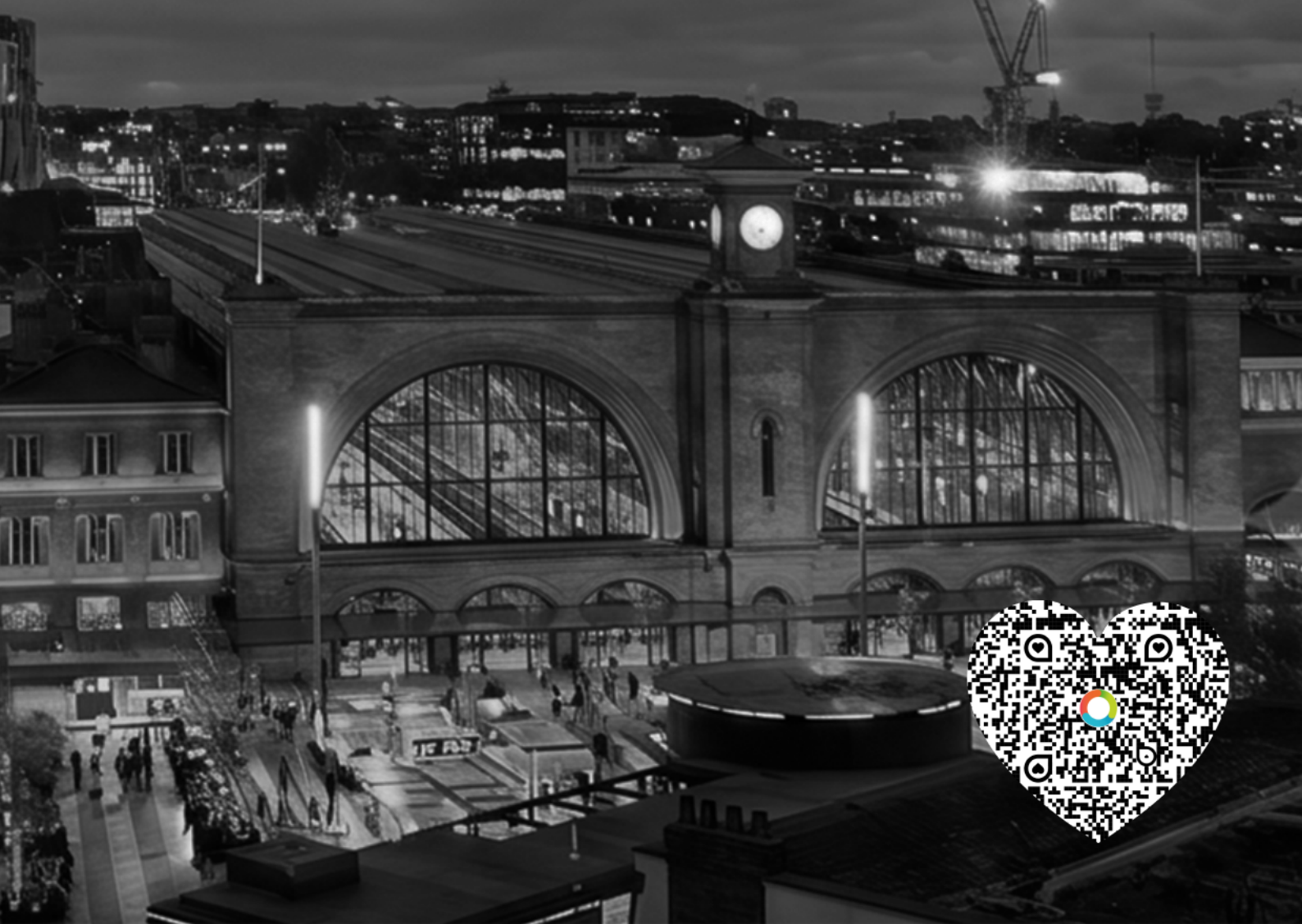
Open:UK® Open:HQ®

AI Openness: 3 Years on from Llama 2

Fireside Chat: Open foundation models enabling a map of the planet

Professor Anil Madhavapeddy
Professor of Planetary Computing,
University of Cambridge



Index

Index

2.9.4 Fireside Chat: Open foundation models enabling a map of the planet

Professor Anil Madhavapeddy
Professor of Planetary Computing, University of Cambridge

Professor Anil Madhavapeddy has shaped open source systems infrastructure for over two decades. He was part of the original Cambridge team that built the Xen hypervisor in 2002 - the first open source type-1 hypervisor, which went on to underpin the modern cloud - and has since worked on containerisation at Docker, the OCaml language and formal verification, and unikernels. Having trained in his early days at NASA, it is perhaps unsurprising that his current focus is planetary computing. He is a key member of the team at TESSERA, an open foundation model for the Earth's terrestrial surface that releases its embeddings, weights and code openly for anyone to build on.

Can you explain what TESSERA is, how it is used, and why open source is so important to it?

TESSERA is an open foundation model for the entire terrestrial surface of the planet. It takes a year of multi-modal satellite time-series data - radar and optical imagery from missions such as the European Sentinel-1 and Sentinel-2 - and learns a compact seasonal "fingerprint" for every 10-metre patch of ground on Earth.

Those fingerprints can then be used very cheaply, often without a GPU, to carry out downstream geospatial tasks: finding crops, classifying land use, locating solar panels or waste dumps, even measuring forest density. We release the map embeddings globally and annually as CC0, alongside equally open weights and code, so anyone can build on them.

To use it, you download the fingerprints for your area of interest and train a lightweight model on top - for mapping crops, forests, flooding or habitats - using a handful of labels rather than millions. We have seen state-of-the-art accuracy across very different mapping tasks from the same embeddings, so we place a lot of trust in them being accurate.

Openness matters here because a foundation model of the planet that you cannot inspect, cannot reproduce, and cannot train yourself would be a dangerous thing to derive social land-use policy from. The people who need it most - conservationists, smallholder farmers, government agencies across the Global South - are the least likely to be able to pay big money for proprietary access.

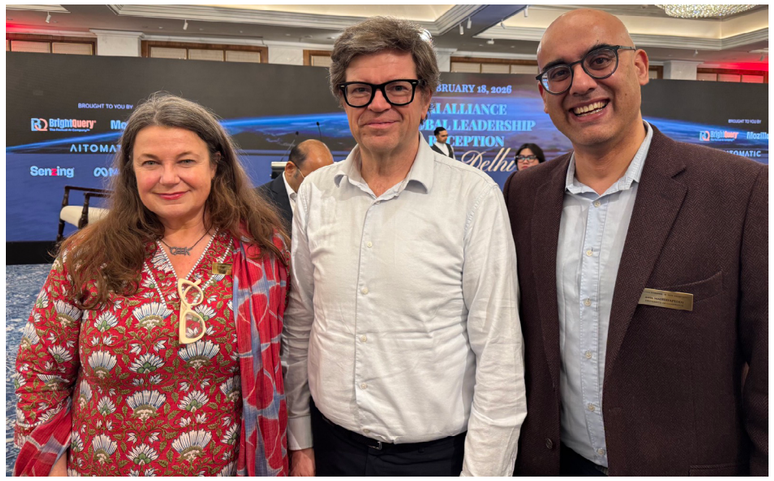
And could you share the hedgehog example our CEO loves?

As a fun but important example, we recently used TESSERA to help understand why Britain's favourite mammal is quietly disappearing. Hedgehogs are now classed as vulnerable to extinction, and rural populations are crashing. With my colleagues Silviu Petrovan,Sadiq Jaffer, Lauren Moore and David Coomes, we are mapping the habitats a hedgehog actually cares about - the "scruffy land" of brambles, hedgerows, and the scrubby green corridors between gardens. Silviu coined the term "digihog", because we are also using GPS-tagged hedgehogs to train models of how they move, so we can predict where habitat preservation is most needed.

TESSERA uses Barlow Twins. How does that connect to Yann LeCun's argument that LLMs are a dead end and self-supervised methods are the future?

The bet in self-supervised learning is that you learn the structure of the world from raw, unlabelled observations rather than from human annotation. That is a perfect fit for satellite data, because the planet pours petabytes of unlabelled imagery down at us every day, and it is enormously expensive to label all of it by hand. With self-supervised methods like Barlow Twins, we never tell TESSERA what a forest or a flood is; it learns a representation of the planet's surface from the raw signal, and we only add human labels at the very end, for the specific task at hand. Barlow Twins, JEPA, LeJEPA and their many variants are all particularly good at building "world models" for physical environments, which is exactly what our geospatial work needs.

Yann argued at the Indian AI summit that the fair way to build open models is for each country to contribute its own data and federate them into one shared model. That may be too late for language models, but it is well suited to satellite data: we already have a public global baseline through Landsat (NASA) and Sentinel (EU), and many countries fly their own higher-resolution task satellites over their own territory. So TESSERA may be the ideal place to actually trial his idea of a global training cooperative, because the incentives line up in a way they simply don't for frontier LLMs.



Claude Fable was used as part of TESSERA. Why was it cut off, in your view, and what has the impact been?

The practical impact is that I am now forced to be reactive rather than proactive.

I maintain a number of open source projects - OCaml, Docker and others - and I receive a steady stream of reported security vulnerabilities against them. It is incredibly frustrating not to be able to run my own scans. Losing access means I have to wait and react to vulnerabilities as they come in, rather than proactively scanning for them myself.

Could this moment be considered a positive turning point?

It can be. Open source is the one form of technology transfer that doesn't require trusting the supplier indefinitely, and that is an enormous advantage when alliances are as fluid as they seem to be these days. The tension is that open source is inherently a global commons. I don't think we resolve that by pretending the borders aren't there; we resolve it by building neutral, shared infrastructure that is valuable enough to everyone that no single nation wants to break it - like the Internet itself. A planetary dataset is a good candidate for exactly that: nobody owns the weather, and everybody needs reliable maps.

What is the replacement, and why DeepSeek specifically? Was there a cost saving or any loss of quality - and does swapping a US proprietary model for a Chinese open one just trade one dependency for another?

DeepSeek is still very expensive to run locally, though the cost of local inference is dropping every day. We had to spend quite a bit of cash to get it working at all, so losing access to hosted models is not ideal in the short term.

And yes, in the long run it does trade one dependency for another. That is precisely why we must ultimately have sovereign models built on open training data - otherwise we are always at the mercy of a supplier. It comes back to the case for open source itself.

Is US policy inadvertently pushing the world towards open source, and towards China's products? Are you inclined to choose non-US commercial technology?

We are "anti-lock-in" more than we hold any geopolitical stance.

As Mark Carney recently noted in his Davos speech, "Middle powers must act together because if we're not at the table, we're on the menu."

Those middle powers include much of the world's Global South, who desperately need every data tool they can get in order to embark on climate adaptation for their populations. Everything we can do to keep TESSERA open is vital to that adoption.

Do you expect your story, and others' experience, to drive more technologists to build on open source - and will that extend to end users?

Every hackathon we have run has been filled to capacity, and our GitHub issues are overflowing with requests. That is simply because building on an open foundation model like TESSERA is the faster, cheaper way to get to a working result, and once a generation of geospatial engineers has tasted that, they don't go back to proprietary tools. Why would they take on all the extra hassle?

The risk we have to manage is community, not adoption. When everyone is coding with AI assistance at high speed, we have to work much harder to keep people talking to each other - mentoring, trading tips of the craft, making sure they actually understand how it all works. Our end users will benefit from the vastly cheaper and more flexible geospatial tasks that TESSERA enables, and that benefit extends to open source use of it too.

What first drew you to open source, and why has it stayed central to your work for over two decades? Has your reason changed?

I was on the original Cambridge team that built the Xen hypervisor back in 2002 - the first open-source type-1 hypervisor, which ended up underpinning the modern cloud. What drew me in was entirely selfish at first: open source was simply the most fun way to learn how the Internet worked. I got started on OpenBSD and PHP contributions in the late 90s, and you could hack on any part of it. It stayed central because I turned that early hacking into a career - in systems infrastructure like Docker, in formal verification with OCaml, and in novel operating systems with unikernels. Nothing proprietary I have worked on has had anything like that reach or longevity.

My reason has changed over the years. In the early days it was about fun and accessibility. In the middle years, around Docker, it became more of a corporate experience - not necessarily worse, because people get paid now. But more recently, working on the planet's climate and biodiversity crisis has turned my position into an ethical one: we must ensure that the poorest people in the world can use open source as a leapfrogging technology to equitably improve their own lives.

In simple terms, what is TESSERA, and what problem does it solve?

TESSERA is a free, open "map-maker" for the whole planet. The problem it solves is that understanding the Earth from satellites used to require a supercomputer, vast storage, and a small army of experts to label imagery by hand. TESSERA does the expensive part once, for everyone, and hands users a ready-made "seasonal fingerprint" of every patch of ground on Earth. After that, answering a huge variety of questions - where the forests are, where the floods will hit, where the hedgehogs live, whatever you can observe from space - becomes about as easy as ordinary programming. And with open coding models now available, that means just about anyone can have a go.

6. Formalities

6.1 Authors of this Report

Professor 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.

Dr. Jennifer Barth, Head of Research, Symmetry Research at FSP and Chief Research Officer, OpenUK

Jenn holds a DPhil from the University of Oxford and is Head of Research at Symmetry Research at FSP, where she brings a rare blend of ethnographic depth, commercial acuity, and ethical clarity to the AI conver- sation. Jenn's work sits at the intersection of technology, organisational culture, and socioeconomic change. She is known for translating complex technological shifts into human-centred insights that resonate deeply with executive audiences. Her measured, critically informed perspective — grounded in qualitative research and lived experience — enables leaders to move from abstract concern to practical, principled action. Alongside her role at FSP, Jenn serves as Chief Research Officer at OpenUK.

Karan Saini, Research Manager, OpenHQ

Karan Saini is a New Delhi-based security researcher, public interest technologist, and open source contributor. He was most recently a Senior Information Controls Fellow with the Open Technology Fund, hosted by the Internet Governance Project at Georgia Tech, where he researched the scale of DNS-based web censorship in India. He regularly writes for leading publications and has contributed to award-winning media projects.

6.2 Creators of this Report

OpenUK

OpenUK is the unique open tech industry organisation 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 manner delivering on three pillars: community, legal and policy and learning. OpenUK's Community is recognised through our world-leading recognition program including the Open UK Awards (the Oscars of Open Source) now in their 7th 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. .

The State of Open Con has become one of the world's leading open source conferences since its inception by OpenUK in 2023. In 2027 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 in London. For 2026 we are on the road, with smaller events across the UK, meeting our audience. Contact mailto:admin@openuk.uk

Symmetry Research

Symmetry Research, an FSP company, 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. We take a research-led approach to helping organisations transform through and with technology. Symmetry Research'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-we do independent research for high-stakes business decisions.

6.3 Contributors

Anastiasia Stasenko
Anastasia Stasenko is the co-founder and CEO of Pleias, a Franco-German startup developing full open source small reasoning models trained exclusively on permissively licensed data. A senior associate lecturer in data analysis and digital strategy at Sorbonne-Nouvelle University, she coordinated the creation of Common Corpus, the largest fully open dataset for training LLMs, with over two trillion tokens.

Professor Anil Madhavapeddy
Anil Madhavapeddy is the Professor of Planetary Computing at the Department of Computer Science and Technology in the University of Cambridge.

David Soria Parra
David Soria Parra is a Member of Technical Staff at Anthropic based in London. He co-created the Model Context Protocol (MCP). He leads its technical direction, governance, and ecosystem development. Before joining Anthropic in 2024, David spent nearly a decade at Meta leading engineering organizations, and served as an Entrepreneur in Residence at Sutter Hill Ventures. He is a long-time open source contributor with a track record spanning PHP, Git, and Mercurial.

Hannah Foxwell
With over a decade of experience in engineering leadership, Hannah Foxwell has always advocated for the human aspects of technology transformation and evolution.Hannah is relentlessly curious about the tools, technologies, processes, and practices that make life better for the people involved in software development. This curiosity has naturally evolved alongside technological changes—from Agile to DevOps, DevSecOps, Cloud Native, and Platform Engineering. In 2026 Hannah co-founded BIMP, bringing together her knowledge of cybersecurity, platform engineering, and AI to deliver software supply chain security at the speed of AI.As the creator of "AI for the rest of us" Hannah leads a community of over 2,000 enthusiasts in the UK, providing free monthly events and an annual conference aimed at demystifying the technology and improving AI fluency.

James Drayson
James Drayson is the CEO and co-founder of Locai Labs, a UK company based in London that develops on-prem sovereign AI so enterprises can own their intelligence. James holds a Bachelor of Science in Biology from the University of California, Los Angeles, and a Master of Science in Innovation, Entrepreneurship and Management from Imperial College London. He is Forbes featured and a UKRI Young Innovator Award winner.

Matthew Barker
Matt Barker is co-founder and CEO of BoltMCP, focused on helping enterprises securely connect AI to their data and control how it's used. Previously, he was CEO and co-founder of Jetstack, where he helped build one of the most widely adopted projects in the cloud native ecosystem, cert-manager. Under his leadership, Jetstack bootstrapped to a team of 50, and was acquired by Venafi in 2020.

Rohit Ghumare
Creator & Maintainer, Agentmemory, CreateMVP, and AI Engineering from Scratch AI-native course. Open source Leader and Cloud-native engineer. He writes and talks about AI agents, harnesses, AI Inference, and the infrastructure underneath, and is currently building iii.dev. Contributor to major open source global projects.

OpenUK AI Openness: 3 Years on From Llama 2, July 2026

Sachiko Muto
Sachiko Muto is the Chair of OpenForum Europe, a senior researcher at RISE Research Institutes of Sweden, and a Board member of the Drupal Association. She originally joined OFE in 2007 and served for several years as Director with responsibility for government relations and then as CEO. Sachiko has degrees in Political Science from the University of Toronto and the London School of Economics; she received her doctorate in standardisation policy from TU Delft.

Sal Kimmich
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.

Vincent Koc
Vincent Koc is Chief Architect at the OpenClaw Foundation, where he leads engineering for one of the fastest-growing open source projects in AI. An engineer and technologist with nearly two decades of experience across the Asia Pacific and the US, he has contributed to major open source projects across the data and AI ecosystem and teaches applied AI at universities including MIT. He joined OpenClaw as a community maintainer focused on security and observability before moving to the Foundation full-time.

Quoted Individuals
Andrew Martin, Co-Founder and CEO, ControlPlane; Adrian Mouat, DevRel Engineer, Chainguard; Chris Yiu, Head of Policy Northern Europe, Meta; Eric Schmidt, Former CEO of Google, Chair of the Special Competitive Studies Project; Jean-Baptiste Kempf, President, VideoLAN and Kyber; Jensen Huang, CEO, NVIDIA; Juliette Denny, Founder, Growth Engineering and Zavmo; Luke Hinds, Founder, nolabs, Inc; Sam Altman, Founder and CEO, OpenAI; Toran Bruce Richards, Founder and CEO, AutoGPT; Yann LeCun, Founder and Chairman, AMI Labs

6.7 Sponsors

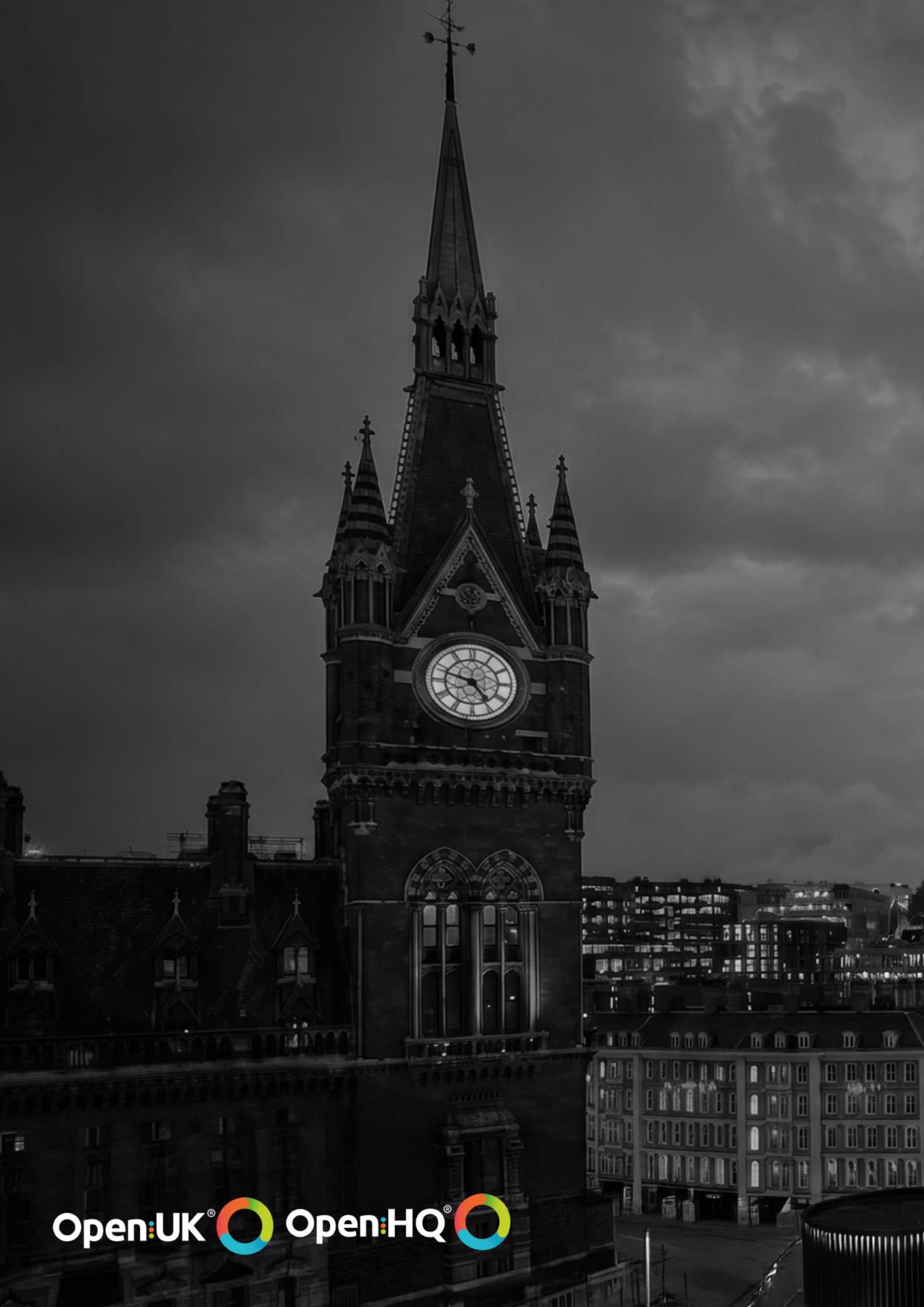
OpenUK is grateful to its many sponsors and supporters without whose support its work including its report would not have been possible in particular, its general sponsors Arm, GitHub, Google, Linaro, Microsoft, Red Hat.

Download the full report here

openuk.uk/report/aiopenness2026

OpenUK Newsletter

Our Conferences - State of Open Con


Open:UK®
Open:HQ®