A policy agenda for building a

UK Research Cloud

Google

Summary

Following the announcement of a pilot of a national ‘data research cloud’ by the UK Government and UK Research and Innovation (UKRI), this Google Cloud white paper outlines the rationale behind establishing a UK Research Cloud (UKRC). Built on public cloud, a UKRC can help deliver on the United Kingdom's strategic vision of becoming the most innovative economy in the world. This proposal also builds on the US National AI Research Resource (NARRI) initiative which is aiming to provide AI researchers access to high performance computing, cloud computing and data.

With daily and significant advances in AI we believe the UK has a unique opportunity to harness these new developments for everyone and deliver sustained economic opportunity. We believe that the development of a UKRC can also support the objectives of the UK Government’s National Data Strategy and its pro-innovation approach to regulating AI.

Ultimately, a UKRC will provide affordable access to state-of-the art artificial intelligence and machine learning technology, computing power, datasets and collaboration tools to academics, institutions, NGOs, businesses and startups, to ensure that every section of society can fully exploit the potential of AI for research, development and economic opportunity.

Background

The UK Government has taken a number of steps in recent years to position the UK as one of the leading countries for innovation. The launch of the UK Science and Technology Framework sets out the government's approach to making the UK a science and technology superpower by 2030. The National AI Strategy launched in 2021 has also been joined by the AI white paper - a pro-innovation approach to regulating AI - seeking to transform the UK economy and society.

Recent announcements in relation to compute and quantum - such as the creation of a: National Quantum Computing Centre and the development of a new national quantum computing strategy have the potential to revolutionise the digital transformation of many key industries across the UK such as healthcare, finance, and manufacturing.

Recognising the potential significant opportunities from democratising access to AI technologies while increasing the availability of public datasets in the UK, Google developed the UK Research Cloud (UKRC) concept in 2020. The premise of the UKRC is focussed on how best to drive innovation and research across the country through adoption of the latest AI tools, algorithms, AI principles and public cloud infrastructure coupled with access to public datasets - and a drive to uplift digital skills across the country to meet future demands from UK employers. Since developing these initial plans, we have been engaging with UK Government departments, public bodies and businesses to test the idea and gather feedback.

Other relevant areas of policy development such as the UK Research and Innovation Strategy point to the time being right to explore the opportunity of a UKRC - converging these important policy initiatives of government alongside a vibrant UK research community and the best innovation the private sector has to offer.

This white paper sets out the primary opportunities of delivering a UKRC and suggests possible models of delivery. In the US, Google has been fortunate to participate and advise, as a member of the National Taskforce for the US NAIRR (National Artificial Intelligence Research Resource) - which has made a number of recommendations to Congress. Our contribution to the debate in the UK is very much in this same spirit of partnership and supporting policy delivery.

Overview

The vision for a UK Research Cloud

A UKRC could increase access to cutting edge AI tools and high-end computational resources for researchers, businesses, and government data scientists - as well as hosting large-scale government-held datasets in a secure cloud environment. We believe that a national consortium could provide this public cloud infrastructure and access to data for all UK researchers and R&D operations, democratising access across different regions and contributing to the UK Government's growth agenda while strengthening collaboration between academic research hubs across the four nations. This would involve a partnership between universities, technology companies, industry and government, and could significantly increase the UK's expertise in AI research.

An UKRC will

- Enable researchers, academics, and government data scientists to keep pace of the rapid changes in technology
- Enable participation and collaboration in further development of AI elements on a wide range of subjects
- Provide UK institutions with direct access into the world of insight across AI to help them identify and develop strategic priorities

Why now?

The UK Government's vision is for the UK to become a science and technology superpower, and a world-leader in AI research and innovation. For this vision to be delivered, it is crucial that the government, combined with investment in basic and foundational research, creates access to meaningful public datasets - and leverages world leading R&D championed by private companies.

Meaningful support for the national research community and academia can be provided through access to the variety of technological innovations in this space. As AI becomes a major driver of global GDP growth, we are entering an era of deployed AI and machine learning (ML), where this prominent technology is being more effectively implemented as part of broader cloud innovation.

The COVID-19 pandemic accelerated increased interest in data analytics and deployed AI to help organisations and governments make smarter decisions, improve their operations, and assist citizens with fast access to insightful data. Facilitating access to AI and ML technology, related research capacity and resources, can be achieved through cloud, and with collaboration between the public, private sector, UK and global research community.

Delivering a UKRC

Our concept proposes establishing a national research cloud infrastructure using public cloud to deliver secure access to the latest cutting edge AI tools and high-end computational resources. Alongside these tools, researchers, academics and government data scientists could also access large-scale government-held and industry shared datasets in that secure public cloud environment.

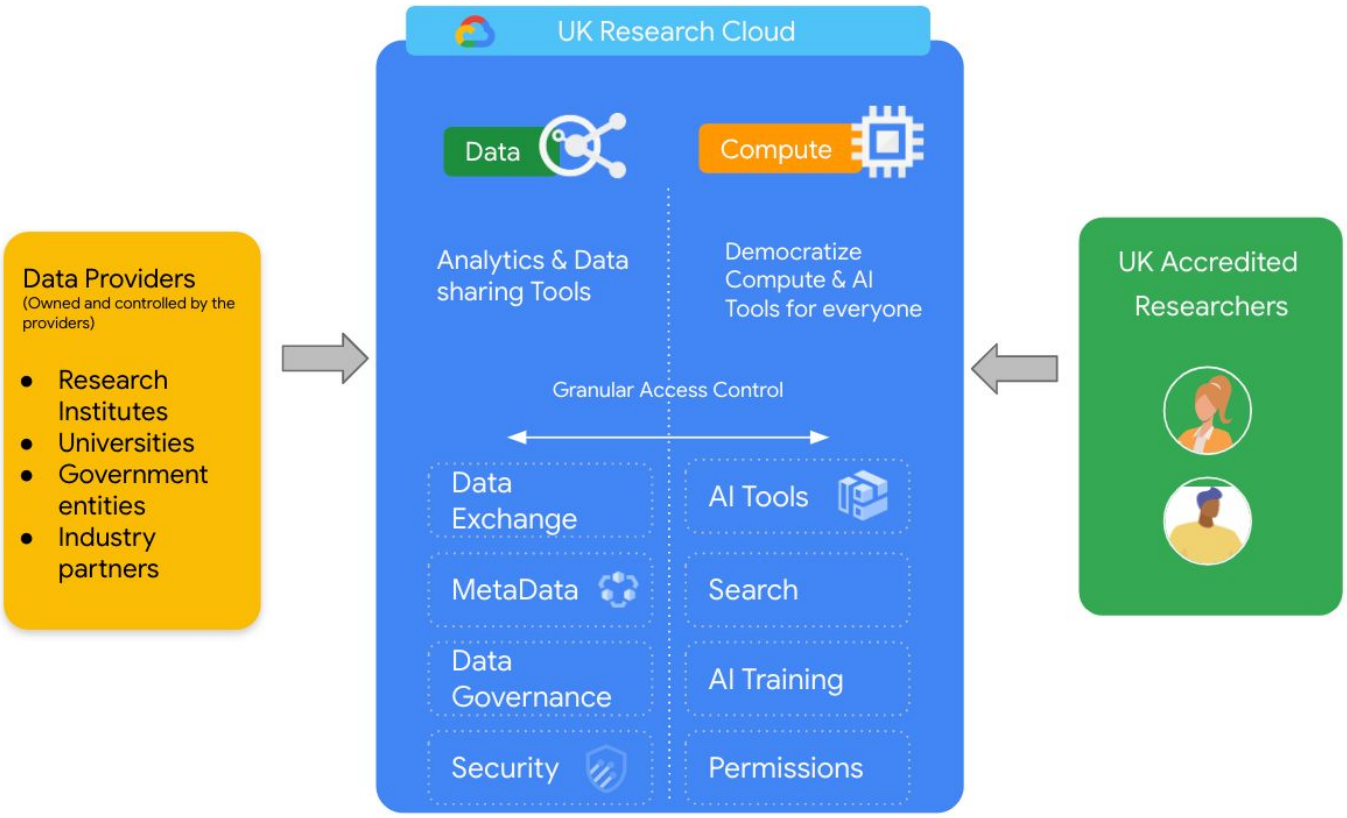
Through provision of a national-level capability, the UKRC would create and build capacity for increased collaboration across academia, government and industry while complementing existing national infrastructure investments. This approach could remove key blockers to innovation and collaboration across research hubs - such as data silos - and accelerate digital transformation.

Managing research data responsibly would also be key to building confidence and a UKRC should be designed while taking into consideration the Findable, Accessible, Interoperable and Reusable Research Data (FAIR) principles and the 7 principles for research data management as recommended by UK Research and Innovation (UKRI). For example, to maintain security and confidentiality around data sharing, the UKRC could provide conditional access as a requirement for sensitive data while retaining open access to public data.

The UKRC infrastructure would bring the double benefit of an increased focus on advancing AI adoption while educating and training the next generation of researchers with more advanced tools and techniques. It would also allow industry partners like Google to share their open source ML algorithms, quantum simulation tools, and access to its Large Language Models (LLM). This would open up access to the latest AI tools that the industry and public cloud has to offer - accelerating researchers' understanding and paving the way for further development, research and applications of AI.

The UKRC could also host its own public dataset program with potential integrations with existing platforms. By maintaining a single copy of research critical , large datasets on public cloud, the UKRC would be able to maintain and provide shared data access to its users and industry partners. This will also help curate the data to maintain its relevance and avoid data duplication across various institutes that drives up storage infrastructure costs in the long run.

The UKRC and its users would also benefit from public dataset programs offered by public cloud providers, e.g: Google's public dataset programme offers more than 200 publicly accessible datasets on a wide range of categories (at the moment its volume stands at 20 Petabytes) along with access to more than 1,000 geo-spatial and satellite imagery datasets (over 80 PetaBytes) through its Earth Engine program.



Exploring the wider impact of a UKRC

The delivery of a UKRC would support many of the ambitions and the vision set out by the UK Government on AI, access to data and innovation. There are potential benefits for policy alignment - helping the government deliver existing policy objectives more effectively. The wider economic imperative of accelerating research across AI is also widely recognised - the UKRC would help position the UK in the global drive for competitive advantage. New and existing infrastructure investments could also benefit through increased design, construction and operation optimisation.

Helping strategic alignment across key policy priorities

The UKRC could be developed to complement existing national infrastructure and support key government policies and initiatives such as:

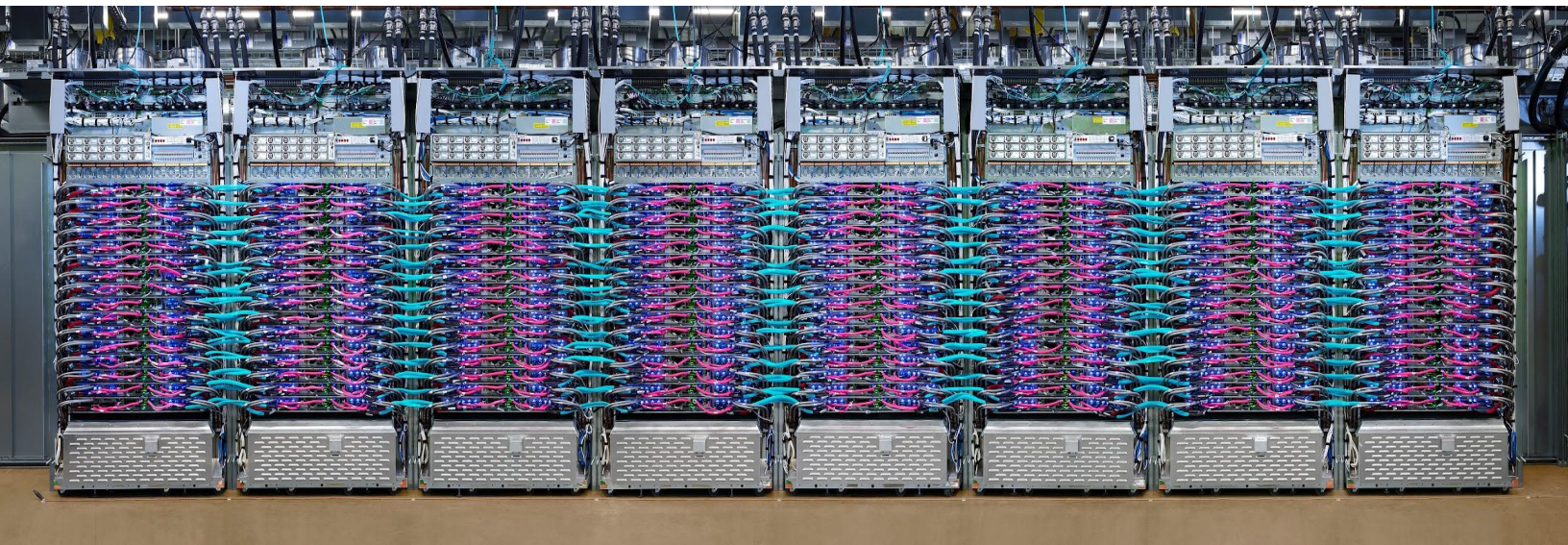
- Supporting the UK National AI Strategy by democratising AI R&D in a responsible manner to benefit citizens, science, industry and technology.
- Disseminating AI principles, tools, methods for building trustworthy AI, explainability, safety and bias mitigation techniques through widely accessible infrastructure and training.
- Enabling affordable access to state-of-the art AI and ML technology and computing power via public cloud - with an objective to boost UK national research potential and articulate the nation's leadership in AI and R&D at the global arena.
- Supporting the UK Government's R&D roadmap's stated ambition to 'develop the UK's digital research infrastructure capability by building an internationally leading national digital research infrastructure.
- Complementing the National Data Strategy (NDS) and the UK Government's plans to encourage better coordination, access to and sharing of quality data between organisations in the public (e.g ONS) and the private sector.

Accelerating AI research across the UK

- Strengthening UK universities' research excellence in AI and attract high-calibre AI researchers - all core components of the UK Government's AI strategy.
- Supporting university startup incubator programmes to help train the next generation of researchers, empower future entrepreneurs and help SMEs grow and scale
- Supporting the growth of computational science, including the use of ML to accelerate breakthroughs in a wide range of scientific fields, such as genomics, healthcare and structural biology.

Multiplying existing investments and infrastructure opportunities

- Enabling UK Government data scientists to leverage UKRC resources aligned with the national R&D objectives and public sector priorities - including improving the quality of citizen services across different focus areas (e.g. transport, digital finance, meteorology, statistics etc.).
- Supporting the UK Government's efforts to more evenly distribute R&D activity across the UK, and outside of London and the south east



Supporting support job creation and workforce development

While increased AI led digital transformation across industry will bring significant change and uncertainty, there will also be job creation opportunities that can be maximised by early workforce development interventions driven by government and the private sector.

Ensuring that the talent and skills mix is in place across the UK is a key concern across many sectors and for the UK Government in general. This is particularly challenging when it comes to AI, and we welcome the drive to create more AI and data conversion courses to help young people from underrepresented groups into the industry.

The UK Government has committed more broadly to supporting universities in their efforts to create opportunities across the country. In recent years we have seen funding for new R&D centres, and initiatives to improve access to higher education in a more balanced way across the UK.

We believe UKRC can also help by providing access to affordable and reliable IT infrastructure and services in every area of the country. This would support businesses and organisations in all parts of the UK, to innovate and grow - and consider how best to invest in future skills.

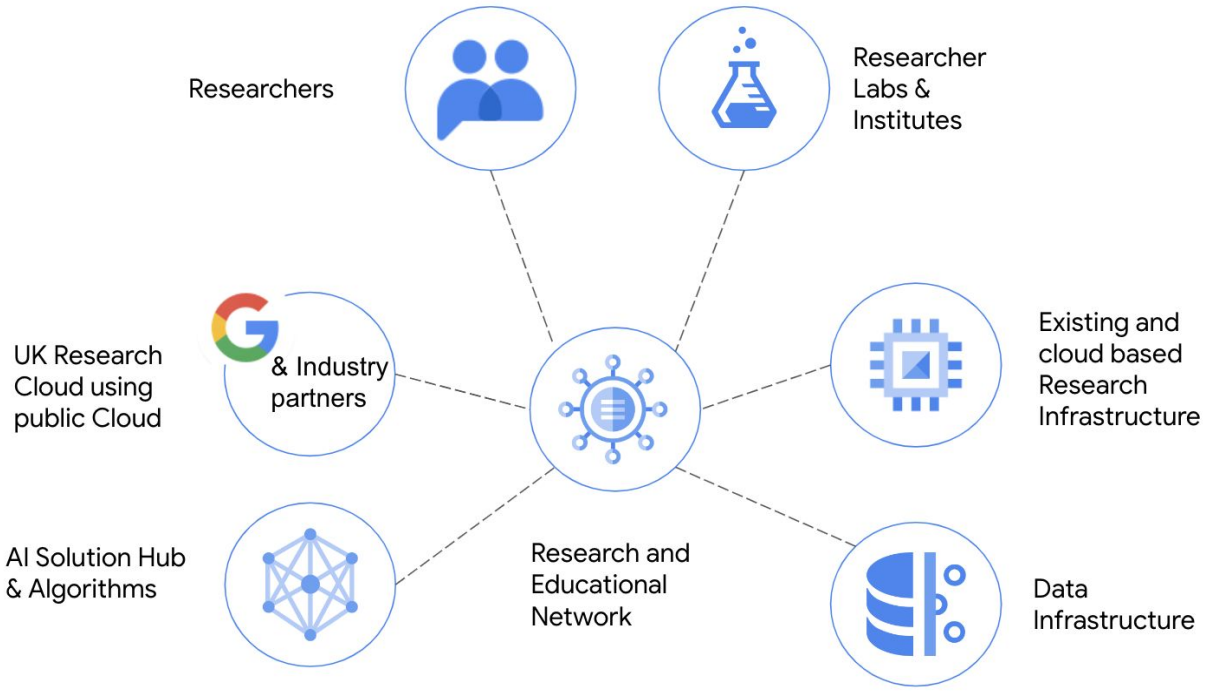
AI, data and cloud computing have the potential to play a significant role in helping the UK to create a more prosperous and inclusive economy. The development of a UKRC can not only provide opportunities in and around research, academia and universities across the UK, it makes it easier for businesses - wherever they are - to benefit from new technology without the significant levels of investment in technology that would have been required in the past. This will improve productivity, reduce costs and create new business opportunities.

This new reality will also create opportunities and challenges in employment and skills. Existing UK Government programmes such apprenticeships, digital bootcamps and the AI scholarship scheme - alongside free training programmes we operate at Google in career development and Google Cloud certification - are playing an important role in helping people develop the skills that they need to fill the gaps in the current job market.

However, in areas such as cyber security, these challenges are more profound and will require increased emphasis from government and business. To address this, a UKRC is likely to foster greater collaboration across the economy to help decision makers better understand the future skills challenges that we face and lead to more impactful interventions from government and industry.

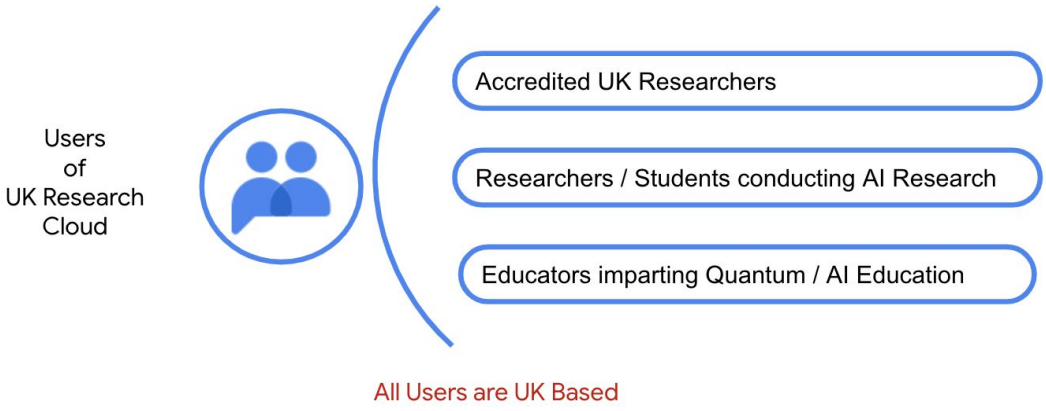
Increasing collaboration

Research and innovation is a collaborative effort, where great minds come together. By bringing together people, organisations, AI tools, data and research infrastructure, a UKRC could enable easier data sharing and collaboration across public and private industries. Collaborating researchers will have open access to all public datasets and limited or secure controlled access to private data where needed.



Maximising access for all

The UKRC should be accessible to a wide range of accredited and diverse users from UK research institutes, educational institutions and government bodies - opening access to advanced AI tools, methods, data and training resources for all. Centralised access and identity management or adoption of existing persistent digital identifiers for researchers will be key to ensure that this public infrastructure serves people as intended.



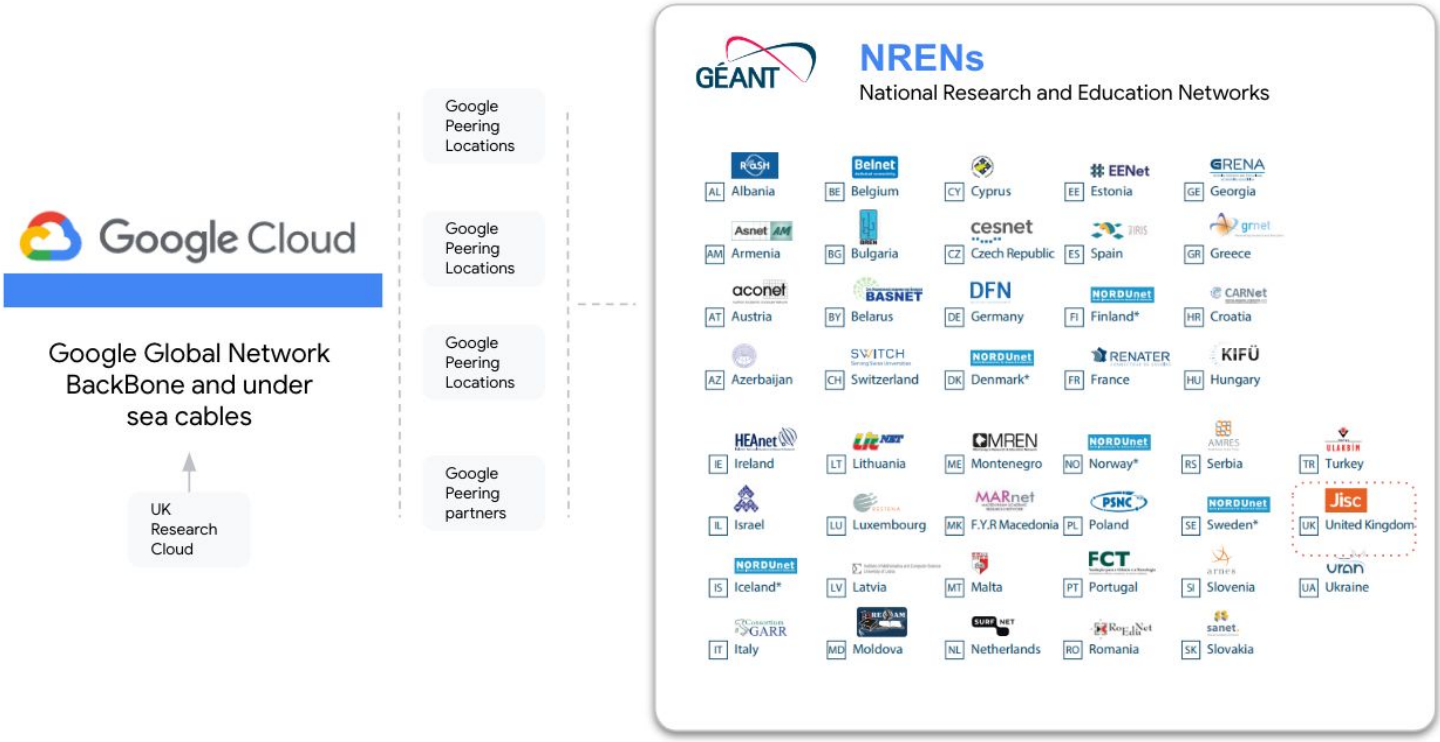
Improving connectivity and security

The UKRC could have priority focus on providing national and international connectivity to its research users with a security model based on zero trust. By adopting a layered security model providing transparency, cyber security, auditability, Threat intelligence and protection - UKRC could protect both data and research that it hosts and that of the researchers and users of the platform.

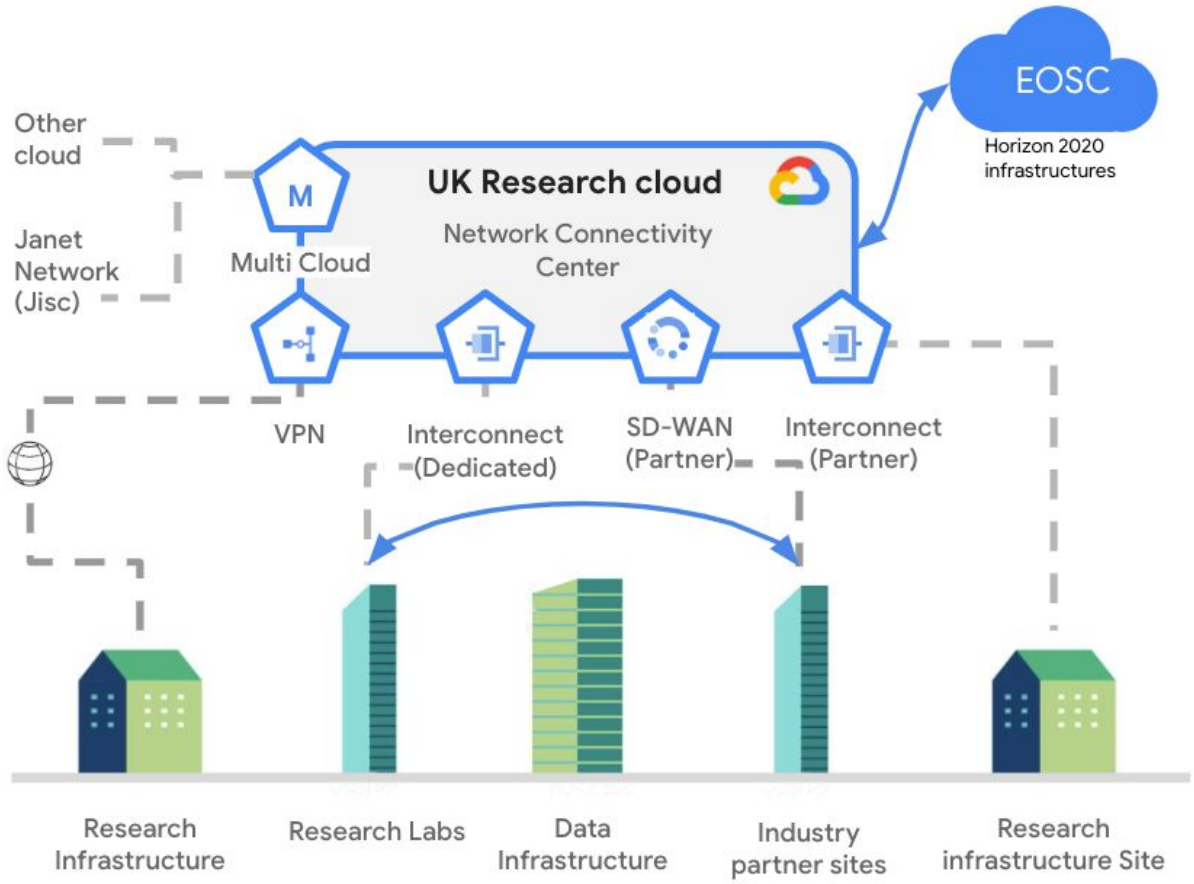
Connecting to critical national infrastructure and global research networks

The UKRC could seek to build on a model that enhances international cooperation by delivering capability that connects UK institutes with research organisations across Europe and the rest of the world. This could be achieved through both existing government research investment in infrastructure and networks - and in networks complemented by UKRC's use of Public cloud infrastructure - e.g. Google's global network backbone infrastructure.

As an example, public clouds such as Google Cloud and partners have already established relationships with both UK (JANET) and European National Research and Education Networks (NREN) through the Open Clouds for Research Environments Project (OCRE). A UKRC would be well placed to maximise the potential of these existing partnerships.

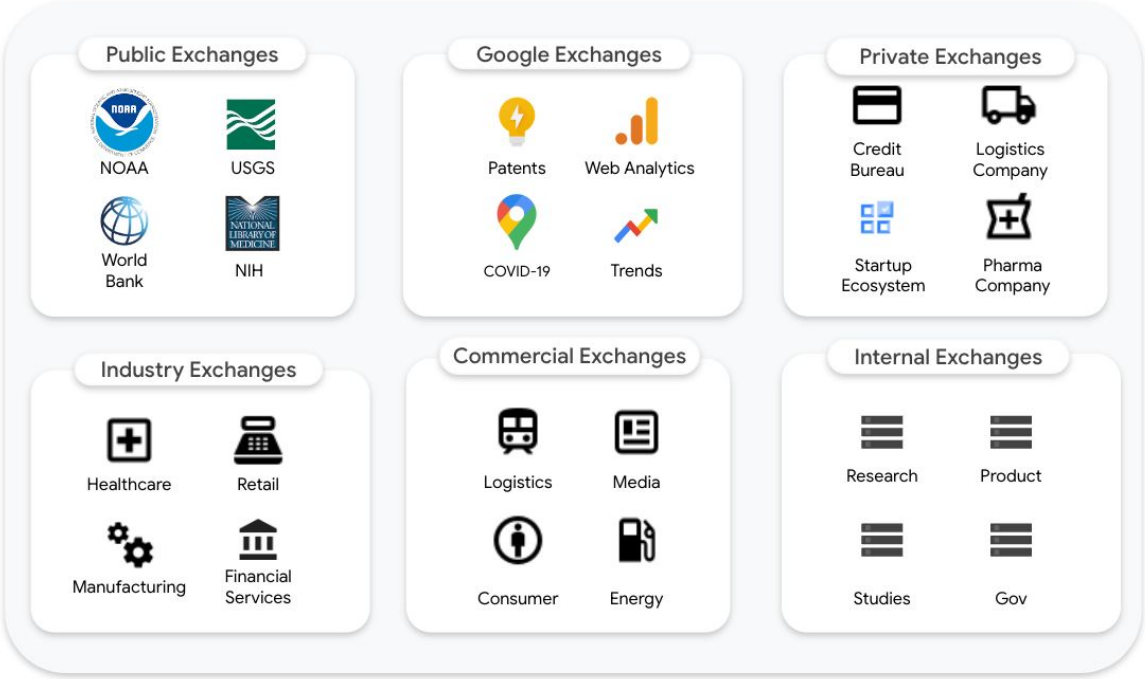


UKRC could also provide the capability as a connectivity hub when needed, for transit connectivity between critical national infrastructure, deployed on-premises, industries and between other cloud providers to enable collaboration and provide more network and connectivity options.



Data ecosystem - data exchange and sharing

The UKRC could host a data exchange consisting of open datasets for research, industry hosted data and data that requires to be privately shared between collaborating entities and researchers. Building this data ecosystem will be key in bringing together participation from industry, academia, research organisations and the government.



The UKRC data exchange would enable

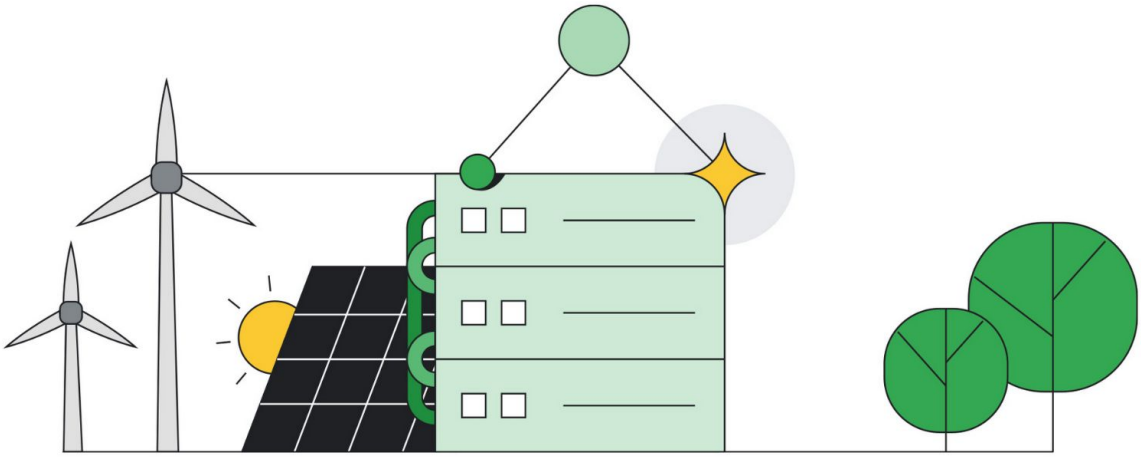
- Public Exchanges to host open datasets that benefit a wide range of research topics in a readily consumable format.
- Exchanges for Data to be shared by technology partners like Google and Google Cloud - e.g Google Trends, Earth Engine, Covid19 etc.
- Private exchanges between organisations that are commercially sensitive in nature.
- Industry exchanges for data in areas such as manufacturing, retail and health that have significant impact at a national level.

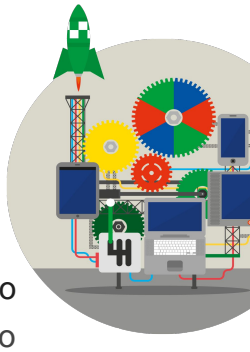
- Commercial exchanges where data generators will have the ability to charge a fee to recuperate their costs on data generation and storage.
- Internal exchanges that can be setup for data sharing within an organisation , e.g. a government department, a research institute or a university..

Public cloud providers such as Google Cloud already host curated public datasets on their infrastructure. Such datasets could also be used or surfaced within the UKRC data ecosystem as it would be designed on a public cloud infrastructure.

Sustainability in everything we do

The UKRC infrastructure would look towards using datacenters powered by sustainable sources of energy or carbon-free energy. A UKRC should also work with public cloud providers that pursue new carbon-free energy generation and storage technologies that work with governments, utilities, and policymakers to deploy those technologies and drive system-level change. By supporting public policies that advance robust global climate action, we would create new pathways to a carbon-free economy for the future.





What next?

We know that AI has the potential to solve some of society's most pressing problems. To realise that potential we believe that the creation of a UKRC - connecting talent to accessible cutting-edge AI tools underpinned by public cloud infrastructure - would be a positive step in the right direction. A UKRC would enable us to innovate, accelerate and deliver world leading science, research and technology - here in the UK - for citizens and for the benefit of everyone around the world.

The UK Government is already making significant steps towards advancing UK capability in relation to AI and through the development of initiatives such as the Integrated Data Service from the ONS. We are particularly supportive of the plans to pilot a data research cloud and this white paper is part of our contribution towards the thinking around this significant initiative.

As the UK Government seeks to develop international cooperation on the responsible use of AI, we believe that a UKRC can also play a positive role to ensure that AI is developed in a way that is aligned with global values and standards - and that data is shared securely across borders - helping produce more accurate and better outcomes.

With plans for a data research pilot already underway we believe there is an unique opportunity to go even further and establish a national consortium to deliver on this commitment and help strengthen the UK's compute ecosystem. We believe that a partnership of this nature - between business, academia, technology companies and government - would be the right approach and the UK Government should look to bring together interested organisations to build on this proposal in the coming months.

This white paper is our contribution to this important debate and we look forward to working with industry partners and government on this exciting journey.

Sources and References

HM Government - National AI Strategy

UK Research and Innovation Strategy

UK Levelling up Executive Summary

Science and Technology Framework

NAIRR Final report to US Congress

National Quantum Strategy

Google case studies and blogs

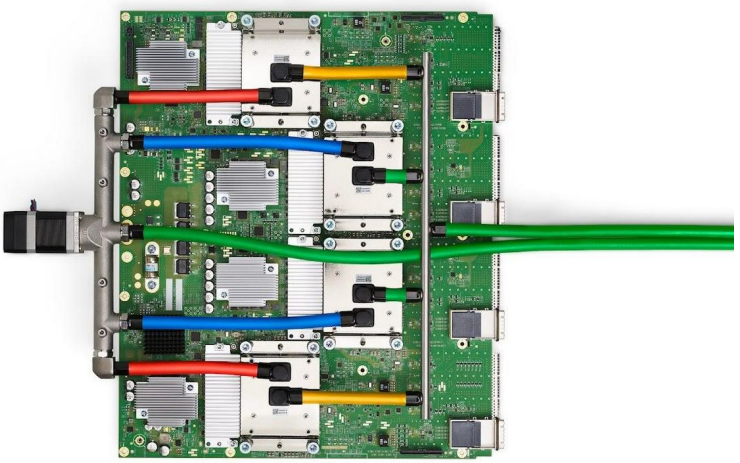
A policy agenda for responsible AI

EMBL-EBI - Strategic partnership

Alphafold - using-open-data-and-ai for 3D protein structure prediction

A responsible path to generative AI in healthcare

Introducing PaLM2



A policy agenda for building a

UK Research Cloud