3



Figure 19  
Source: GitHub, May 2025

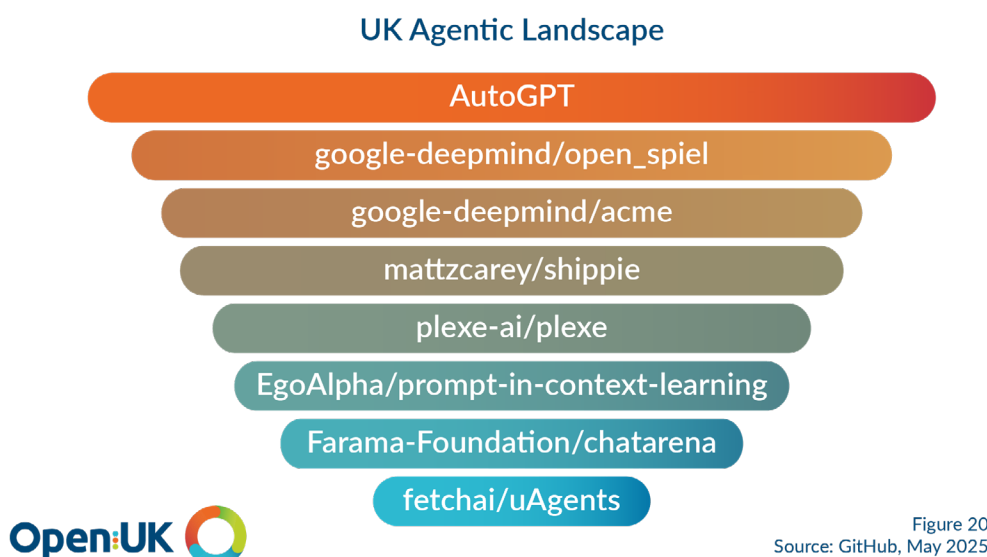
We see a shift in the conversation from LLMs in 2024 to Agentic and small models.

## UK Agentic Landscape

8 repositories on GitHub meeting the criteria of being agentic and holding 1k+ Git Hub Stars.

### Top UK Agentic AI Repositories with 1k + GitHub Stars

In association with  Runa  
Capital



With its position as the world's fastest growing open source repository in 2024, it is hardly surprising that Scotland's AutoGPT tops the UK's Agentic and AI repositories in 2025.

## 3.5 Cybersecurity Data

A recent [McKinsey](#) report found that 62% of technologists and AI developers globally were worried about cybersecurity regarding open source AI. In the UK, this is 64%.

In 2024, the [UK DSIT published a report](#) assessing cybersecurity risks specifically related to AI by identifying vulnerabilities across the AI lifecycle, evaluating their exploitation and impact, and providing real-world and theoretical case studies of cyberattacks revealed a significant lack of readiness and internal expertise in addressing AI-specific cybersecurity risks. The report conducted case study interviews and found that most organisations "confirmed a lack of AI or ML applications in their workflows, excluding standard AI tools.

Many were "unaware of AI's use and consequent cyber security risks within their operations and where organisations did recognise the risks, they lacked internal expertise for risk assessment and management. None of the clients had yet developed incident response plans specifically for cybersecurity incidents affecting AI systems and there is a lack of awareness of any specific cybersecurity regulations tailored to AI" among stakeholders. This may be linked to one of the key points made in the [CapGemini](#) report that 77% of executives prefer proprietary models for AI implementation as, among other things, they are perceived to provide better security and reduce risk. McKinsey echoes this in reflecting the stated preference of organisation leadership for use of open vs proprietary AI technologies - 72% chose security, risk, and system control as the top of 3 choices.



## 4. Agentic AI

2025 has seen a shift in the conversation from a focus on Large Language models to a focus on Agentic AI. This can be seen in the literature review at 5, with an increasing focus on agents and in Equally as with Large Language Models openness in agents is critical to access for all and the democratisation of AI.

As agents have risen up the conversation so too have concerns about their security and the unprecedented access they are granted to a company's systems. An agent in this context is like a legal representative, holding agency to act on a human's behalf. In March 2025, JP Morgan Chase's Chief Information Security Officer, Patrick Opet issued an open letter to the organisation's supply chain, stating:

"Further compounding the risks are specific vulnerabilities intrinsic to this new landscape: inadequately secured authentication tokens vulnerable to theft and reuse; software providers gaining privileged access to customer systems without explicit consent or transparency; and opaque 4th-party vendor dependencies silently expanding this same risk upstream. Critically, the explosive growth of new value-bearing services in data management, automation, artificial intelligence, and AI agents amplifies and rapidly distributes these risks, bringing them directly to the forefront of every organization."

"The most effective way to begin change is to reject these integration models without better solutions. I hope you'll join me in recognizing this challenge and responding decisively, collaboratively, and immediately."

In this section, the creator of the general purpose open AutoGPT agent, one of the most important agents globally, Toran Bruce Richards writes about the agentic landscape in 2025, clarifying that the "comprehensive access - the thing that makes security teams wake up in cold sweats - is exactly what makes them transformational".

Approaches to the security challenges faced by users of agents explored in this report include both the use of identity to secure agents and an overview of a recent DSIT led hackathon with many hyperscalers and the UK's Control Plane. At this stage in agentic's evolution opinions are mixed and this is reflected in the thought leadership we share.

OpenUK will be hosting an unconference this autumn to look at the agentic challenge.

## 4.1 Thought Leadership: Agentic AI

Toran Bruce Richards,  
Founder, AutoGPT



### The Agentic AI Opportunity: Building Your Digital Workforce the Right Way

Here's what nobody tells you about building AI agents: the moment they actually start working is the moment they need access to everything. And that comprehensive access - the thing that makes security teams wake up in cold sweats - is exactly what makes them transformational.

I've been building AutoGPT long enough [since 2023] to see this pattern play out dozens of times. A company deploys their first agent for something simple, maybe automating invoice processing. It works great within its little sandbox. Then someone realises: wait, if this agent could also check inventory levels, vendor history, and delivery schedules, it could handle the entire procurement workflow. Suddenly, that neat little box we put it in becomes the very thing limiting its value.

### The early implementations tell the story

**Hiscox Insurance:** compressed their underwriting process from three days to three minutes using AI that consolidates information from multiple sources into structured quotes.

**DigiFabster's AI Quote Agent:** transforms manufacturing quotes from multi-day ordeals into minute-long tasks, by orchestrating across production schedules, material costs, and logistics platforms simultaneously. Healthcare networks are reducing patient waiting times through agents that coordinate specialist availability and case routing.

The pattern is clear: **constrain access, constrain value.**

### Why Access Isn't Optional

We're finally building systems that work the way businesses actually operate - messily, across boundaries, with information scattered everywhere. Traditional software forced us to be the glue between disconnected tools. Now we're building digital workers that naturally span those gaps.

Think about how you actually get work done. You're constantly context-switching between tools, copying information from one system to another, mentally tracking dependencies across platforms. That digital paperwork is the tax we've been paying for decades of building isolated systems. Agents eliminate that tax, but only if they can see the full picture.

A customer service interaction is the perfect example. When a customer asks about their order, they don't care that order history lives in your commerce platform, inventory data sits in your warehouse system, and delivery tracking exists in yet another tool. They just want an answer. An agent with proper access can provide that answer in seconds. An agent with limited access becomes just another silo requiring human intervention.

What surprises people is how agent value compounds. It's not linear. Give an agent access to two systems and you don't get 2x value - you often get 10x, because suddenly it can handle entire workflows instead of just tasks. The real productivity gains aren't from agents doing tasks faster, but are from eliminating the coordination overhead that consumed so much human bandwidth. When your finance agent can directly verify budget approvals, check compliance requirements, and coordinate with project timelines, you've removed three meetings and a dozen emails from the process. The numbers back this up with 82% are already deploying agents, with 98% planning expansion. But the ones seeing the real return on their investment aren't the ones with the most agents. They're the ones whose agents can actually do meaningful work across systems.

## The Pattern That Works

Let's be honest about what comprehensive access means. [SailPoint's research](#) shows 39% of companies have observed agents accessing unauthorised systems. Another 23% report agents occasionally sharing credentials inappropriately.

These aren't inevitable problems - they're what happens when organisations try to build agent infrastructure from scratch without proper architectural controls. Modern agent platforms solve credential visibility and access scoping at the architectural level. But that's not to say security is 'solved' It's just moved up a layer. Now your platform credentials become the keys to the kingdom, creating a different but equally important security consideration.

The organisations succeeding aren't the ones limiting agent capabilities out of fear. They're the ones who have deployed agents that physically can't make dangerous mistakes by design, and so don't need to hold them back. They're implementing:

**Architectural isolation** where agents never see the credentials they're using

**Access boundaries** built into the agent design itself - where available tools are explicitly defined at appropriate points in the workflow, rather than arbitrarily selected by the agent

**Systems** where security comes from architecture, not from policies and monitoring

This is the breakthrough: when security is built into the architecture, comprehensive access stops being a liability and becomes your competitive advantage. You can give agents the keys to everything because you've designed the locks.

## The Window Is Now

We're in that brief moment where competitive advantage is up for grabs. The companies figuring out how to deploy agents comprehensively - with appropriate design but without crippling limitations - will define how business operates for the next decade. The acceleration is remarkable. Security leaders are refreshing their AI strategies constantly because the opportunity cost of waiting is too high. Every month of delay is another month of manual processes your competitors might be automating.

When agents can seamlessly orchestrate across systems, you can design workflows that were literally impossible before. You can make decisions based on real-time synthesis of information that previously took days to compile.

## Building Forward

The question isn't whether to give agents comprehensive access - it's how to do it intelligently. With the right architecture, we've solved the obvious problems. Agents can't leak credentials they can't see. They can't break out of their designed workflows. Security is structural, not supervisory.

But let's be honest about what that means. When you solve credential sprawl by consolidating access through a platform, you're making a trade. Yes, you've eliminated entire categories of security failures. But you've also created a new challenge: platform credentials become incredibly powerful. This concentration of access demands a different kind of vigilance.

The organisations that win won't be the ones that pretend this trade-off doesn't exist. They'll be the ones that embrace it while building the right safeguards. They'll treat platform security with the gravity it deserves while still moving fast enough to capture the opportunity. The transformation is happening with or without perfect answers. The organisations that embrace the complexity while building towards solutions will own the future. Those waiting for complete clarity will find themselves buying that future from others.

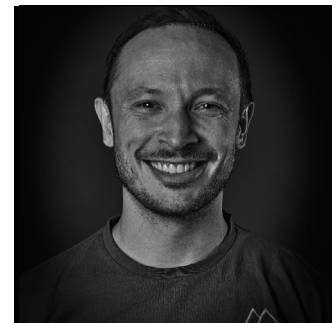
I'm certain that today's agents will look primitive in retrospect.

The capabilities we're unlocking now - as powerful as they seem - are just the foundation.

But the architectural decisions we make today about access and security will determine who can actually harness what's coming.

## 4.2 DSIT, Identity, and AI Agent Security

Andrew Martin,  
Co-Founder and CEO  
ControlPlane



AI agents present us with an intriguing paradox: an attempt to secure autonomous systems with identity frameworks designed for distinctly non-autonomous humans. As [1 Password observes](#): Legacy IAM, IGA, and MDM tools... assume interactive logins, static access patterns, and human oversight. AI agents don't fit this model. Single-sign-on and multi-factor authentication (MFA) are designed for human users. We can memorise passwords without writing them down, and carry physical second-factor authentication tokens.

AI agents that operate continuously and programmatically, requiring runtime authentication and fine-grained access control that traditional systems simply weren't designed to handle. This is an architectural mismatch. Like human users, AI agents can access regulated data. As they're trained on the corpus of human information, they operate at a similar or greater level of inherent fallibility to humanity.

Their actions must be auditable and compliant with data and security standards (like SOC 2, HIPAA, and GDPR). The problem space widens when we consider industry predictions. According to [Gartner®, TSP 2025 Trends: Agentic AI – The Evolution of Experience](#), February 2025 by 2028, at least 15% of daily work decisions will be made autonomously through agentic AI, up from 0% in 2024. This rapid shift from human-mediated to autonomous decision-making suggests this is a sensible moment for a fundamental rethinking of our security and identity management approaches.

### The UK Government Response: GDS Zero Trust Hackathon

Building on this threat model, the UK government took proactive action through the GDS Zero Trust Hackathon, bringing together major technology providers and government stakeholders to prove interoperability between technical suppliers into the UK government and public sector.

Government bodies including the UK National Cyber Security Centre (NCSC), UK Government Security Group, and the Government Digital Service (GDS) collaborated with major cloud providers Google Cloud, Microsoft, Github, Oracle, IBM, Cloudflare, Okta, Red Hat, and ControlPlane. The event centered around innovation in identity management and its intersection with government policy. Participants explored challenges in digital infrastructure security, with special attention to workload identity (agents are an autonomous “workload”) and ensuring seamless interoperability between different providers.

This work extends beyond traditional workloads to include AI agents, recognizing that the future of government services will increasingly involve autonomous systems, and that nondeterministic AI agents may not respect traditional boundary concepts even when instructed to do so. The hackathon identified three critical interoperability levels that must function seamlessly:

**Intra-Domain Interoperability:** Securing communication between workloads within the same domain (e.g. cluster, service, cloud, organisation etc), the foundation upon which everything else builds.

**Cross-Cloud Interoperability:** Enabling secure interactions between workloads across different cloud providers. This addresses the multi-vendor reality that most organisations now face.

**Cross-Domain Interoperability:** Supporting secure communication between workloads in different security domains or organisational boundaries. This is the zero trust architecture that organisations are still struggling to embrace.

Agents just like their human counterparts operate across all levels simultaneously, demanding a level of interoperability sophistication we've never attempted so widely before.

The team is developing an open artifact to demonstrate secret-less cloud and multi-vendor architectures to benefit the open source community. You can follow the progress at [github.com/co-cddo/zero-trust-cloud-identity](https://github.com/co-cddo/zero-trust-cloud-identity).



## Agents and the Workload Identity Landscape

As the agent ecosystem evolves, two major protocols are emerging to address different aspects of agent communication.

**MCP** focuses on the interaction between LLMs and your app's tools and context.

**A2A** (Agent-to-Agent) focuses on enabling inter-agent task exchange.

For more details on the architectural differences, see the [MCP Architecture Specification](#).

The browser-driven ChatGPT-like interface that everybody knows and loves only represents the beginning of AI. When we want to connect Claude or ChatGPT to our email or GitHub, we use Model Context Protocol (MCP) servers to facilitate this connection. This is an evolution from remote procedure calls, message buses, and APIs, and hopefully not the end.

MCP functions as an API layer designed specifically for LLMs and other AI agents, theoretically supporting sophisticated authentication and authorisation flows using OAuth 2.1. The MCP specification has standardised the [OAuth 2.1 IETF DRAFT](#) authorisation flow with PKCE as best-practice according to the [MCP specification](#) though current implementations vary widely.

But while OAuth was designed to connect two systems (like ChatGPT and Gmail) and supports fine-grained access controls, current MCP implementations have a significant limitation: in the goldrush to ship the first MCP connector for each service, many developers have bypassed the fine-grained access controls that OAuth was designed to support.

Many MCPs are “unofficial”, and not supported by the services they interact with. This creates a false sense of security: the appearance of robust authentication masking inadequate authorisation and data controls.

**The Cloud Security Alliance identifies the core issue, [Cloud Security Alliance - MCP OAuth 2.1 PKCE and the Future of AI Authorization](#)**. While the MCP Authorization Specification clearly defines the process for acquiring an access token, a critical aspect remains open: defining how non-human entities and autonomous workloads authenticate to the Authorization Server. This is a significant concern for platform engineers managing AI deployments who need to establish secure workload identities for these agents.

This observation cuts to the heart of our challenge. OAuth excels at delegating human authority, but what happens when there's no human to delegate from?

OAuth provides a solid foundation for delegating human behavior. But this does not prevent the agent from acting out of character, or in a way misaligned with the user's interests. Tooling and papers including [Agent Smith](#) (which demonstrates exponential speed of infection for jailbroken agents) and [AgentDojo](#) (a testing lab for attacks on agentic platforms) demonstrate that this is a near and present danger to current systems and approaches.

**The Security Challenge:** From Human to Agent-to-Agent authentication and authorisation closely parallels the microservices paradigm. Traditional approaches using slow-to-rotate credentials or passwords are insufficient for this new landscape.

The solution lies in workload identity, which takes a secret or identifier for each microservice or agent and exchanges it for a short-lived token managed by a central identity system.

This enables zero trust architectures where every connection is verified, and can be implemented using various technologies including OAuth, SPIFFE/SPIRE, JWTs, or Kerberos. This level of zero trust, where we verify each connection every time, is required for modern, secure systems. Tools like Istio provide this wrapper of workload identity automatically for cloud native applications: accumulative attestations provide metadata for ABAC level control, and ultimately zero trust systems based on properties of the inter-service connection. For example a service, at a time of day, connecting to another service, with a specific credential, provides a basket of data with which to make an authorisation decision on the path and content of the connection.

Robust agentic identity and access management controls are crucial due to regulations like the EU's AI Act, which demand transparency and accountability. So what could possibly go wrong?



Consider a user who wants to book a doctor's appointment. Where they may have previously asked an assistant, they now ask an LLM. In both cases, they trust the delegate not to succumb to social engineering attacks, and to faithfully carry out the instructions without leaking confidential information.

So the user asks an LLM, which spins up agents to: pull medical data from the email inbox; arrange a doctor's appointment; and then book transport to the surgery. The medical data retrieved from email should never enter the transport booking system. Similarly, only required inbox contents should not be fed to any other agent.

Data oversharing becomes a primary concern—each agent spun up by a master agent should only have access to the data it needs, but without proper controls, we risk sending confidential information intended for recipient A to recipient B. Access control challenges compound this issue, as agents receiving full access to resources like email can potentially access and share information beyond their intended scope.

And perhaps most critically, the path of data between agents must be both minimum viable and verified, creating verification challenges that existing systems struggle to address. Zero trust, with versioned data schemas and data loss prevention is the appropriate paradigm, but we have not embraced these practices across the industry yet.

### The Agent Ecosystem

Agents may take many forms, but the current state of our understanding covers four primary cases, and all permutations of interaction between them (A Novel Zero-Trust Identity Framework for Agentic AI: Decentralised Authentication and Fine-Grained Access Control, [Huang et al. \(2025\)](#)):

**Persistent Agents:** Long-lived services maintaining state across interactions, requiring session management and credential rotation strategies.

**Ephemeral Agents:** Task-specific instances created and destroyed dynamically, challenging traditional identity lifecycle management.

**Scaled-Out Duplicates:** Parallel instances handling similar tasks simultaneously, requiring consistent identity without collision.

**Hierarchical Agent Structures:** Master agents spawning specialised sub-agents, creating complex delegation chains that must be securely managed.

These structures can be mapped to social interactions in daily life: guards at a gate, contractors completing a specific task, groups of stewards at a large event, and management structures. Agent ecosystems are not unique, they are just “artificial”, and humans can be equally dangerous and misled too (as phishing and con artists have taught us recently, and eternally).

The complexity of architectures and systems requires governance over behavior, examining and monitoring agentic actions to determine possible compromise of the model even with correct authentication. By creating a new layer between humans and machines, agents must be identified with unique identifiers so that logs, network behavior, and data access can be audited. And there is a clear identity between the behavior of the agent and the human it represents.

### **The Path Forward**

To succeed with secure agentic AI deployment, we must transcend our persistent issues with human security and embrace those learnings in workload identity, zero trust for agents. We have the capability to secure these agents with existing tools, but for truly autonomous distributed systems greater levels of data classification and behavioral monitoring are required than with traditional microservice systems.

While MCP provides a useful human-to-agent delegated authority model, agent-to-agent and future decentralised verifiable identity frameworks must become easier to deploy and reason about in order to enable secure, widespread adoption. Ensuring compliance with emerging AI regulations entails a building of sustainable trust in autonomous systems. The work being done through initiatives like the GDS Zero Trust Hackathon provides a foundation, but the industry must continue to innovate and collaborate to meet the challenges ahead.

The path forward requires not just technical solutions, but a fundamental shift in how we think about identity, trust, and security in an age where autonomous agents will become integral to our digital infrastructure.

### 4.3. Thought Leadership: Identity

Sal Kimmich  
CEO, Gadfly AI



#### Time becomes the 4th Dimension of Trust

In today's hyper-distributed, Cloud Native ecosystem, identity is no longer a human-first problem. Instead, software workloads-containers, virtual machines, and serverless functions-must authenticate and authorise each other dynamically, at scale, and with provable integrity. This fundamental shift in computing has birthed a new class of identity: workload identity.

At its core, workload identity refers to a unique, verifiable identifier assigned to a non-human actor, allowing it to prove who it is across systems and clouds. In practice, this might take the form of a SPIFFE ID, bound to a short-lived X.509 certificate or JWT (SVID) issued at runtime.

But this isn't a new problem-it's an evolution decades in the making. Tracing its origins reveals a lineage of security protocols: the Needham-Schroeder mutual authentication model, Kerberos ticketing systems, X.509 public key infrastructures, and federated identity systems like OAuth2 and OIDC-all precursors to what SPIFFE/SPIRE would eventually standardise.

Understanding workload identity begins with the concept of trust boundaries-those invisible lines across which authentication must be reasserted. Whether it's a user logging into a platform, a microservice calling another API, or a Virtual Machine requesting access to a secure resource, each of these interactions crosses a boundary that demands scrutiny.

Where there are boundaries, there must be anchors-roots of trust like TPMs, secure enclaves, or cloud-native certificate authorities. These anchors validate the claims identities make as they cross domains. Trust binding, then, is the act of associating a workload's identity to real-time evidence, answering not just "Who are you?" but "Why should I trust you right now?"

The introduction of the Secure Production Identity Framework for Everyone-SPIFFE-formalised this model. Its reference implementation, SPIRE, automates the secure issuance, rotation, and revocation of identity credentials through attestation. Through SPIRE, ephemeral SVIDs-X.509 or JWT-based-can be issued dynamically, cryptographically bound to runtime metadata, and delivered without embedding secrets in code or configuration. This makes mutual TLS between workloads not just possible but automatic, cementing SPIFFE as a foundational building block of zero-trust systems.

But even SPIFFE isn't a silver bullet. Its design elegantly handles workload identity within a single trust domain, but as soon as identities cross federated, multi-cloud, or confidential computing environments, things get complex. SPIFFE Federation allows multiple trust domains to exchange bundles of trusted cryptographic material-public keys, CA roots, and metadata-so workloads from one domain can verify the identities issued by another. In effect, SPIFFE becomes to workloads what SSO is to humans: a portable, automated identity layer that eliminates the risk of static secrets and enables scalable trust across systems.

This becomes especially critical in today's environments, where workloads are not only short-lived but hyper-ephemeral. Containers start and stop in seconds. Serverless functions scale to zero. Autoscalers spin up hundreds of instances on demand. Manual secret distribution simply doesn't scale. In these conditions, workload identity must be issued just-in-time, enforced through policy, and expired rapidly. SPIRE was built with this in mind-but even so, real-world implementations face latency and revocation challenges. SPIRE relies on short-lived SVIDs to mitigate the need for traditional revocation mechanisms like CRLs or OCSP, but this makes instantaneous invalidation difficult. In federated systems, this becomes a pain point. Emerging patterns-like JWKS caching, STS-based credential brokers, and mesh trust proxies-aim to smooth out this complexity by scaling cryptographic validation while preserving flexibility.

At the core of this evolution is the shift from static trust to dynamic trust. Where older models relied on long-lived credentials and pre-established policy, today's systems must prove themselves at runtime. This is where confidential computing enters the picture. Trusted Execution Environments (TEEs) like Intel SGX, AMD SEV, and ARM CCA enable workloads to run in isolated enclaves where neither the OS nor cloud provider can interfere. When workload identity is paired with attestation from a TEE, it enables a new level of runtime verification-not just that a credential is valid, but that the software presenting it is running in a secure, verified state.

Depending on the architecture, this attestation may be platform-level (e.g., an SGX quote) or application-level (e.g., an IMA measurement or enclave signature). In either case, attestation becomes the runtime heartbeat of trust, verifying not only identity, but integrity. It's no longer enough to ask "Who are you?"-we must ask "What code are you running, where, and has it changed?"

This model is particularly compelling as we enter the age of AI-driven infrastructure.

Workloads are no longer static binaries; they are models, agents, and orchestration layers capable of adapting, evolving, and reasoning. An autonomous agent might modify its behavior or dynamically request new permissions. In this context, AI observability must be layered with identity observability. Workload identity systems can provide the foundation for anomaly detection-flagging lateral movement, unexpected privilege use, or access patterns that diverge from learned behavior. Machine learning models can detect the drift, but only if identities are granular, auditable, and cryptographically verifiable.

This is also where time becomes the 4th dimension of trust. Identity must no longer be seen as static or eternal. Instead, it should be contextual and ephemeral-a reflection of the workload's current behavior, not just its intended role. A credential that made sense 5 minutes ago might be irrelevant-or dangerous-now. Static credentials, even if rotated, offer no protection against behavioral drift. The future requires temporal governance: strict enforcement of identity expiration, context-aware attestation, and runtime policy enforcement.

Policy engines like [Open Policy Agent \(OPA\)](#) allow real-time decision-making based on identity claims, request metadata, and even the attestation timestamp itself. This enables systems where identities don't just expire-they age out of trust based on time and context. In Cloud Native settings, tokens might be issued with a TTL of 15 minutes, rotated continuously, and invalidated automatically based on risk signals or revocation events.

In this model, temporal binding becomes a security control. It's not enough to know who a workload is-we must know when that claim is valid, and under what conditions. If an identity is reused outside of policy bounds, or after behavioral drift, it should be automatically revoked or denied. This is essential in environments where workloads are spawned by AI agents, or where systems self-replicate or evolve in real time. The old security perimeter is gone; the new perimeter is trust, and trust must have a half-life.

This brings us to sovereign infrastructure. In the context of rising geopolitical pressure and regulatory fragmentation, sovereign cloud architectures are being demanded by many. It's not just about verifying who a workload is. It's about verifying where it was run, by whom, under what controls, and with what governance. For a cloud spanning France, Italy, and Switzerland, attestation may need to prove geographic location, operational control, and runtime measurement. In these cases, workload identity isn't just a technical control-it becomes a legal proof of sovereignty for those who require it. Ultimately, workload identity is evolving beyond "Who is this service?" to "Why do we trust this actor now, under these conditions, for this limited time?"

It is identity as a verb-not a label.

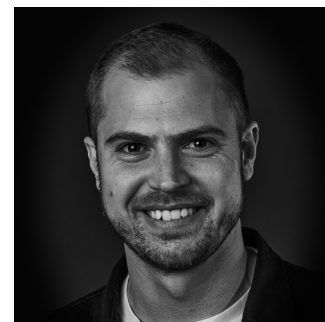
It's about active trust, driven by cryptography, policy, attestation, and real-time context. It's about treating non-human actors-code, containers, models-as first-class citizens in a zero-trust world. And it's about embracing the fact that trust is no longer just a relationship. It's bound to a moment in time and the context of that time.

Understanding workload identity, trust boundaries, and time-bounded credentials is not just preparation-it's prerequisite knowledge for securing evolving AI architectures. As we shift from securing static code to securing reasoning agents, the principles of dynamic trust, remote attestation, and contextual identity must evolve to account for decision-making logic, emergent behavior, and recursive autonomy. What we've learned from securing workloads will become the scaffolding for a new era of cybersecurity-one where we must prove not only what a system is, but why it chose to act.



## 4.4 Thought Leadership: Securing Agentic AI

**Matthew Barker**  
VP Workload Identity CyberArk



It took almost 6 years of dedicated effort to reach 12,000 Github stars with cert-manager. So in early 2023 when I heard AutoGPT hit over 50,000 stars in a matter of days, I needed to know more.

As we already know, AI needs prompting. But what happens when you want something more than just an answer to the question? What if, for example, you have a complex, multi-step task to complete, and a huge variety of ways to go about it?

As it happens, a Scottish engineer running a game studio in Glasgow created and released a model which did exactly that. AutoGPT gave the user an ability to set AI a complex, long-term objective and then sit back and wait for the results. If, for example, you wanted to generate viral videos from trending topics. You can get an agent to read topics on Reddit, identify what's trending and then automatically creates a short-form video based on the content. In this case, the AI agent is independently working out how to solve the problem over multiple steps, taking its own decisions on how to achieve the goal - all without further prompting or intervention on the part of the human.

This paradigm set the tech world ablaze with ideas on how the world of work could be transformed by a swarm of AI 'agents' working independently on your behalf.

We were told to imagine a whole army of the world's smartest employees working in unison to give you the results you need, only limited by the cost of the tokens and the limit of your own creativity. Back then it sounded too good to be true, but fast forward a couple of years, and the realities of using agents are coming into reality. Take for example, Lovable and Bolt. Built on the principles of agentic AI, they have become two of the fastest growing companies in the world, offering an incredible ability to test and build new applications by focusing specifically on code development.

Now we're at a point where the value has been proven, I have never seen more excitement from enterprises as they explore how this technology can radically change the way they do business.

This is all well and good, but before we unleash thousands of smart agents that have the ability to independently make decisions, we should perhaps take a small pause to consider the security implications. Yes, agents are goal oriented, but unlike humans they don't understand "right" or "wrong," "secure" or "insecure." So what's the worst that can happen? In my role, this is the type of question we wake up thinking about.

So where can it go wrong? Here are a few areas identified:

### **Access Control Loopholes**

Like many technologies before us, speed and convenience get prioritised over security. Teams, eager to get productive quickly, may assign "super-admin" roles to agents. The logic? Easier to do, faster to finish, less to manage. The reality? An agent with overly broad permissions making the 'wrong' decision, or getting compromised, could cause huge issues across large swathes of your network.

### **Data Privacy and Confidentiality**

Imagine an AI agent with access to your customer database. It processes transactions, stores sensitive data and personalises experiences. Now, what if the agent makes a bad decision, or is compromised? Every piece of customer information becomes susceptible to theft or misuse in an instant. As new regulations around data privacy emerge, a breach of this nature means more than just operational headaches. It could invite heavy compliance, fines and erode hard-earned customer trust.

### Model Security Vulnerabilities

AI models depend on properly trained and secured systems. However, these are prone to exploitation. For example, inputs can be manipulated to skew results, enabling adversaries to introduce bias or even destabilise the entire model. Now, imagine an unguarded AI agent deployed in financial services, making false predictions following a poisoning. This could lead to catastrophic losses either through impact on share price, regulatory fines, or brand damage.

When we think about agents in the context of security, they pose a unique risk. This is because they are incredibly easy to set up - in many cases by people who have little understanding of tech or security - and yet they are massively powerful, and potentially highly privileged. This is a mix that is bound to cause huge issues as we learn the hard way how to protect, and protect ourselves from AI agents.

### So what can you do to help mitigate these risks?

Here are a few things to think about:

**Security-First Architecture:** Every AI initiative must include dedicated security architects from project inception, not as a later-stage checkpoint. This ensures security considerations shape system design rather than constraining it after implementation.

**Identity as First Principle:** Treat AI agent identity management as a fundamental first principle. This means dedicated resources, executive oversight and integration with existing identity platforms rather than ad-hoc solutions. Our ability to secure AI agents and ultimately have a kill switch if they misbehave or are compromised depends on strong identity.

**Built-In Kill Switch:** Once every AI agent – and each of its thousands of copies – have a unique identity, the right level of privilege controls can be applied. Zero standing privileges work seamlessly. And if agents misbehave or are compromised, identity can be used as the kill switch for the 21<sup>st</sup> century.

**Vendor Due Diligence:** Evaluate AI platform providers through the lens of identity and security capabilities, not just functionality. Organizations should prioritise vendors demonstrating mature approaches to AI agent identity management and granular access controls.

The organisations that address AI security challenges strategically will capture sustainable competitive advantages. They'll scale AI deployments confidently, knowing each agent operates within defined boundaries. They'll avoid the incidents that will inevitably impact competitors treating AI agents as conventional software. Most importantly, they'll build trust with customers and stakeholders by demonstrating responsible AI governance. The alternative carries significant business risks. Security breaches involving autonomous AI systems could result in regulatory penalties, customer defection and reputational damage that extends far beyond traditional data breaches. When an AI agent operates autonomously with excessive privileges, the potential for widespread impact increases exponentially.

As AI becomes central to business operations across industries, security transitions from a technical concern to a business imperative. The executives who recognise this shift early - and invest accordingly - will position their organizations to harness AI's transformative potential while avoiding the pitfalls awaiting those who prioritise speed over security.

The time to build that foundation is now, before the damage gets done, and autonomous systems become too embedded in your operations to secure retroactively.

#### 4.5 Thought Leadership: Quantum Readiness, Securing AI and Public Trust



Isabelle Mauny  
EU Field CTO WSO2



Richard Steel  
Technology Advisory Services

The convergence of AI advancements and quantum computing breakthroughs has created an urgent need for quantum safeness—the proactive adoption of post-quantum cryptography (PQC) to safeguard critical infrastructure, AI systems, and maintain public trust.

##### The Quantum Threat to Data Confidentiality and Trust

Quantum computers threaten to break widely used encryption methods that underpin data confidentiality and digital communications.

##### Key risks include:

**“Store Now, Decrypt Later” attacks:** Adversaries are already collecting encrypted data today, with the intent of decrypting it once quantum computers become powerful enough. Sensitive information such as financial transactions, customer records, and intellectual property could be compromised retroactively.

**Transitioning to post-quantum cryptography (PQC):** before quantum computers render current encryption obsolete. All industries are potentially impacted but banking, finance, healthcare and government are the most critical.

**The timing challenge for quantum-safe transition:** can be understood through Mosca’s inequality: if migration time plus data confidentiality period exceeds the time until quantum computers become crypto-relevant, organisations face inevitable security compromise. If an industry mandates long-term data storage, organisations must act within a limited window to encrypt that data using quantum-safe methods. Any information encrypted with RSA or ECC today will be at risk once quantum computers become capable of breaking these algorithms.

**Erosion of Public Trust:** Breaches enabled by quantum attacks could destabilise confidence in any digital services like AI-driven services, smart cities, or automated decision-making.

**AI Model Vulnerabilities:** Compromised encryption could enable adversarial attacks on AI training data or model weights, undermining reliability.

The UK’s National Cyber Security Centre (NCSC) warns that PQC migration is unavoidable, setting a 2035 deadline for full transition. Moreover, according to the agenda of the EU Cooperation Group on Network and Information Security (NIS), all EU member states should at least start switching to post-quantum cryptography (PQC) by the end of 2026. The experts also warn that PQC should be used in critical infrastructures such as the energy and telecommunications sectors “as soon as possible, but by the end of 2030 at the latest”.

##### Post-Quantum Cryptography: Technical Foundations

Today’s encryption relies on mathematical problems that are extremely difficult for conventional computers to solve. RSA encryption, which protects everything from online banking to government communications, would take conventional computers billions of years to break—essentially the lifetime of the universe. This also applies to ECC (Elliptic Curve Cryptography), and other foundations of modern digital security.

Quantum computers change this equation fundamentally. Shor’s algorithm can solve these same mathematical problems exponentially faster—what would take conventional computers billions of years, quantum computers will accomplish in days. NIST has already standardised post-quantum algorithms, as per the table below. Additionally, symmetric cryptography (AES) remains quantum-resistant with adjusted parameters, offering a bridge for legacy systems.

Post Quantum Algorithms

Approach	Examples	Status
Lattice-based	CRYSTALS-KYBER (i.e. ML-KEM, ML-DSA)	NIST-standardized (FIPS 203/204)
Hash-based	SPHINCS+	NIST-standardized (FIPS 205)
Symmetric Key Updates	AES-256	Doubling key sizes counters Grover



Figure 21  
Source: NIST

PQC Adoption Strategies and potential challenges

Strategies:

**Risk Assessment (now):** Inventory cryptographic dependencies in data storage, communications and AI implementations.

**Hybrid Deployments (2025–2030):** Combining classical (C) and post-quantum (Q) algorithms, we can ensure resilience against classical or quantum attacks, while allowing users to gain trust in the new quantum algorithms. Additionally, web browsers such as Google Chrome will use quantum-safe algorithms whenever available. As an example, starting with [WSO2’s Ballerina Swan Lake Update 9](#), all service-to-service communications use hybrid schemes, ensuring quantum-safe connections while maintaining compatibility with existing systems.

**Full Transition (2030–2035):** Prioritise high-risk sectors (healthcare, energy) and align with [NCSC’s 2035 deadline](#).

Potential Barriers

**Establish awareness:** Quantum threats are still not widely known, and in particular the imminence of the threat is not well understood. It is critical to invest in quantum education.

**Establish trust:** Post-quantum algorithms are gradually gaining prominence but require time to establish trust within the industry. Larger Key Sizes and Potential Performance Overhead: Post-quantum cryptographic algorithms require in general larger key sizes compared to classical cryptography, leading to increased bandwidth consumption, storage demands, and computational latency, especially in resource-constrained environments like IoT ecosystems or high-traffic networks. Recent performance analysis has however demonstrated that in the case of modern machines, the latest NIST-certified algorithms deliver much increased security at a very high performance.

Impact on PQC Adoption

**Bandwidth and Latency:** Larger keys increase data transmission requirements, affecting real-time systems.  
**Storage and Memory:** Key storage needs expand by 2–5× compared to RSA/ECC, straining embedded devices.

**Cost and Complexity:** Enterprises report limited awareness of PQC’s urgency amid competing investments.

Open Source and Global Governance

The AI Action Summit highlighted open-source tooling as critical to democratizing PQC adoption:

**Post Quantum Cryptography Alliance (PQCA):** provides open source reference implementations of PQC algorithms and supported by companies like Google, Microsoft, Meta and others.

**ROOST Initiative:** A €30M suite of open-source safety tools promotes transparency in AI and cryptographic governance.

**Public Interest Foundations:** Coalitions like Mistral’s €400M fund aim to align PQC migration with SDGs, ensuring equitable access.

Recommendations for Policymakers and Practitioners

**Harmonise Global Standards:** Ensure alignment of NIST, NCSC, and EU PQC timelines to prevent fragmentation.

**Fund Open-Source PQC Libraries:** Accelerate community-driven development,  
**Mandate PQC in AI Ethics Frameworks:** Link cryptographic resilience to AI accountability protocols.

Conclusion

Quantum safeness is not merely a technical challenge but a societal imperative. By integrating PQC into AI governance and public-interest frameworks, stakeholders can preempt existential threats while fostering trust in next-generation technologies.

As [Yann LeCun](#) noted at the AI Action Summit, “openness is the root to safety”-a principle that must guide both AI innovation and cryptographic resilience.

## 5. AI Global Literature Review

### 5.1 Key themes from the 2025 reporting landscape

An expansive literature review across almost 30 reports in 2025 show some key themes.

The reports reviewed are:

1. McKinsey, Open Source Technology in the Age of AI, 2025
2. Tech Nation, UK AI Sector Spotlight 2025
3. MIT & Databricks, The Great Acceleration: CIO Perspectives on Generative AI, 2025
4. Snowflake, AI & Data Predictions 2025
5. Perforce OpenLogic, State of Open Source Report 2025
6. Stamford and Human Centred Artificial Intelligence, Artificial Intelligence Index Report, 2025
7. IBM, Open Source AI in 2025
8. Accenture, Technology Vision 2025
9. Writer, New Data on Navigating the AI Adoption Gap, 2025
10. Forbes, The 5 AI Trends in 2025
11. KI Company, the 5 Best Open Source AI Models in 2025
12. IEEE Spectrum, 12 Graphs that Explain the State of AI in 2025
13. Nature, AI Race in 2025 is Tighter than Ever Before
14. Thoughtworks, Technology Radar 2025
15. World Economic Forum, What is Open Source AI and How Could DeepSeek Change the Industry? 2025
16. Arm AI Readiness Index, 2025
17. Linux Foundation commissioned by Meta, The Economic and Workforce Impacts of Open Source AI, 2025
18. Policy Connect, Skills in the Age of AI, 2025
19. Insight, State of the AI Agents Ecosystem, 2025
20. Forbes, AI 50 List, 2025
21. Demos, The Open Dividend, 2025
22. Centre of the Governance of AI, Open Sourcing Highly Capable models, 2023
23. QuantumBlack AI by McKinsey, Seizing the Agentic AI Advantage, 2025
24. European Commission, Generative AI Outlook, 2025
25. Daniotti, et al. Who Is Using AI To Code? 2025
26. CapGemini, AI In Action, 2025
27. RStreet, Mapping the Open Source AI Debate, 2025
28. Red Hat, Why Open Source is Critical to the Future of AI, 2025
29. Computer Weekly, AI Models Explained, 2025

### 5.2 The UK's AI Landscape

According to [Tech Nation](#), The UK stands as Europe's leading AI market, home to over 2,300 VC-backed AI firms with a combined valuation of \$230 billion. Investment in Q1 2025 hit \$1.03 billion, a 3-year high, reflecting strong investor confidence. Importantly, 76% of UK tech leaders report that AI positively influences business growth, with half citing product and service improvements. However, the sector is constrained by talent shortages and limited access to capital. [Policy Connect's report](#) underscores these issues, revealing that over half the UK workforce lacks essential digital skills, with 19 million people experiencing digital poverty. A key recommendation is embedding AI literacy into national education and investing in localised skills infrastructure.

With regards to policy, the UK's AI Opportunities Action Plan represents a key initiative to boost national capabilities. Tech Nation explains that the plan involves expanding public computing power, establishing a National Data Library, and launching AI Growth Zones to expedite infrastructure development.

The UK has an opportunity to lead in the domain of AI openness, as recognised in past OpenUK reports and a recent [Demos' report](#) proposing that "openness" be placed at the core of UK AI strategy, citing transparency, adaptability, and public trust as strategic advantages. Almost 15 years into the UK's "open source first" public sector policy it is perhaps time that this is taken as a given and put into actions not policies and



Demos go on to recommend openness underlies the Action Plan.

[MIT and Databricks' report](#) also reveals a growing focus on governance is also apparent, with over half of organisations supporting a unified model for managing AI and data governance.

### 5.3 Trends Shaping the AI Landscape

Several trends define the current development of AI based on current reporting.

[Forbes'](#) claims that autonomous agents capable of reasoning and decision making without human input are gaining traction, powered by advancements in reinforcement and deep learning. Similarly, multi-modal AI systems that integrate text, audio, and visual data are now standard across sectors like healthcare and customer service. This reflects a clear shift in the reporting discussion in 2025 to agentic AI.

Another key movement, noted by Stanford's [HAI report](#), is the growing focus on resource efficient models or small models that deliver competitive performance with reduced parameters. Microsoft's Phi-3-mini, for example, achieved high benchmark scores with just 3.8 billion parameters - compared to legacy models with over 500 billion. Deep Seek's R1 is an open model example and [KI Company's articles](#) states that this trend is epitomised by R1. A Chinese open model trained on a streamlined dataset using stripped-down Nvidia H800 chips sees training costs through distillation reduced to just \$5.6 million. KI Company places R1 in the top spot in their list of the top 5 best open model AI of 2025. DeepSeek R1 rivals GPT-4 and Llama 3 in a variety of tasks and has even surpassed ChatGPT in app store downloads. Its success marks a milestone for cost effective, high performance open model AI.

[Accenture's](#) notion of "Cognitive Digital Brains" - AI systems that embed company specific workflows and knowledge - further underscores AI's strategic role in reshaping enterprise operations. These systems coincide with what Accenture terms the "Binary Big Bang": a shift from static software to dynamic, agentic, intent driven systems.

### 5.4 AI Adoption & Maturity

MIT & Databricks' report shows that in 2022, organisations had a generally cautious approach to AI. 94% of organisations globally are adopting AI in some form, but only 14% aim for enterprise wide integration by 2025.

3 years on we see organisations increasingly integrating AI into core functions, with 85% of decision makers reporting progress in executing AI strategies, and 47% seeing positive Return On Investment according to [IBM](#).

However, some cautiousness remains as [Arm](#) reports that while 82% of organisations use AI tools, only 39% have clear, comprehensive AI strategies.

Despite this, [Snowflake's report](#) reflects a positive trend with enterprises moving beyond experimentation to focus on AI solutions that deliver tangible and measurable value. Arm's report also supports this, noting that investment in AI is robust, with 80% of organisations having dedicated AI budgets and 87% expecting these budgets to grow in the coming years. Snowflake argues that a critical component enabling this shift is observability - tools and processes that allow organisations to monitor and understand AI system behaviors. Improved observability, Snowflake argues, not only enhances operational reliability but also supports compliance and transparency efforts.

This surge in adoption is underpinned by significant cost reductions. Stanford's HAI reports that since 2022, inference costs have plummeted, with some tasks becoming 9 to 900 times cheaper annually, dramatically lowering barriers for broader AI deployment.

Arm's report notes operational efficiency as another primary driver for AI investment, cited by 80% of leaders as the central focus of their 2025 AI strategy.

However, as Accenture points out, building trust is now viewed as a cornerstone of successful AI integration, with 77% of executives surveyed by Accenture stating that the full benefits of AI will only be realised when users can trust digital systems.

## 5.5 The Public Sector

According to the recent CapGemini 2025 report, [Data Foundations for Government](#), across all regions surveyed, “9 in 10 public sector organisations” are to focus on agentic AI in the next 2-3 years, but data readiness is still a challenge. Globally, 75% of public sector organisations are either exploring or actively working on gen AI initiatives.

### Globally, the report sees:

**An execution gap:** While 87% of public sector organisations report having a national AI strategy, only 34% have implemented AI at scale.

**A data quality deficit:** 68% of respondents cite poor data quality as a major barrier to AI execution, while only 28% have robust data governance frameworks.

**Fragmented infrastructure:** Legacy systems and siloed data repositories hinder interoperability. Only 22% of agencies report using centralised data platforms.

**Correlation of Strong leadership and talent investment:** with successful AI implementation. Countries with appointed Chief Data Officers and data governance boards show higher rates of AI maturity.

**Regular audits conducted by only 41% of organisations:** The report warns that insufficient transparency and accountability could erode public trust in AI applications. The report included public sector organisations from the UK as part of the Europe region, which constituted 51% of the total organisations surveyed, the UK being 9% of these. Key findings from the UK suggest strong implementation in the UK and indications show that the UK is leading Europe in AI adoption.

For example, the UK government uses an AI-powered “Consult tool” to extract patterns and themes from thousands of public consultation responses, which helps policymakers make quicker, more informed decisions with fewer resources. Moreover, the UK’s NHS is implementing an AI tool nationwide that predicts a patient’s risk of falling with 97% accuracy. This tool is expected to prevent up to 2,000 falls and hospital admissions daily, reducing hospitalizations by up to 70% and saving approximately £2 billion annually in associated costs.

And the UK government’s AI-powered tool “Connect” aims to assist the Department for Energy Security and Net Zero (DESNZ) in managing grid connections for renewable energy projects. This tool is projected to reduce connection times from over 5 years to 6 months and could generate an economic benefit of £75 billion. All of this indicates a strong move towards transformation in the UK public sector.

At the same time, the proportion of public sector organisations in the UK that have scaled AI remains low, reflecting a broader challenge in deploying Gen AI into production in the public sector across both the US and Europe.

## 5.6 Generative and Agentic AI

GAI seems to have had an impact on the cautious approach to AI we saw in 2022, encouraging CIOs to democratise AI across business functions. GAI is now used for content creation, fraud detection, and data analytics according to MIT & Databricks and [CapGemini](#). It also enables unstructured data utilisation that was previously inaccessible.

Positively, Stanford HAI reveals that in 2024, 71% of businesses used GAI in at least 1 function while CGI reports a 1.7x Return On Investment from GAI, with deployment growing from 20% to 36% between 2024 and 2025.

However, as [Writer’s report](#) states, challenges remain. 71% of C-suite leaders say that AI is still developed in silos, and 68% report internal friction over implementation.

It’s clear from Insight, Accenture and [McKinsey’s](#) report, that agentic AI is expanding across domains including finance, development, and cloud operations. Accenture describes this moment as the “Binary Big Bang”,<sup>41</sup>

where AI is embedded into workflows as autonomous cognitive systems. These AI agents are used in supply chain, finance, and operations, increasing efficiency by up to 45% and reducing errors by 40%, according to CGI.

Still, [Insight](#) claims that these systems require clear guardrails and raise difficult integration and pricing questions. McKinsey proposes an “agentic mash” architecture to allow seamless integration of various AI agents, with features like traceability and vendor agnosticism.

Snowflake notes that agents work best when specialised but can deliver even greater value when designed to collaborate with other agents.

## 5.7 Security

Security remains one of the biggest barriers to AI adoption.

McKinsey, Arm, and [Perforce](#) all report concerns about compliance, lifecycle management, and system vulnerabilities. Perforce warns that 26% of firms still use unsupported software, tripling audit risks.

Arm identifies model extraction and data leakage as major threats, while [R Street](#) highlights specific vulnerabilities in frameworks like DeepSeek and Ray.

The [Centre for the Governance of AI](#) cautions that releasing powerful models without safeguards could lead to cyberattacks, misinformation, or bioengineering misuse. It proposes a spectrum of disclosure options, such as gated access, auditor-controlled releases, and hybrid systems to mitigate risks. HAI’s report also notes a 54% increase in AI related incidents in 2024, prompting new safety benchmarks like HELM and AIR-Bench.

## 5.8 Scaling Challenges

Despite significant progress, certain obstacles persist. The Arm AI Readiness Report reveals that while 82% of global business leaders report using AI applications, only 39% have a clearly defined, comprehensive AI strategy, potentially limiting effectiveness and scalability.

MIT & Databricks on the other hand, note that data remains the primary challenge for 72% of CIOs, prompting shifts toward unified analytics platforms like data lakehouses.

Meanwhile, [Perforce](#) claims that technical complexity undermines confidence, with nearly half of organisations struggling to manage big data stacks using software like Hadoop and Kafka. Compounding this is a widespread shortage of skilled personnel: over 75% cite inexperience or staffing gaps as barriers to deploying open source data technologies. Additionally, according to Arm’s report only 29% of organisations can automatically scale compute resources to meet AI demands, and 34% report a lack of AI talent.

In the UK context, Tech Nation’s report claims that the AI sector’s dynamism is offset by a “brain drain,” as firms like Wayve and DeepMind seek growth opportunities abroad.

Limited capital access and talent availability are major hurdles for domestic AI startups in the UK, threatening long-term competitiveness.

## 5.9 The Rise of AI Openness and Public Good AI

Openness in AI has become central to AI innovation and deployment.

As McKinsey’s report shows over 50% of organisations use AI technologies that are open across the AI stack, from data management to user-facing applications.

Additionally, Stanford HAI’s report shows that a significant 65.7% of foundation models released in 2023 were open, up from 44.4% in 2022.

These trends highlight a key shift, noted by IBM's report: from isolated model development to building comprehensive, open AI systems that integrate models, tools, and workflows.

This adoption is strategic. Organisations leveraging openness report lower implementation costs (60%) and higher developer satisfaction (66%), according to McKinsey's report.

Additionally, IBM reports that companies using open source AI see higher Return On Investment (51% vs. 41% for closed AI users).

Furthermore, those who view AI as a competitive differentiator are 40% more likely to adopt open solutions. Transparency and customisation-particularly crucial in regulated sectors-are major draws.

UK organisations follow global patterns.

Openness in AI allows them to fine-tune LLMs using internal data while maintaining IP control, as noted by MIT & Databricks' report. Hybrid approaches are increasingly common, blending proprietary tools with open source components to achieve scalability and customisation, seen by both MIT & Databricks, and McKinsey.

## 5.10 Challenges of AI Openness and its commercial adoption

Despite its benefits, AI openness is not without complications.

Security and compliance remain top concerns, cited by 56% of organisations, along with uncertainty about long-term support in McKinsey's report.

Organisations often opt for proprietary models to gain tighter control over risk and system governance. Furthermore, according to Perforce, failure to maintain open source(e.g., relying on end-of-life software) can jeopardise compliance efforts.

The rise of AI adoption underscores the need for robust governance of open projects.

Encouragingly, in 2025, 59% of organisations now scan open source software for vulnerabilities, and 35% have formal security policies in place, according to [Perforce](#). Active participation in open source communities is also growing, with 38% contributing to projects or foundations. Trends which we may expect to see replicated in AI that is open.

## 5.11 Future Outlook for Open Source AI

Looking ahead, [McKinsey's report](#) notes that 76% of organisations expect to increase their reliance on AI that is open.

Perforce finds that investment in open source software technologies, cloud-native infrastructure, and containerisation continues to rise, as companies seek scalable, cost-efficient AI solutions. The transition from models to full-stack systems-complete with classifiers, data pipelines, and orchestration tools-signals a new era of holistic AI openness development, according to IBM. A trend that may also be reflected in AI.

The open ecosystem around AI is also driving equity in AI innovation.

By lowering entry barriers, openness enables startups and smaller firms to build competitive AI capabilities without relying on expensive proprietary APIs, noted by [Forbes](#).

This democratisation could be key to sustaining the UK's leadership in AI, provided ongoing investment, skills development, and policy alignment are maintained.

## 6. Conclusion

**Dr Jennifer Barth,**  
Founder Symmetry and Research Director OpenUK



This report reflects a defining moment in the AI trajectory globally and in the UK. Against a backdrop of accelerating global adoption, technical breakthroughs, and rising policy tensions, the report presents a clear picture: the UK is uniquely positioned to lead on AI openness and public-interest technology - but only if it can address systemic challenges and harness its collaborative ecosystem.

Across this report - in the policy reflections, the data analysis, the literature review and the contributions from key voices - we see the outlines of what such a strategy backed by clear practical steps, could be. We also see, with increasing clarity, what happens if the UK doesn't act fast enough.

From a policy perspective, it's clear that the global landscape is fracturing.

In the US, nearly 2,000 AI-related bills have surfaced, while the Biden Executive Order has been withdrawn. As Harvard Fellow Ben Brooks highlights in his Fireside Chat, this legislative noise masks a deeper uncertainty - that no one is quite sure how governance at scale should now work.

In the EU, enforcement of the AI Act remains in question, even as efforts to develop a voluntary Code of Conduct move forward. In contrast, the UK's decision to initiate the AI Safety Summits in 2023, and its continued presence at this year's Paris AI Action Summit and in building the next summit in India, suggests an ambition to shape global norms.

The Paris Summit itself brought with it the launch of two important initiatives - ROOST and Current AI - both placing openness and public interest at the centre of future AI development. Martin Tisné's reflection on Current AI brings a reminder that transparency, shared tooling, and public oversight are not luxuries; they are preconditions for trust in this new infrastructure.

Turning to the data, we see both strength and warning signs.

In GitHub's repositories, the UK remains in a solid third position globally, in AI that is open, with 135 repositories exceeding 1k+ stars - a 19% increase from 2024.

But that growth, while respectable, is being outpaced.

India grew by 38% in the same period, France by 26%, the US by 23%. China, often treated as a slow but steady comparator, saw a 16% increase. It's not that the UK is slipping - but it is no longer growing fast enough to comfortably hold its place.

The numbers remind us that leadership in AI, especially in open ecosystems, is not measured by intention or press statements and promises, but by participation and output.

The story is similar when we look at the agentic AI space - the shift from passive language models to tools that can reason, plan, and act. The US now hosts 114 notable agentic AI repositories. China, 36. The UK? Just eight.

Early engagement, yes, but not yet at the scale needed to be a thought leader, let alone a market one. The UK's leading AutoGPT remains an open submarine in the UK's AI economy.

Despite this, there are signs of movement. Scarf's data suggests that around 20,000 UK companies interacted with open source AI in the past year, with 8,500 active in just the past month - triple the figure from



a year ago. That kind of acceleration, particularly in commercial adoption, shows that the appetite is there. But the infrastructure to support it - in compute, in skills, in national coordination - is not yet matching pace. The literature review was perhaps the most revealing in painting a complete picture of the landscape. With over 30 reports reviewed in a very extensive review, the key patterns are hard to ignore.

**The UK leads Europe in AI valuation:** Home to more than 2,300 VC-backed firms worth a combined \$230 billion. Tech Nation's framing of this as a continental advantage is backed by numbers but also challenged by structural fragilities.

**The Digital skills gaps remain widespread:** 19m people in the UK experience some form of digital poverty. Despite ambitious programmes, including the Prime Minister's AI Skills Boost, the workforce is not yet AI-ready.

**GAI is no longer theoretical:** 71% of businesses now use it in at least one function. CIOs have shifted from cautious experimentation to scaled deployment. The enthusiasm is real, but so are the barriers - data infrastructure, system integration, internal resistance, and above all, trust. Reports from IBM, Arm, McKinsey and others show that while adoption is growing, confidence in safe, accountable systems is not keeping pace. That gap matters.

The thought leadership in this report adds texture to this strategic position. Toran Bruce Richards, founder of AutoGPT, makes a strong case for protecting the autonomy and capability of agentic systems - and warns that security measures that restrict their function too severely risk removing their value altogether. Sal Kimmich and Matt Barker tackle governance and identity, proposing practical structures for trust. Isabelle Mouray and Richard Steel invite us to think bigger still, about the convergence of AI with quantum systems. And Andrew Martin, through the lens of the DSIT Hyperscaler hackathon, shows what it looks like when public and private sectors build together in a genuinely open way.

This report, in its depth and breadth, leaves little doubt that we are at a crossroads.

The UK has the opportunity - perhaps even the obligation - to shape a distinctive model for AI development rooted in public good. Not as a counterweight to the US or China, but as an alternative: one that prioritises transparency, inclusion, and long-term value over short-term optimisation.

The UK's edge is openness. In building ecosystems that are transparent, interoperable, and geared toward public value. This isn't just a philosophical stance; it's a competitive one.

In a world increasingly wary of AI's opaque, centralised power, systems built in the open - with public infrastructure, shared safety tools, and collaborative development - will carry more legitimacy, not less.

If the UK is to lead on Public Good AI, it needs to move decisively - to scale its investments in open infrastructure, to align its regulation with its values, to accelerate skills development in a way that is inclusive and localised, and above all, to embed openness at the core of its national AI agenda.

Public Good AI is not a side track - it is the track.

This report, thanks to the contributions of researchers, policymakers, engineers and strategists alike, makes clear what it will take to stay on it.

## 7. Formalities

### 7.1 Contributors

#### **Amanda Brock, CEO, OpenUK**

OpenUK CEO, Amanda's built one of open source's most recognised and impactful organisations. Executive Producer of State of Open Con (2023- 2025), Amanda's a globally sought-after keynote speaker. A lawyer with 25 years' experience, 5 as GC of Canonical, she's been instrumental in shaping open source's legal frameworks, as she was internet law during the early 2000's. Regularly contributing to tech press, she edited 'Open Source: Law, Policy and Practice', (2022). Recognition: Computer Weekly 50 Most Influential Women in UK Tech (2023, 2024); Computing IT Leaders 100 (2023, 2024); Lifetime Achievement Award WIPL (2022); Women Who Will Changemaker (2023); INvolve Heroes (2022, 2023); Novi Awards (2024) and Ambassador, Open Charge Alliance. Advisory Appointments: UK Cabinet Office Open Standards Board; UKRI Digital Research Infrastructure; UKRI Exascale; KDE; commercial boards – Mimoto, Scarf, FerretDB and Space Aye; and is Fellow Open Forum Academy; Distinguished Fellow Rust Foundation; and European Representative, OIN.

#### **Andrew Martin Co-Founder and CEO, ControlPlane**

Andrew has an incisive security engineering ethos gained building and destroying high traffic web applications. Proficient in systems development, testing, and operations, he is at his happiest profiling and securing every tier of a cloud native system, and has battle hardened experience delivering containerised solutions to enterprise and government.

#### **Ben Brooks, Fellow, Berkman Klein Center, Harvard**

Ben is a Fellow at the Berkman Klein Center, Harvard, where he scrutinises the regulatory and legislative response to AI models. He served most recently as Head of Public Policy for Stability AI, custodian of Stable Diffusion. Ben has testified on AI regulation before the US Congress and UK Parliament, and engaged policymakers around the world to protect open innovation in future regulation. Previously, Ben championed the safe, open, and durable regulation of emerging technologies, including for drone delivery at Alphabet, ridesharing at Uber, and digital assets at Coinbase. He has worked with authorities on the ground in over 25 countries as they navigate complex reforms in high-stakes or permission-based domains.

#### **Isabelle Mauny, Co-founder and field CTO, WSO2**

Isabelle Mauny is the European Field CTO at WSO2, where she leads technical evangelism and strategic technology initiatives for enterprise clients. With extensive experience working alongside Fortune 500 companies, Mauny specialises in API security architecture and AI security frameworks, helping organizations navigate the evolving landscape of digital transformation and emerging threats. Her expertise spans quantum cryptography applications in enterprise security, positioning her at the forefront of next-generation cybersecurity solutions. As a recognised thought leader, she regularly speaks on the intersection of API governance and security, and more recently quantum-resistant cryptographic implementations in large-scale enterprise environments.

#### **Dr Jennifer Barth, Founder Symmetry and Research Director OpenUK**

Jenn has more than 15 years of experience leading independent research on the intersections of emerging technologies and socioeconomic change. She provides companies with independent thought leadership and media engagement opportunities on global issues impacting and shaping our current and future technical-social lives. Her work spans the digital through to social and economic change. She has looked at sustainability, workforce skills and organisational competitiveness strategies through and beyond the pandemic with Microsoft and many other big and small organisations and works as the Chief Research Office researching the role of open source software and its potential to fuel the circular economy with OpenUK. She has experience working on the human impact of artificial intelligence (AI) through fieldwork experiments with IBM Watson, Microsoft and other providers. She is skilled at blending research methods and working with people to bring to life the stories behind numbers. Dr Barth earned her DPhil in Geography from the University of Oxford.

### **Martin Tisne, Chair, Current AI Interim Board**

AI Collaborative, an initiative of The Omidyar Group created to help regulate artificial intelligence based on democratic values and principles and ensure the public has a voice in that regulation. Martin brings over 15 years of investment and leadership experience to his role, including advising several heads of state on AI policy, serving as a board member of the Partnership on AI, and helping establish two multi-stakeholder initiatives and three NGOs. He founded the Open Government Partnership (OGP) alongside the Obama White House and helped OGP grow to a 70+ country initiative. He also initiated the International Open Data Charter, the G7 Open Data Charter, and the G20's commitment to open data principles. Additionally, Martin founded and led the Transparency and Accountability Initiative, a donor collaborative bringing together the world's largest open government funders, co-founded Publish What You Fund, a global campaign for foreign aid transparency, and co-founded Integrity Watch Afghanistan, the country's leading anti-corruption NGO.

### **Matt Barker, VP & Global Head, Workload Identity Architecture, CyberArk**

Matt was the co-founder & CEO of Jetstack, a Kubernetes company he started in 2015 and bootstrapped before being acquired by Venafi in 2020. Venafi was subsequently acquired by private equity company Thoma Bravo, and then the cybersecurity company CyberArk. Jetstack is best known for its open source project 'cert-manager', which is downloaded millions of times a day to secure cloud native infrastructure. Cert-manager was donated to the CNCF in 2020 and graduated in 2024. Matt has played a number of product and leadership roles since being acquired, and now leads workload identity at CyberArk. In January 2021 Matt was awarded as a top 100 Open Source Influencer by OpenUK.

### **Richard Steel, CIO, Technology Advisory Services**

Richard Steele is a seasoned C-suite technology executive and freelance CTO, renowned for delivering transformative IT and digital solutions across telecoms, transport, agriculture, and startup sectors. With over two decades of experience, Richard has led the design and management of large-scale cloud platforms, including the national telecoms switching hub for the UK, and has a proven track record in enterprise IT strategy, vendor management, and technical leadership. Richard is also deeply engaged in emerging technologies, with a particular interest in quantum computing and post-quantum cryptography. He actively pursues professional development in advanced machine learning and quantum security, and is committed to exploring how quantum technologies will shape the future of secure digital infrastructure.

### **Sal Kimmich, Technical Director, GadflyAI**

Sal's work is centered on advancing secure computing practices and pioneering privacy-enhancing technologies to safeguard sensitive data. At the Confidential Computing Consortium (CCC) within the Linux Foundation, they play a pivotal role in the development and implementation of Open Source projects that establish Trusted Execution Environments (TEEs). These environments are designed to provide secure areas within hardware that protect data in use, addressing critical security challenges across various industries. My contributions are particularly significant in sectors such as finance, healthcare, and emerging fields like computational justice, where data confidentiality and integrity are paramount. In my previous roles, they've led initiatives to build and nurture ISO 27001 compliant cultures within organizations. This includes establishing robust security-first standards that are integrated into all engineering practices, ensuring that security considerations are embedded from the ground up.

### **Toran Bruce Richards, Founder, AutoGPT**

Toran Bruce Richards, a prominent figure in AI innovation from the UK, is chiefly known for creating Auto-GPT, an Autonomous AI System. Launched Open Source in March 2023, Auto-GPT represents a lead in the application of advanced large language models (LLMs), capable of independently managing complex, multi-step tasks without continuous human input. The scale of the project and its community impact is significant, with Auto-GPT's GitHub Repository achieving the #23 rank globally, amassing over 157k stars. Additionally, it has fostered a vibrant community of over 50k AI enthusiasts on Discord, marking its position as a prominent and influential project in the 'open source AI' landscape.

## 7.2 About the Creators of this Report

### OpenUK

OpenUK is the unique open tech industry body for the business of open technology in the UK. It spans the opens – software, hardware, data, standards and AI and is the convening point for the UK’s business, academic and contributing communities across open tech. Our work supports the UK’s journey to become “The State of Open”. Our organisation is run with the support of our volunteer community and their leadership in the tradition of open source delivering on 3 pillars: community, legal and policy and learning. Our Community is recognised through our world-leading open tech recognition programme including the OpenUK Awards (the Oscars of Open Source) now in their 5th year, New Year’s Honours Lists and Ambassador Scheme.

OpenUK undertakes research and reporting both on its own account through its “State of Open Reports” and on a commissioned basis for third parties. Case studies, Thought Leadership, Surveys and desk-based research are included in our reporting which pushes the envelope and leads the way. Our Research and Reporting Show and Tell events coalesce the global open source research communities digitally to regularly update and share research practices and topics. OpenUK’s new OpenUK Fellows Network for postgraduate researchers is launching in 2024 to encourage more academic research across the opens. The community’s strength is channelled to enable a cohesive voice that responds to legislative proposals and sets policy. We have set the agenda in policy matters across openness in the UK and beyond. OpenUK’s Policy work leads the conversations around open source licensing and commercialisation, AI openness and cloud computing and other key topics across open source, as they emerge. Engagement with UK policy makers is supported by a volunteer Policy Advisory Board and by experts across our volunteer Advisory Boards and the open source communities. Our Advisory Boards span AI, Communications Tech, Data, Finance, Hardware, Healthcare, Security, Software, Space, Sustainability and Quantum Computing. We are able to provide industry experts across the opens for speaking engagements, consultancy and advisory boards.

OpenUK is the second organisation established anywhere in the world with open source policy as its purpose, our approach is holistic to and representative of the entire open ecosystem. OpenUK undertakes a broad range of activities in support of its policy work and is a day one member of GaiaX and UK’s GaiaX Hub Coordinator, hosted one of the biggest tech events at COP26, and was the first organisation in open tech to put a Sustainability Policy and Chief Sustainability Officer in place. Skills and Learning form our third pillar and our Learning work has spanned initiatives for children including our award winning Kids Camps which teach coding, open source and sustainability in a real world context; and exploring the business of open source through our Founder training. We have shared several hundred hours of digital training. Our ambitions include a UK apprenticeship module and adding open source to the UK curriculum.

The State of Open Con has become one of the world’s leading open source conferences since its inception by OpenUK in 2023. In 2025 we expect to host 1000 people across 8 tracks and plenary sessions, with at least 50 partners in our delegate experience space and over 200 speakers. Our small events team deliver to the highest standards a series of unique events through the year and our community organise UK-wide OpenUK meet-ups. Contact OpenUK [admin@openuk.uk](mailto:admin@openuk.uk)

### Symmetry

Symmetry looks beyond the surface and behind the curtain of the fundamental innovations and trends shaping our society, markets, culture, and values. We are academics and researchers looking at the intersections of emerging technology and socioeconomic impact, producing independent research for thought leadership and business solutions. Symmetry’s mission is to share and grow knowledge about the interaction of technology and everyday lives. We want to understand the past, present, and future of human interaction with emerging technologies and socioeconomic changes-from behaviour to context, nature to nurture, origin to experiences-helping our clients engage their clients and public imagination.

### 7.3 Acknowledgements

The research was led by Dr Jennifer Barth, Founder and Research Director at Symmetry and OpenUK's Chief Research Officer in partnership with Amanda Brock, CEO OpenUK. Thank you to Runa Capital, GitHub and Scarf for sharing data. Thank you to our team of economists, psychologists, data scientists and social scientists to all who contributed, and in particular to Lucy Anderson, Zin Nwe Zaw Lwin and Elefteria Kokkinia. We are grateful to the individuals who participated and provided us with case studies, fireside chats and thought leadership to bring the key issues to life.

We are grateful to Scarf, Runa Capital, Hugging Face and GitHub for their data.

### 7.4 References

All sources used in establishing our thinking on this report, whilst only some are quoted:

- Arm (2025). Arm AI Readiness Index <https://armkeil.blob.core.windows.net/developer/Files/pdf/report/arm-ai-readiness-index-part2.pdf>
- AWS (2025). AWS Security Token Service. <https://docs.aws.amazon.com/STS/latest/APIReference/welcome.html>
- Accenture (2025). Technology Trends 2025. <https://www.accenture.com/gb-en/insights/technology/technology-trends-2025>
- Bank of England (2025). Financial Stability in Focus. <https://www.bankofengland.co.uk/financial-stability-in-focus/2025/april-2025>
- Belcak, et al. (2025). Small Language Models are the Future of Agentic AI. <https://research.nvidia.com/labs/lpr/slm-agents/>
- BIS (2024). Project Leap. [https://www.bis.org/about/bisih/topics/cyber\\_security/leap.htm](https://www.bis.org/about/bisih/topics/cyber_security/leap.htm)
- Capgemini (2025). Data Foundations for Government. [https://www.capgemini.com/wp-content/uploads/2025/05/Capgemini-Research-Institute-report\\_Data-foundations-for-government\\_From-AI-ambition-to-execution-3.pdf](https://www.capgemini.com/wp-content/uploads/2025/05/Capgemini-Research-Institute-report_Data-foundations-for-government_From-AI-ambition-to-execution-3.pdf)
- Capgemini (2025). AI in Action. <https://www.capgemini.com/wp-content/uploads/2025/06/Final-Web-Version-Report-AI-in-Business-Operations.pdf>
- Centre for the Governance of AI (2025). <https://www.governance.ai/research>
- Cloud Security Alliance (2025). <https://cloudsecurityalliance.org/blog/2025/05/28/mcp-oauth-2-1-pkce-and-the-future-of-ai-authorization>
- Computer Weekly (2025). Execs Shy Away from Open Models and Open Source AI <https://www.computer-weekly.com/news/366626106/Execs-shy-away-from-open-models-and-open-source-AI>
- Congress (2022). <https://www.congress.gov/bill/117th-congress/house-bill/7535>
- Current AI (2025) <https://www.currentai.org>
- Current AI (2025). Building Public Interest AI. <https://www.currentai.org/latest-updates/building-public-interest-ai---current-ai-next-chapter>
- Current AI (2025). <https://www.currentai.org/#Current-AI-Initiatives>
- Current AI (2025). Launch Press Release. <https://www.currentai.org/latest-updates/launchpressrelease>
- Daniotti, et al. (2025). Who Is Using AI to Code?. <https://arxiv.org/pdf/2506.08945>
- Datatracker (2025). Draft-ietf-oauth-v2-1-12. <https://datatracker.ietf.org/doc/html/draft-ietf-oauth-v2-1-12#name-authorization-code-grant>
- Debenedetti, et al. (2024). AgentDojo. <https://arxiv.org/abs/2406.13352>
- Demos (2025). The Open Dividend. <https://demos.co.uk/research/the-open-dividend-building-an-ai-openness-strategy-to-unlock-the-uks-ai-potential/>
- Digifabster (2025). AI-Powered Quote Agent. <https://digifabster.com/product/features/ai-quote-agent/>
- European Commission (2025). European Approach to Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>
- European Commission (2025). Generative AI Outlook Report. [https://media.licdn.com/dms/document/media/v2/D4D1FAQFOfR-cDXysQA/feedshare-document-pdf-analyzed/B4DZdp\\_3mxH4Ac-/0/1749830056779?e=1750896000&v=beta&t=EyBb\\_safOP5bf8ZmHBUf\\_5sM1a8LBT20\\_susMNDJnEI](https://media.licdn.com/dms/document/media/v2/D4D1FAQFOfR-cDXysQA/feedshare-document-pdf-analyzed/B4DZdp_3mxH4Ac-/0/1749830056779?e=1750896000&v=beta&t=EyBb_safOP5bf8ZmHBUf_5sM1a8LBT20_susMNDJnEI)
- European Parliament (2024). The United Kingdom and Artificial Intelligence. [https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/762285/EPRS\\_ATA\(2024\)762285\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/762285/EPRS_ATA(2024)762285_EN.pdf)
- Forbes (2025). The 5 AI Trends in 2025. <https://www.forbes.com/sites/solrashidi/2025/02/28/the-5-ai-trends-in-2025-agents-open-source-and-multi-model/>
- Forbes (2025). 2025 AI 50 List. <https://www.forbes.com/consent/ketch/?toURL=https://www.forbes.com/lists/ai50/>



Gartner (2025). TSP 2025 Trends. <https://www.gartner.com/en/documents/6202687>

GitHub (2025). Agent Smith. <https://sail-sg.github.io/Agent-Smith/>

Glide (2025). Top 5 Source for AI Stats in 2025. <https://www.glideapps.com/blog/ai-stats-2025>

Global Risk Institute (2024). Quantum Threat Timeline Report. <https://info.quintessencelabs.com/hubfs/PDFs/Global-Risk-Institute-Quantum-Threat-Timeline-Report-2024.pdf>

GOV.UK (2025). Cyber Security Risks to Artificial Intelligence. <https://www.gov.uk/government/publications/research-on-the-cyber-security-of-ai/cyber-security-risks-to-artificial-intelligence>

GOV.UK (2025). International AI Safety Report 2025. <https://www.gov.uk/government/publications/international-ai-safety-report-2025/international-ai-safety-report-2025>

HAI (2025). HAI AI Index Report 2025. [https://hai-production.s3.amazonaws.com/files/hai\\_ai\\_index\\_report\\_2025.pdf%20https://hai.stanford.edu/ai-index](https://hai-production.s3.amazonaws.com/files/hai_ai_index_report_2025.pdf%20https://hai.stanford.edu/ai-index)

Heise Online (2025). "Cryptocalypse". <https://www.heise.de/en/news/Cryptocalypse-EU-demands-quantum-safe-encryption-partly-by-2030-10456642.html>

Hiscox (2023). Hiscox and Google Cloud Collaborate on AI in Lead Underwriting from the London Market. [https://www.hiscoxgroup.com/news/press-releases/2023/12-12-23?utm\\_source=chatgpt.com](https://www.hiscoxgroup.com/news/press-releases/2023/12-12-23?utm_source=chatgpt.com)

Huang, et al. (2025). A Novel Zero-Trust Identify Framework for Agentic AI. <https://arxiv.org/abs/2505.19301>

IBM (2025). 2025 Open AI Trends. <https://www.ibm.com/think/news/2025-open-ai-trends>

ICO (2025). UK GDPR Guidance and Resources. <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/>

IEEE (2025). AI Index 2025. <https://spectrum.ieee.org/ai-index-2025>

Insight (2025). The State of the AI Agents Ecosystem. <https://www.insightpartners.com/ideas/state-of-the-ai-agent-ecosystem-use-cases-and-learning-for-technology-builders-and-buyers/>

ISACA (2025). Post Quantum Cryptography. <https://www.isaca.org/resources/news-and-trends/industry-news/2025/post-quantum-cryptography-a-call-to-action>

KI Company (2025). The 5 Best Open Source AI Models in 2025. <https://www.ki-company.ai/en/blog-beit-raege/the-5-best-open-source-ai-models-in-2025>

Lemly, et al. (2025). The Mirage of Artificial Intelligence Terms of Use Restrictions. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5049562](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5049562)

LinkedIn (2025). AI and the Global Economy. <https://economicgraph.linkedin.com/research/ai-skills-resources>

LinkedIn (2025). AI Skills Trends in the UK. <https://economicgraph.linkedin.com/content/dam/me/economic-graph/en-us/PDF/ai-skills-trends-uk.pdf>

Linux Foundation (2025). The Economic and Workforce Impacts of Open Source AI [https://www.linuxfoundation.org/hubfs/LF%20Research/lfr\\_market\\_impact\\_052025a.pdf?hsLang=en](https://www.linuxfoundation.org/hubfs/LF%20Research/lfr_market_impact_052025a.pdf?hsLang=en)

Linux Foundation Projects (2025). Post Quantum Cryptography Alliance. <https://pqca.org>

McKinsey (2025). Open Source Technology in the Age of AI. [https://www.mckinsey.com/~media/mckinsey/business%20functions/quantumblack/our%20insights/open%20source%20technology%20in%20the%20age%20of%20ai/open-source-technology-in-the-age-of-ai\\_final.pdf](https://www.mckinsey.com/~media/mckinsey/business%20functions/quantumblack/our%20insights/open%20source%20technology%20in%20the%20age%20of%20ai/open-source-technology-in-the-age-of-ai_final.pdf)

McKinsey (2025). Seizing the Agentic AI Advantage. [https://media.licdn.com/dms/document/media/v2/D561FAQGXB1aHZG-3fg/feedshare-document-pdf-analyzed/B56ZdvemaQHUAc-/0/1749921987019?e=1750896000&v=beta&t=o0Mc7CTSZlUrPJl\\_ioUsMTO5WNQRA9oftZJuZPt3QCg](https://media.licdn.com/dms/document/media/v2/D561FAQGXB1aHZG-3fg/feedshare-document-pdf-analyzed/B56ZdvemaQHUAc-/0/1749921987019?e=1750896000&v=beta&t=o0Mc7CTSZlUrPJl_ioUsMTO5WNQRA9oftZJuZPt3QCg)

MIT & Databricks (2025). The Great Acceleration. <https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report>

Model Context Protocol (2025). Architecture. <https://modelcontextprotocol.io/specification/2025-03-26/architecture>

Model Context Protocol (2025). Authorization. <https://modelcontextprotocol.io/specification/2025-03-26/basic/authorization#oauth-grant-types>

National Cyber Security Centre (2025). Timelines for Migration to Post-Quantum Cryptography. <https://www.ncsc.gov.uk/guidance/pqc-migration-timelines>

National Telecommunications and Information Administration (2024). Dual-Use Foundation Models with Widely Available Models with Widely Available Model Weights Report. <https://www.ntia.gov/programs-and-initiatives/artificial-intelligence/open-model-weights-report>

Nature (2025). AI Race in 2025 is Tighter than Ever Before. <https://www.nature.com/articles/d41586-025-01033-y>

New York Times (2025). Some AI Companies Face a New Accusation: 'Open Washing'. <https://www.nytimes.com/2024/05/17/business/what-is-openwashing-ai.html>

NIST (2025). AI Risk Management Framework. <https://www.nist.gov/itl/ai-risk-management-framework>

NIST (2025). Post-Quantum Cryptography. <https://csrc.nist.gov/projects/post-quantum-cryptography>

NTIA (2024). Dual-Use Foundation Models with Widely Available Model Weights Report. <https://www.ntia.gov/programs-and-initiatives/artificial-intelligence/open-model-weights-report>

Open Policy Agent (2025). <https://www.openpolicyagent.org>  
Open Source Initiate (2025). Key Insights from the 2025 State of Open Source Report. <https://opensource.org/blog/key-insights-from-the-2025-state-of-open-source-report>  
OpenUK (2025). AI Action Summit. <https://openuk.uk/wp-content/uploads/2025/02/AI-Report-FINAL.pdf>  
Perforce OpenLogic (2025). 2025 State of Open Source Report. <https://www.openlogic.com/resources/state-of-open-source-report>  
Policy Connect (2025). Skills in the Age of AI. <https://www.policyconnect.org.uk/research/skills-age-ai>  
Post-Quantum Cryptography Alliance (2025)  
Quantum Insider (2025). UK Sets Timeline, Road Map for Post-Quantum Cryptography Migration. <https://thequantuminsider.com/2025/03/20/uk-sets-timeline-road-map-for-post-quantum-cryptography-migration/>  
Red Hat (2025). Why Open Source is Critical to the Future of AI. <https://www.redhat.com/en/blog/why-open-source-critical-future-ai#:~:text=No%20single%20vendor%20can%20provide,collaboration%20across%20projects%20and%20industries.>  
Robust Open Online Safety Tools (2025). <https://roost.tools>  
RStreet (2025). Mapping the Open Source AI Debate. <https://www.rstreet.org/research/mapping-the-open-source-ai-debate-cybersecurity-implications-and-policy-priorities/#the-ldquo-open-rdquo-approach-to-aidevelopment>  
SailPoint (2025). SailPoint Research Highlights Rapid AI Agent Adoption, Driving Urgent Need for Evolved Security. <https://www.sailpoint.com/press-releases/sailpoint-ai-agent-adoption-report>  
Snowflake (2025). AI and Data Predictions 2025. <https://www.snowflake.com/en/lp/snowflake-ai-data-predictions/>  
SPIFFE (2025). <https://spiffe.io>  
TechNation (2025). UK AI Sector Spotlight 2025. <https://technation.io/uk-ai/>  
Thoughtworks (2025). Technology Radar. <https://www.thoughtworks.com/en-gb/radar>  
Time (2024). Yann LeCun On How An Open Source Approach Could Shape AI. <https://time.com/6691705/time100-impact-awards-yann-lecun/>  
WEF (2025). What is Open Source AI and How Could DeepSeek Change the Industry. <https://www.weforum.org/stories/2025/02/open-source-ai-innovation-deepseek/>  
Writer (2025). Enterprise AI Adoption Survey. <https://go.writer.com/enterprise-ai-adoption-survey>  
Wso2 (2025). Post-Quantum Hybrid Encryption with Ballerina. <https://wso2.com/library/blogs/post-quantum-hybrid-encryption-ballerina/>  
Wso2 (2025). Towards Quantum-Safe Applications. <https://wso2.com/library/blogs/towards-quantum-safe-applications/>  
Wso2 (2025). Quantum Threats Are Closer Than You Think. <https://wso2.com/events/webinar/quantum-threats-are-closer-than-you-think-act-now-to-stay-secure/>  
1Password (2025). Secure Agentic AI. <https://1password.com/solutions/agentic-ai>

## 7.5 Sponsors

OpenUK is grateful to its many sponsors and supporters without whose support its work including its report would not have been possible in particular, its general sponsors Arm, GitHub, Google, Linaro, Microsoft, Red Hat and SUSE. We are also grateful to WSO2 for their support of our launch event.

Sign up to OpenUK's newsletter



Find out more about our annual  
conference State of Open Con 26







[openuk.uk](https://openuk.uk)

X: [@openuk\\_uk](https://twitter.com/openuk_uk)

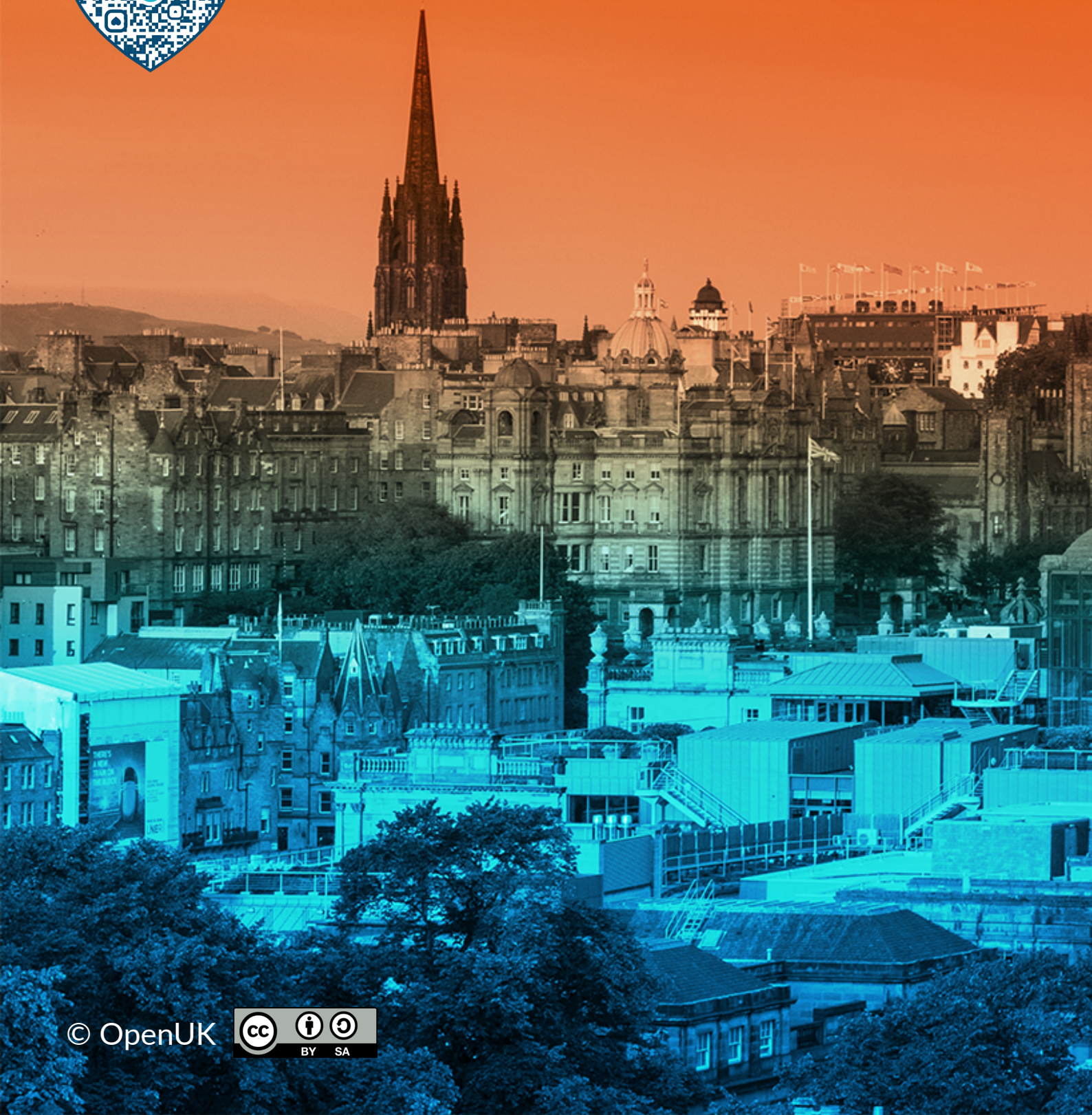
Mastodon: [@openuk@hachyderm.io](https://hachyderm.io/@openuk)

Bluesky: [@openuk.bsky.social](https://bsky.social/@openuk)

LinkedIn: [www.linkedin.com/company/openuktechnology](https://www.linkedin.com/company/openuktechnology)

Slack: [openuk.slack.com](https://openuk.slack.com)

GitHub: [@OpenUK](https://github.com/OpenUK)



© OpenUK

