# Innovation mapping now

Innovation Mapping Team, Nesta

March 2019

nesta

## Acknowledgements

## About Nesta

Nesta is a global innovation foundation. We back new ideas to tackle the big challenges of our time.

We use our knowledge, networks, funding and skills - working in partnership with others, including governments, businesses and charities. We are a UK charity but work all over the world, supported by a financial endowment.

To find out more visit **www.nesta.org.uk**

If you'd like this publication in an alternative format such as Braille, large print or audio, please contact us at: **information@nesta.org.uk**
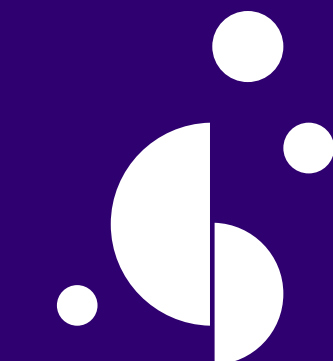
nesta

# Innovation mapping now

Innovation Mapping Team, Nesta

# Executive summary

**There is a glaring disconnect in our networked era. We create brilliant machines but our productivity growth is stagnant. Amazing scientific advances and technological innovations are shadowed by increasing inequality, the threat of climate change and societal discontent.**

Economic and innovation policymakers seek to fix this disconnect with ambitious missions, challenges and national strategies that align technology development with societal needs and make innovation more inclusive. Traditional data sources, with their static taxonomies, aggregate statistics and simple summaries of a complex economy fall short of these policymakers' evidence needs.

New, more detailed and timely data, powerful analytics and ways of presenting information could help them. Business websites, science and technology databases and social media platforms can be used to monitor the emergence of new technologies and the structure of innovation systems, while interactive visualisations and open datasets empower more users to answer questions about their own industries, locations and communities. But these novel data sources, methods and tools come with their own risks - their quality is unknown, they are hard to interpret, and they raise ethical questions. As a consequence they are not trusted and their adoption in policy has been sluggish.

Nesta's innovation mapping team seeks to overcome these barriers with an approach that is user-driven, pragmatic, eclectic and open: we deliver innovation maps that can inform today's policies and perhaps inspire tomorrow's. We make our tools open so others can build on them, and to build trust around them. We seek to strengthen the innovation mapping skills of our policy partners and collaborate with researchers so that our work doesn't just create more data, but also advances our understanding of innovation.

Ultimately, we want to help bring innovation mapping to the mainstream of innovation and growth policy because we believe that will make it smarter, more inclusive and fit for the future. This is an ambitious mission, but one worth pursuing.

# 1

# Why we need innovation mapping

**Imagine using a 2007 newspaper to try to understand what is happening in the world today. It would not be very useful. Yet this is what we do with the economy. The UK's economic statistics are based on industrial codes agreed in 2007. This means that they cannot measure the new industries that have appeared since, from modern applications of artificial intelligence (AI) to the immersive economy and clean tech. Yet these are precisely the sectors that could transform the economy and help address some of our biggest social and environmental challenges (while admittedly creating new ones).**

Policymakers want to monitor the development of these new technologies and industries as well as their impact, but this is not easy with traditional data. Paraphrasing US economist Robert Solow, you could say that we see innovative firms everywhere except in the official statistics.[1]

But identifying where new ideas come from and understanding their impact is more important now than ever. The recent productivity record is abysmal, and the gaps between leader and lagging firms, as well as those between creative cities and 'left behind places' are widening, raising questions about market power and creating political instability. There is an apparent disconnect between measured R&D investments on the one hand and prosperity, wellbeing and sustainability on the other.

Governments across the world are trying to tackle this productivity puzzle with activist industrial and innovation policies. The UK government is pursuing sector deals and setting up ambitious missions and challenges to strengthen specific sectors and technologies like AI or life sciences and to deal with major economic, social and environmental problems like ageing and climate change. Missions will play an important role in Horizon Europe, the new EUR 100 billion framework programme for R&D in the European Union. Meanwhile, more than twenty governments across the world have created national strategies to develop their AI industries safely, another example of the desire to achieve technological leadership while delivering social goals.

But how will we know if all these policies are headed in the right direction, and at what speed? How do we avoid going down dead-ends, or getting prematurely locked in to the wrong innovation trajectory?

This report argues that we need better maps to inform these new policies. Traditional data sources, such as business and innovation surveys, and aggregate measures of R&D spend, STEM graduate supply and patenting, were never designed to create those maps. They were instead created to monitor a simpler, slower-changing economy where decennial updates in the industrial codes were enough to capture the emergence of new industries, and where the market was relied on to allocate resources best. They took manufacturing as the main locus of innovation and national policymakers as their main audience, ignoring the innovation activities and data needs of many other important sectors and audiences, locally and outside government. Many of these assumptions are however being increasingly challenged.

New datasets, analytics methods and interactive visualisations could help address some of these important gaps in the evidence base (the glossary provides brief definitions of technical terms we use here and elsewhere in the document). With today's data and open source tools, we can:

- Analyse millions of documents ranging from patents to business websites to monitor how new technologies and industries emerge almost in real time.

- Make predictions about future technological trajectories and interesting innovation events using machine learning.

- Identify gaps in innovation systems and new opportunities for collaboration and creative recombination through network analyses and complexity science.

- Disseminate the results of these analyses via interactive data visualisations and dashboards, search engines, open databases and open-source software that others can build on.

This report summarises how we are exploring these opportunities at Nesta.

Section 2 sets out the policy context, evidence gaps and new mapping opportunities and challenges.

Section 3 describes innovation mapping at Nesta: what we are trying to achieve, our audiences, projects, methods and outputs.

Section 4 concludes with a vision for policies that are powered by innovation mapping to become smarter, more inclusive and fit for the future, and sets out ways you can work with us to push this agenda forward.

# 2

# Why now?

## The policy problem

**A persistent productivity puzzle**

Western economies have seen their productivity growth stagnate in the last two decades. In the EU, it has declined from 2 per cent annually in 1995 to less than 0.5 per cent now. The situation in the US is not much better, with annual productivity growth of 0.6 per cent since 2011.[2] This stagnation is a cause of grave concern. Productivity growth drives improvements in living standards; our aging population has to be supported by fewer and fewer workers and the sustainability imperative demands more efficient economies.

The source of this problem does not seem to be a deceleration in technological innovation.[3] To the contrary, there is a widespread perception that technological progress is speeding up, with rapid advances in disruptive technologies such as AI, robotics and drones, virtual and augmented reality, cryptography and distributed ledgers, the internet of things, gene editing and clean tech. If anything, policymakers and the public worry about a 'future shock' as these technologies drive change faster than individuals, businesses and societies are able to cope with.[4]

Economists have proposed many explanations for the productivity puzzle, such as:[5]

- A decline in the productivity of R&D investments as innovation processes become more complex and new ideas get harder to find. For example, today it takes 18 times as many researchers to sustain the advance of Moore's Law in electronics as it did in the 1970s.

- Shortages in the skills required to successfully apply new technologies, which is partly why machine learning PhDs straight out of university get paid seven digit figures in the technology sector.

- Lags caused by the need to experiment with business models and processes in order to discover those that work best with new technologies.

- A concentration of the gains from innovation in a small number of superstar firms with the clout to attract the smartest talent and acquire potential competitors.

- Misalignments between R&D priorities and societal needs, as science funding gets captured by vested interests and businesses deploy disruptive technologies such as AI without considering their negative side effects (for example, in terms of algorithmic error or labour market disruption).

Previous rationales for innovation policy based on market failures (the idea that markets fail to provide sufficient incentives to invest in R&D) and systems failure (concerned with disconnections and gaps between different participants in the innovation system such as researchers and industry) do not seem enough to tackle these productivity growth headwinds on their own. New models are needed.

**New wave economic and innovation policies**

In recent years, a new wave of industrial and innovation policies has come onto the scene in Western Economies.[6] These new policies are more directional, activist and holistic.

What do we mean by this?

- **Directional**: They recognise that there are many possible economic and technological trajectories, and that some are preferable to others.
- **Activist**: They seek to steer economic and innovation activities towards accomplishing particular goals.
- **Holistic**: They pay closer attention to the mix of capabilities, skills, resources and infrastructures that need to be in place for successful technology deployment, and care about the inclusiveness of the processes and constituencies that drive these processes forward.

The summary table below compares key features and evidence needs of this new wave with old rationales based on market failures and system failures.

**Summary table: Three frameworks for innovation policy**

| | Market failure | System failure | Emergence failure |
|---|---|---|---|
| **Rationale for policy** | Insufficient investment in R&D because its benefits are not fully captured by the firms carrying it out | Disconnections between researchers and industry due to differences in culture, goals and values | Uncertainty and complexity hinder the application of new technologies to tackle societal challenges |
| **Flagship policy** | R&D tax credit | Collaborative R&D grant | Innovation mission |
| **Key indicator** | Level of investment in R&D | Collaboration between university and industry | Emergence, diffusion and impact of new technologies |
| **Data sources** | Business surveys, scientometric indicators (patents, publications) | Innovation surveys Knowledge exchange surveys | Text and network data from open and web data |
| **Outputs** | Indicators | Scoreboards | Innovation maps |

This new wave draws on the notion of 'emergence failure': the idea that radical new ideas can take a long time to emerge, or even fail to do so for a number of reasons.[7] It might be that key ingredients in their ecosystem (from skills to standards) are missing, they lack a niche to develop and become more competitive, the actors who would need to come

together to develop them are not connected with each other.[8] Examples of these challenges include the slow development of alternatives to fossil fuel or new models for delivering healthcare.

Emergence failure risks locking-in our economy to established ways of doing things, a path of diminishing returns. Traditional policies are of limited use to overcome this inertia because they tend to favour powerful incumbents and incremental varieties of innovation.[9] They support activities that would have happened anyway, or perhaps should not even happen at all.[10]

Advocates of the new wave argue for policies pursuing ambitious, concerted action to bring together distant combinations of knowledge and new alliances of innovators.[11] This can help build new industries and deliver the transformational types of innovation required to address important societal challenges. Grand challenges and missions can act as a focal point for these coalitions. The UK Government's Industrial Strategy, with its challenge areas and missions, and its Industrial Strategy Challenge Fund is pushing in this direction.[12] It is similarly expected that mission-oriented innovation will play a significant role in the EUR 100 billion EU R&D Framework Programme, Horizon Europe starting in 2020. The Green New Deal being called for in the US is another example of this sort of policy.

Another manifestation of increased activism, directionality and holism in innovation policies is the recent proliferation of national strategies to support emerging technologies such as AI.[13] These strategies combine traditional policy levers around research funding and business support, with interventions to increase the supply of other important inputs such as machine learning skills and data. They also incorporate safety and ethical considerations and seek to create incentives for the deployment of AI in sectors where its societal benefits could be considerable, such as health or education.[14]

**Not just old wine in new bottles**

One thing that sets this new wave of economic and innovation policies apart from previous science and technology-driven missions, such as the Manhattan Project or the Apollo program, is their shift away from purely technical or economic problem-solving as they seek to tackle big social challenges that people face in their daily lives.[15]

This broader and in some ways more ambitious remit for missions requires more inclusivity in how they are defined, designed and delivered. A broader set of communities need to be involved to ensure that missions are relevant and legitimate.[16] They also need to bring together disciplines and sectors that go beyond Science, Technology, Engineering and Mathematics (STEM) to also incorporate social sciences and arts and humanities, as well as non-technological innovators in the third and public sector and the creative industries.

New wave innovation policies are also more explicitly experimental than those that came before, acknowledging radical uncertainty by trying out new approaches and adapting in response to new information about what works and what doesn't.[17] In doing this, they try to avoid the frequent criticism that industrial and technological activism distorts the economy by 'picking winners' that eventually disappoint technologically or commercially, and risks capture by vested interests.

These concerns raise an important evidence question: how do we ensure that these new policies are informed by the best possible data in order to realise their potential and avoid their significant pitfalls?

## An innovation mapping opportunity

New wave industrial and innovation policies require a detailed, timely and relational view of economic and innovation systems: what technologies are being developed and diffused today, what are the connections between them, and what are their economic and non-economic impacts.[18]

The last decade has seen the emergence of a new field of science and innovation mapping that uses new and big data sources and analytical methods to generate better representations of complex, fast-changing innovation and economic fields, potentially helping to address these new policy needs.[19]

In this section we outline some 'burning' questions for R&I policymakers today that traditional sources struggle to answer, and where innovation mapping could inform better policies.

**What is the industrial and technological composition of the economy today?**

Traditional economic data, for example from business surveys, are structured around out-of-date industrial taxonomies: they cannot be used to study new sectors like AI or immersive technologies (including Augmented Reality and Virtual Reality) that don't fit in those pre-existing codes. Further, this data takes a long time to collect and analyse so it gives a rear-view of the economy. For example, the latest version of the UK innovation survey refers to the period 2014-2016.

Real-time, detailed data from web sources can give us more relevant and timely information.[20] Text in business websites, social media platforms and open datasets tell us about innovation in software development (in GitHub), fund-raising (in Kickstarter) or networking (in Meetup). Many of these sources are updated regularly, giving us a picture of the economy much closer to real time than we are able to get with official sources.[21]

We have worked with this data in multiple projects. For example, in our map of the UK immersive economy we worked with Glass, a big data startup, to analyse the websites of hundreds of thousands of UK businesses, identifying those that mention keywords related to immersive technologies such as Virtual Reality and Augmented Reality. This allowed us to identify almost two thousand companies in the sector, and to understand its geography and evolution (Figure 1).[22]

In another project we measured activity related to AI in arXiv, a repository of research widely used by computer scientists and engineers. Our research showed that those locations that combine AI R&D with relevant industrial activities tend to develop the strongest clusters in this general-purpose, high-impact technology (Figure 2).[23]

**Figure 1: The immersive economy in the UK**
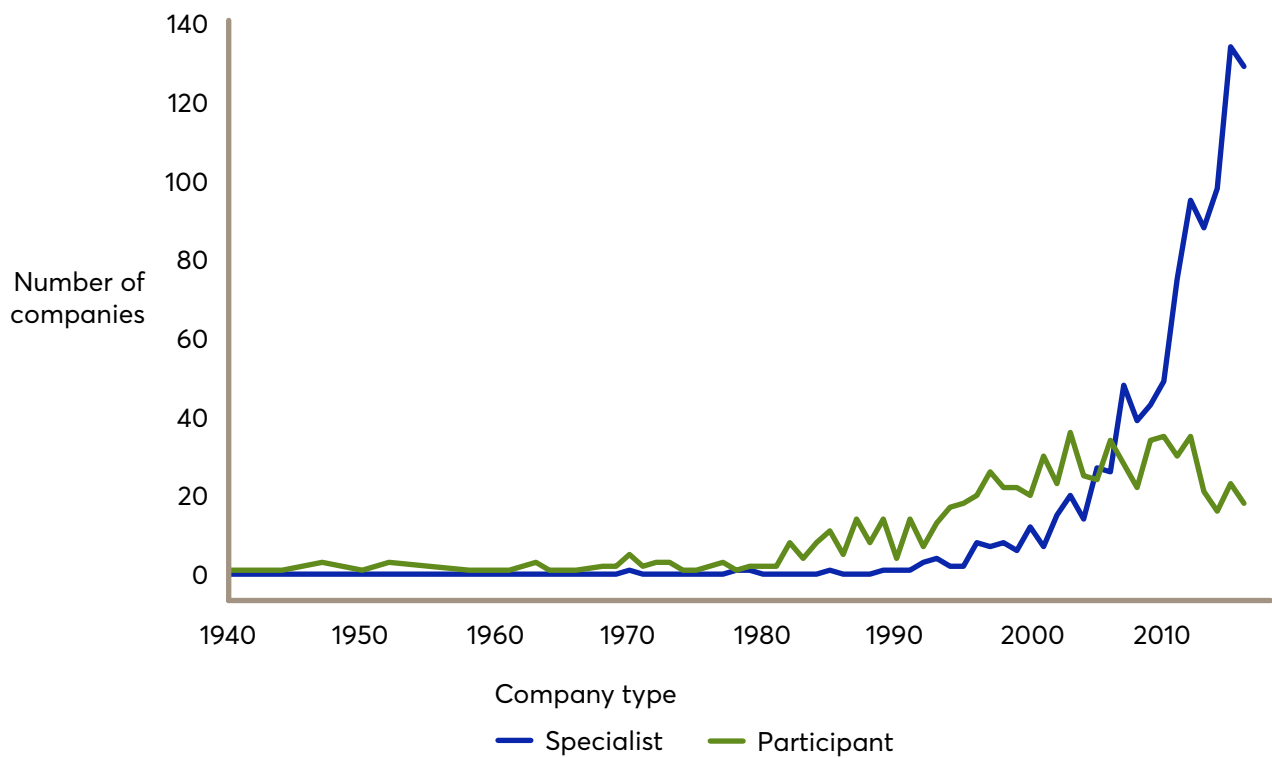


This figure shows the incorporation date for immersive companies identified in our immersive economy analysis. We used web data to identify immersive companies which we then surveyed, helping us to measure their level of specialisation in immersive technologies (based on the level of revenue they generate from those technologies). We then trained a model on those survey responses that determined what business characteristics (and website text) was most predictive of a company's immersive specialisation, and applied it to the population of firms (including those we hadn't surveyed). The chart shows rapid growth in activity, consistent with the emergence of this sector.

**Figure 2: Mapping AI research globally**



- ■ 100 AI papers, specialisation gain since 2012
- ■ 500 AI papers, specialisation gain since 2012
- ■ 1,500 AI papers, specialisation loss since 2012
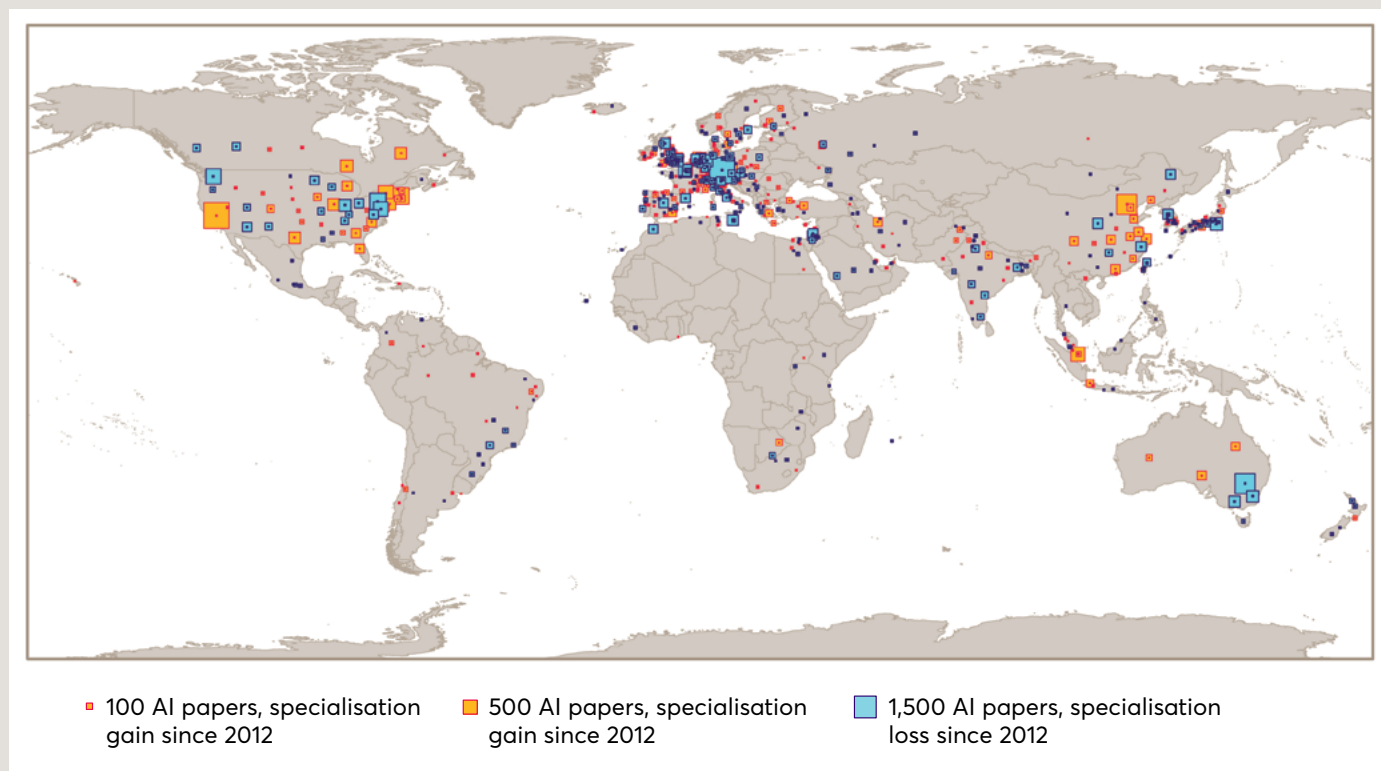
This map shows the level of AI research activity based on pre-publications in arXiv in different regions of the world. The shape of the square represents the number of AI papers and the colour shows whether the region has become more (orange) or less (blue) competitive since 2012. We note that gains in competitiveness in China in the East Coast of the USA, and the slightly more mixed situation in Europe.

These projects helped us measure innovation activities that policymakers are already interested in. The next step is to use text and network data to try to identify emerging technologies that we are not even aware of yet. This is what we are doing with Rhodonite, an open source tool that looks at the evolution of scientific and technological language to identify new combinations of interest for policymakers.

**How much innovation is happening outside of manufacturing and STEM disciplines?**

Traditional indicators such as scientific and technological R&D, patents and supply of STEM graduates exclude important innovation activities in the creative industries and third and public sectors that cover a significant part of the economy and are vital for delivering innovation missions that seek to achieve technological and societal change.

We can increasingly complement paper and patent data with other sources of information that might be more relevant for those 'hidden innovators'.[24] For example, our project mapping the UK video games industry used data scraped from video games directories, wikis and fan sites to measure and map the sector.[25] In doing this, we measured important dimensions of innovation like the platforms being targeted by different games studios that are not captured in traditional statistics and surveys yet are important for understanding the situation of the sector (for example, the skills needs of an indie studio making games for an app store are completely different from those of a console games developer with dozens of employees).

**What is the structure of the R&I system and the connections between different fields?**

Simple R&I indicators and KPIs do not consider connections between technologies and fail to link information across datasets. This creates a siloed view of innovation systems and the economy, and misses the crossover between sectors, which is often precisely where innovation is happening.[26]

Today, we can use network and semantic analyses to draw links between organisations and fields, and connect information from different datasets.[27] In Arloesiadur, our project mapping innovation in Wales, we mapped networks of research collaboration between academia and industry with interactive data visualisations that can be explored to understand who is working with whom - and where are the gaps - in Wales' research and innovation system [Figure 3].[28]

**Figure 3: The innovation systems that are, and the innovation systems that could be**

Existing network          Recommendation network
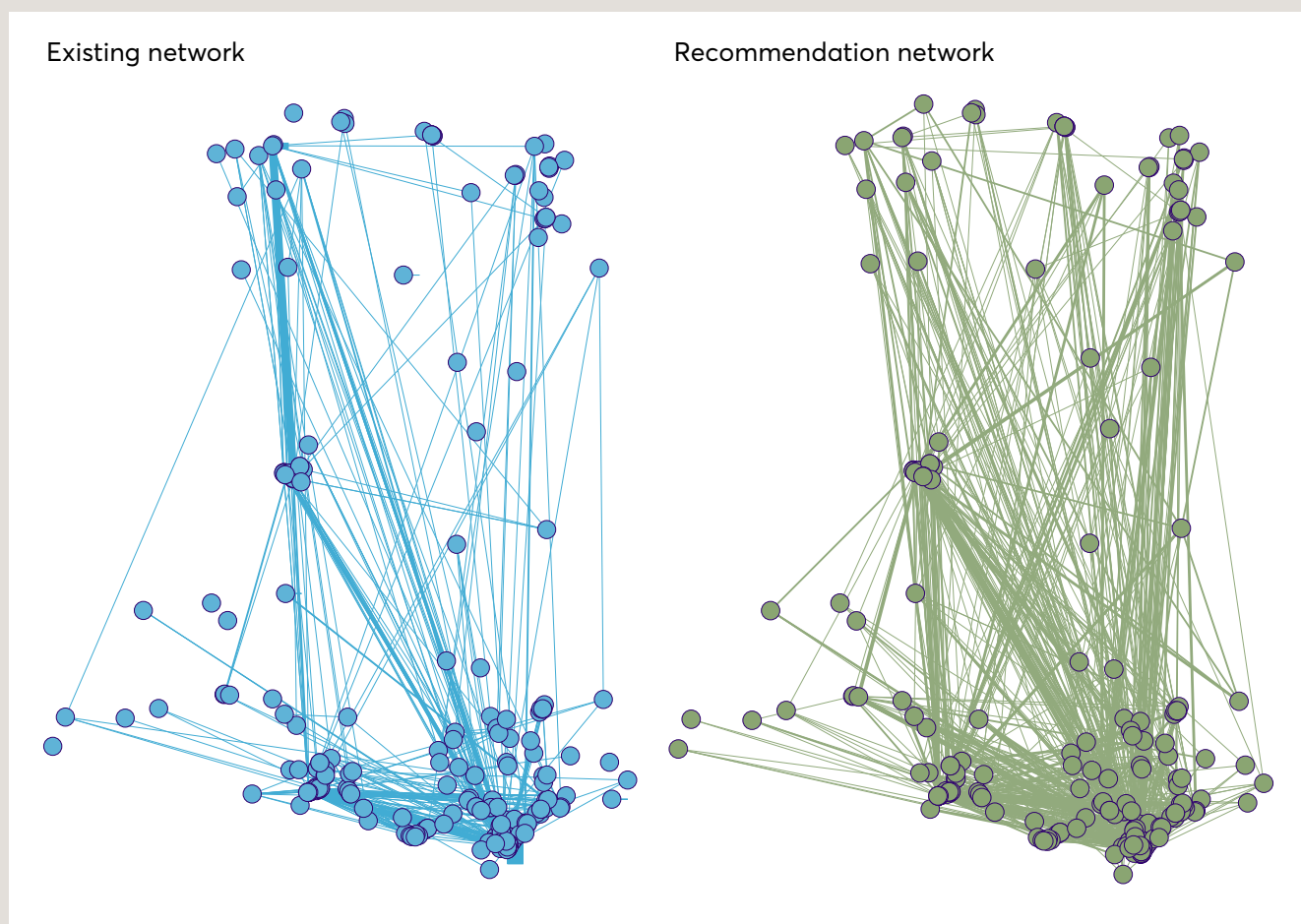


This network visualisation shows research collaborations networks in Wales based on Gateway to Research data. The image on the left shows the actual network as it exists in the collaboration data. Each node is an organisation, each tie is a collaboration. The position of the nodes represents the geographical position of an organisation.

The network in the right shows the same information, except that the ties are based on a recommendation engine that we developed which takes into account an organisation's past collaboration record. Since the network ignores an organisation's location when making a recommendation, it tends to suggest connections with organisations further afield geographically, potentially making the whole system better connected.

**Figure 4: Research trajectories in the AI for chronic diseases mission field**
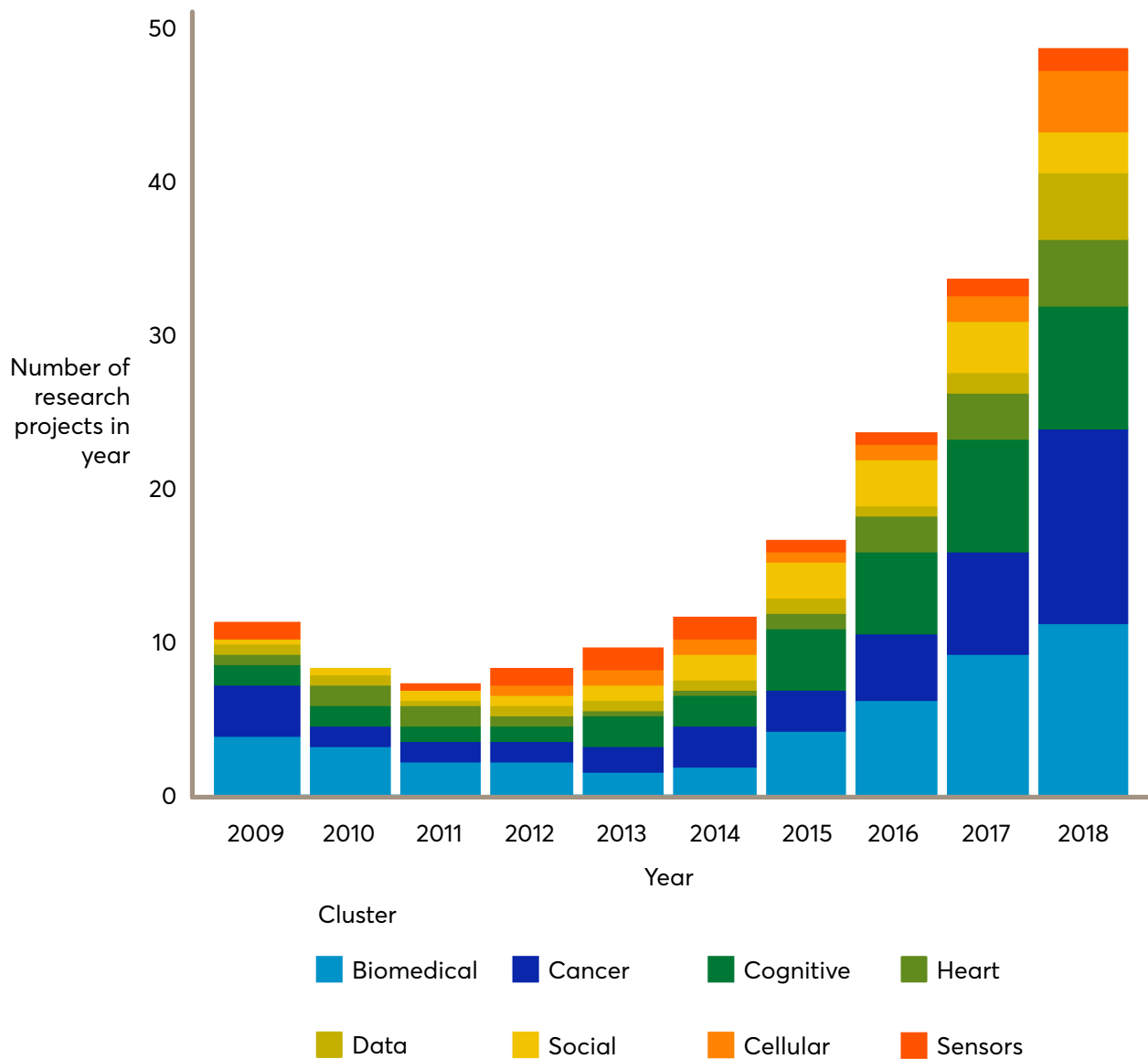


This experimental analysis shows the evolution of UK research activity in different 'components' of the mission field using AI to transform how we prevent, diagnose and treat chronic diseases. We identified each of these components through a clustering analysis that considers salient (distinctive) terms in groups of semantically similar documents. In addition to showing us the increasing level of activity in this mission field even before the UK government identified it as a grand challenge, the chart shows the composition of the mission field and how it seems to be exploring multiple paths (not all of which are technology-led) to deploy AI to tackle chronic diseases.

**Some risks and challenges**

Given all the opportunities we described above, why isn't innovation mapping already in the mainstream of economic and innovation policy?

Leaving aside the non-trivial need to develop new data collection, processing, analysis and visualisation capabilities, there are also concerns about data quality; especially regarding representativity, accuracy and interpretability, and the ethics of new methods.[33]

**Representativity** concerns run counter to claims that with 'big' data, N (the sample size of a study) equals "All". The reality is that participation in web platforms and services often have important biases. For example, Twitter is used to network in digital media and the creative industries, so its levels of adoption in that sector could be expected to be higher than, say, in biotechnology. The same point applies to other websites and services. Relying on biased data sources indiscriminately can give us a skewed view of the economy.

**Data accuracy** is also a substantial concern in a world where digital platforms are often hacked, or used to distribute spam, fake news and propaganda. Even where they are not, changes in the behaviours of their users or in their design can introduce breaks in time series that are unrelated to the activities that we want to map.[34]

**Data interpretability** challenges are two-fold.

First, the analysis of big data sources might involve working with proprietary data or algorithms that cannot be reproduced by other researchers. This lack of transparency makes it difficult to quality-assure the data (i.e. to determine whether it is representative or accurate) and can reduce trust in the findings.[35] Moreover, some of the machine learning algorithms used to analyse big data have high predictive performance but their results can be difficult to explain, again contributing to a loss of trust for policymakers who need a clear rationale for their decisions.[36]

The second aspect of the interpretability problem relates to the challenge of interpreting the meaning of complex analytical outputs such as network graphs representing relations inside innovation systems, or bottom-up, highly detailed industrial and technological taxonomies that gain depth at the expense of simplicity.

Last but not least, there is the issue of **ethics**. In a post-Cambridge Analytica world few will deny the ethical risks raised by the widespread availability of personal data and opaque, potentially biased algorithms making predictions that determine what university you go to, or whether you get a loan or not.[37] Do we want the same unaccountable, risky technologies informing economic and innovation policies that could shape the future of whole industries and regions?

# 3

# Innovation mapping at Nesta

**Until now, we have focused on the what and the why: the economic and innovation evidence gaps that innovation mapping seeks to fill and the data sources and methods it uses to do this, together with some risks and challenges that may hinder its adoption. Here we focus on how Nesta is trying to bring innovation mapping to the mainstream of policy: our mission, audiences, projects, outputs and pipelines.**

## Our approach and mission

Our ultimate goal is to have smarter and more inclusive innovation and growth policies that are fit for the future, with bigger, more widely-shared and better-evidenced impacts.
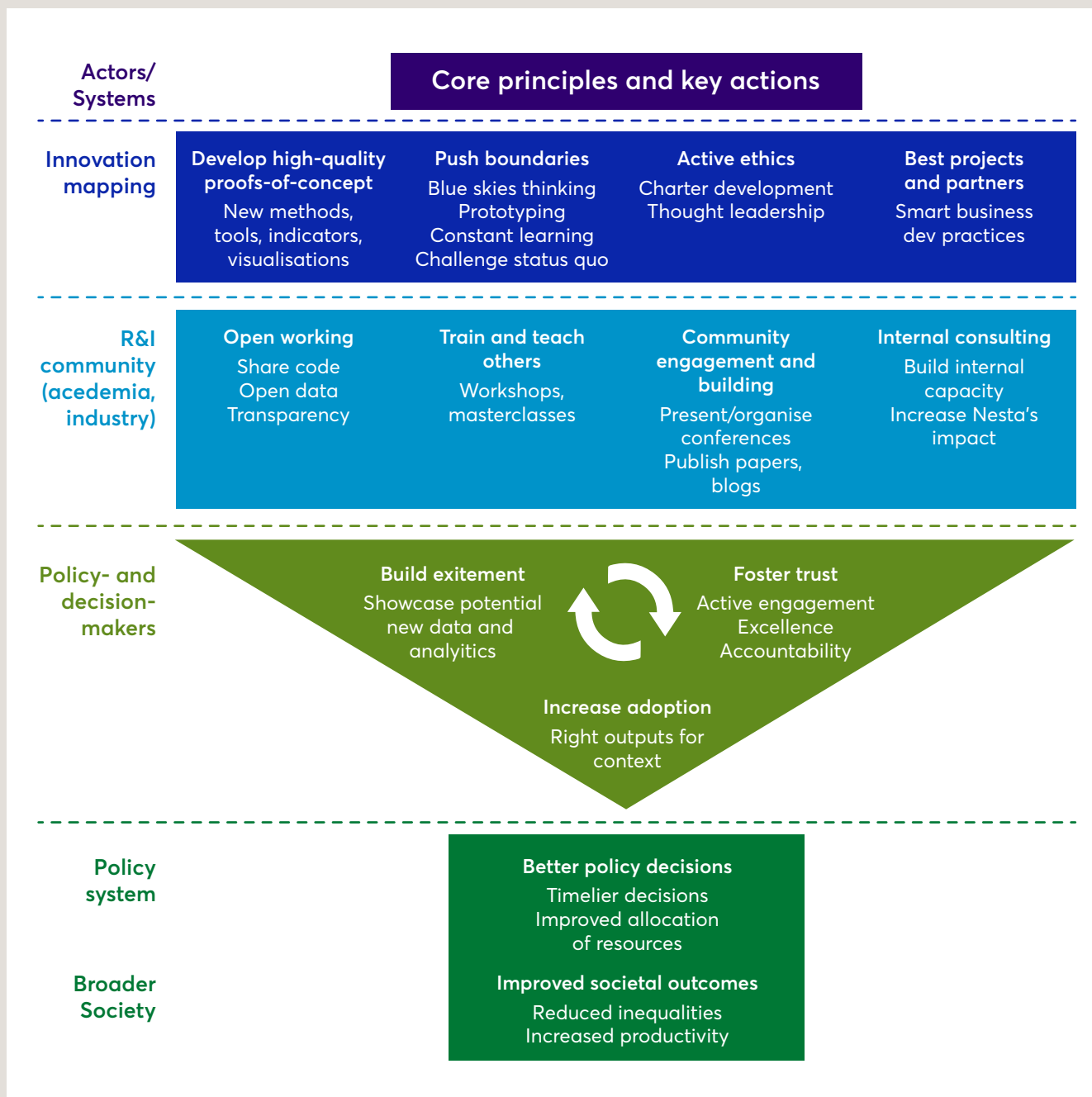
We believe that innovation mapping, based on new data sources and data combinations, analytical methods and ways of presenting information can greatly contribute to this goal, but analytics will, on their own, not be enough to deliver on this promise. This requires an approach to innovation mapping that is:

- **Highly responsive to policy needs**, avoiding the fascination with new data sources and methods that sometimes becomes the main driver for data scientists and researchers.

- **Pragmatic** in its willingness to combine new and untested data sources with traditional and strenuously quality-assured ones.

- **Eclectic**, combining the knowledge of domain and policy experts, and of multiple data sources both old and new in order to validate indicators and results.

- **Open** in its tools and results to maximise reproducibility and ensure that novel methods are suitably reviewed to maintain the highest ethical standards, and achieve the robustness required to inform policy action.

- **Human-centric**, It also needs to **recognise the importance of building 'absorptive capacity'** in policy users through capacity building activities and training.

Our mission builds on these values to transform innovation policy. We seek to:

1. **Bring innovation maps into existing policy decisions**: We create innovation maps that are relevant and trusted by policymakers, and therefore used to make better-informed decisions.

2. **Explore what new types of policy decisions might be possible**: We explore what new policy processes, programmes and ways of working might be adopted to create more value from highly-detailed and timely innovation maps.

3. **Enhance our understanding of innovation**: We encourage social scientists and relevant research communities to adopt innovation mapping data and methods, to validate our data sources, operationalise existing theories of innovation and advance new ones.

4. **Build up a sustainable data and technical ecosystem for innovation mapping**: Wherever possible, we make the data we work with open, and release all our software code with open source licenses in order to maximise our transparency, encourage reuse and develop technical standards.

The figure below presents a generalised version of our theory of change, describing core principles and actions we take, and key actors at each step.

| Actors/ Systems | Core principles and key actions | | | |
|---|---|---|---|---|
| **Innovation mapping** | **Develop high-quality proofs-of-concept** New methods, tools, indicators, visualisations | **Push boundaries** Blue skies thinking Prototyping Constant learning Challenge status quo | **Active ethics** Charter development Thought leadership | **Best projects and partners** Smart business dev practices |
| **R&I community (acedemia, industry)** | **Open working** Share code Open data Transparency | **Train and teach others** Workshops, masterclasses | **Community engagement and building** Present/organise conferences Publish papers, blogs | **Internal consulting** Build internal capacity Increase Nesta's impact |
| **Policy- and decision-makers** | **Build exitement** Showcase potential new data and analyitics | **Increase adoption** Right outputs for context | **Foster trust** Active engagement Excellence Accountability | |
| **Policy system** | **Better policy decisions** Timelier decisions Improved allocation of resources | | | |
| **Broader Society** | **Improved societal outcomes** Reduced inequalities Increased productivity | | | |

## What we do

In recent years, we have tackled complex and ambitious innovation mapping projects. Some milestones include our *Map of the UK Games Industry*, which used website data to identify games developers and studios across the UK; *Tech Nation*, where we combined official and web data to map local technology ecosystems; Arloesiadur, a path-breaking project to develop an innovation dashboard for Welsh Government; *Creative Nation*, an influential map of the creative industries in the UK; and our analysis of the *Immersive Economy of the UK*, where we used web data and machine learning models to study the size, geography and impacts of the UK immersive industry for the first time.

Currently, we are working on the following projects:

- **EURITO (EU Relevant, Inclusive, Trusted and Open indicators for Research and Innovation policy):** In this three-year, Nesta-led Horizon 2020 project with partners in Denmark, Germany and Spain, we are developing new research and innovation indicators for the EU. This includes indicators about emergent technology ecosystems such as artificial intelligence (AI), indicators for mission-oriented innovation policies and the impact of research, inclusive innovation and business investment in R&D, to name a few.[38]

- **Health Innovation mapping:** In this project, developed with support from the Robert Wood Johnson Foundation in the US and in collaboration with Nesta's Health Lab, we are building an open, web-based 'Health Innovation Scanner' mapping health innovations globally, with a focus on digital and social types of innovation that aren't well covered by existing data sources.

- **Mapping innovation in Scotland:** In this project, supported by Scottish Government, we are mapping innovation in Scotland along three themes - innovation in the enterprise sector, the situation of Scotland's Research & innovation system and inclusive innovation.

- **Mapping key enabling technologies and digital transformation in the EU**: In this project, supported by the European Commission, we are working with a consortium led by Technopolis Group to monitor the evolution of technologies that could transform EU industry.

- **Thematic streams**: We are combining all our work across different projects into thematic streams that shed light on particularly important or interesting policy questions. AI and inclusive innovation are two areas of focus for us right now.

- **Open source software**: We are also developing and launching self-standing, general purpose versions of key tools we work with such as Clio, an engine to search innovation data, and Rhodonite, a tool to explore research and technology trends.

We also play a convening role by hosting and co-organising events that bring together data scientists, policymakers and academics to discuss the challenges and opportunities in the innovation mapping space, and we build absorptive capacity in users and fellow researchers through 'hack' events and workshops.

## What we create

We produce a wide variety of outputs addressing the needs of different users, including the following:

- **Innovation maps: Arloesiadur**. We build interactive visualisations and dashboards where different users can explore the results of our analysis at different levels of granularity. Perhaps the best example of this is Arloesiadur, the innovation dashboard for Welsh Government that we published in 2017. Arloesiadur contains information about industrial, research and technology networking activity that can be explored at different levels of industrial and geographical detail to understand the evolution of Wales' economy and the structure of its innovation systems.

- **Policy-oriented reports: *Creative Nation***. We analyse the data that we collect in order to identify interesting patterns and insights, and consider their policy implications. *Creative Nation*, our 2018 report of the UK creative industries, illustrates how this works. There, we combined information from official, open and web sources to build a comprehensive view of the geography, current situation and evolution of the UK creative industries which is now informing government policy.[39]

- **Datasets: *Creative Nation* open dataset**. We launched *Creative Nation* together with an open dataset and dashboard presenting its results in an accessible and explorable format. This information is being used by many different stakeholders, from national policymakers to local governments and creative industries advocates.

- **Papers: *Deep Learning, Deep Change***. In 2018 we published a paper that studies the geography of Deep Learning, a technique at the heart of recent advances in AI. To do this, we combined academic publication data with information about startup activity from CrunchBase (a startup directory), all of which was text mined to identify relevant activities.[40] Our ultimate aim was to understand how AI is spreading into different locations and disciplines, and the factors driving this process. We published the paper together with all our code and data to enable other researchers to validate and replicate our analysis.[41]

- **Software: OmniSlash**. As part of our work, we develop software tools that we then release for others to use. An example of this is OmniSlash, an algorithm written in Python, a programming language that reduces data complexity in order to make analyses more efficient.[42]

- **Workshops: Innovation Growth Lab**. We design and deliver interactive, hands-on workshops and masterclasses to technical and non-technical audiences around the world. For instance, in June 2018 we delivered a workshop to innovation policy experts and researchers at the Innovation Growth Lab conference in Boston. Through activities such as this, we aim to build capacity, increase the uptake of novel methods, as well as to develop new project ideas and collaborations.

## Who we work with

We follow a user-centred approach to innovation mapping, placing stakeholders at the centre of our processes. The personas below are fictional but portray a selection of key attributes of our end users. They illustrate the range of relevant audiences for innovation mapping, including economic and innovation policymakers looking for better statistics, local economic developers who want detailed information about the situation in their own cluster (including about individual organisations to engage with), funders interested in the latest innovation trends, and researchers who want to develop and test new theories of innovation.

In our work, we always keep these user and stakeholder considerations in mind. This introduces helpful creative constraints, prevents self-referential design (i.e. designing for people just like us) and makes it more likely that our work will be taken up. The same persona can influence and help guide our work in many ways. For example, we may work collaboratively with 'Joseph H', the Head Economist of an innovation policy organisation, to build a machine learning model that predicts which types of technologies are emerging in a country, and then write a research note with him as our envisioned end user, focusing of his technology priorities and preferred level of technical detail. In designing an interactive data visualisation, we might expect that Joseph is interested in learning about how the data were collected and what limitations should be taken into account when interpreting the findings, and that he may want to download the data to perform his own analyses.

### Karl M., PhD
Professor Innovation Studies Germany/UK

*"I want to push the boundaries of our knowledge on innovation"*

**Sector**: Higher education

**Role**: Teaches innovation studies at two universities, supervises six graduate students, advises governments globally

**Key interests**: Productivity puzzle, local economic growth

**Use of data**: Perform cutting-edge research, consult decision-makers

### Joseph H., PhD
Head Economist Paris, France

*"I want to develop internationally-comparable innovation metrics"*

**Sector**: Multilateral

**Role**: Leads a team of five, oversees international network of innovation measurement in the digital economy

**Key interests**: Digital economy, global economic growth

**Use of data**: Advice to international partners on innovation measurement

### Cathy Z., BA, MBA
Knowledge Transfer Lead Bristol, UK

*"I want more local industries to leverage local research"*

**Sector**: Public/private

**Role**: Coordinates a UK-wide network that connects businesses and research organisations

**Key interests**: Emerging technologies, skills

**Use of data**: Identifying key actors in R&I ecosystem, understanding geographic variability in activity

### Andrew H., MPA
Policy Officer Glasgow, Scotland

*"I want our innovation policies to be more inclusive"*

**Sector**: Government

**Role**: Works closely with analysts to translate evidence into policy proposals

**Key interests**: Social, geographic and industrially-inclusive innovation

**Use of data**: Develop evidence-based policy proposals, evaluate impacts

### Emily C., MSc
Senior Policy Analyst Brussels, Belgium

*"I want to provide timely, robust insight into policy-relevant questions"*

**Sector**: Government

**Role**: Performs analysis on high-profile sector- and policy-specific issues

**Key interests**: Entrepreneurship, business growth, skills/future of work

**Use of data**: Value-for-money assessments, business and commercial analyses, R&D programme evaluations, statistical briefings

### Fatima L., PhD
Head, Brighter Futures Team United States

*"I want to support the best innovations for societal progress"*

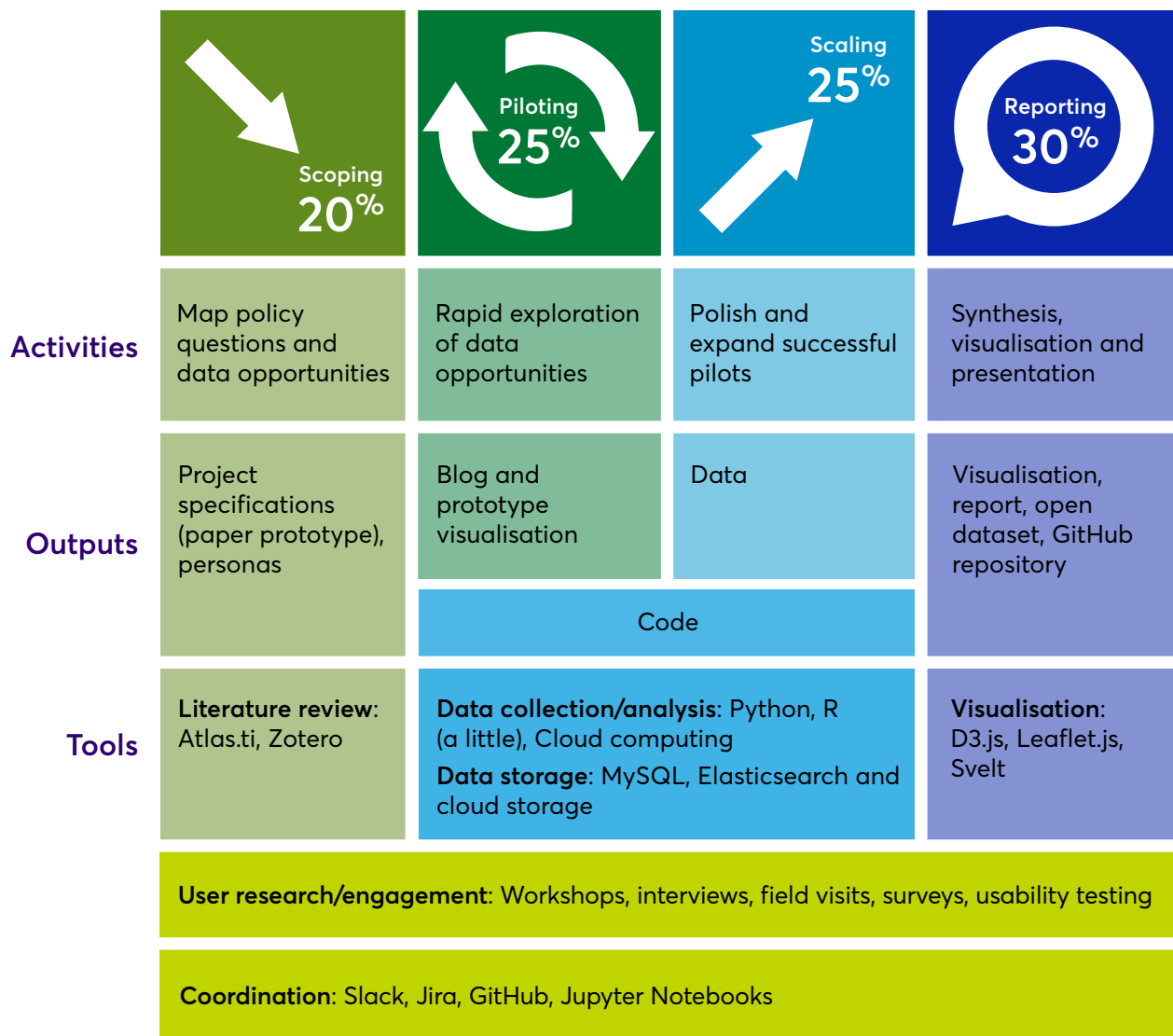**Sector**: Non-profit/third sector

**Role**: Leads team of ten in identifying and funding promising ideas, people and organisations

**Key interests**: Grand societal challenges, reducing inequality

**Use of data**: Guide funding decisions, broaden network of actors, assess emerging trends

# How we do it

Most of our projects involve four key stages: scoping, piloting, scaling and reporting, all of which are informed by user research and stakeholder engagement, and supported by a coordination function.

| | Scoping 20% | Piloting 25% | Scaling 25% | Reporting 30% |
|---|---|---|---|---|
| **Activities** | Map policy questions and data opportunities | Rapid exploration of data opportunities | Polish and expand successful pilots | Synthesis, visualisation and presentation |
| **Outputs** | Project specifications (paper prototype), personas | Blog and prototype visualisation | Data | Visualisation, report, open dataset, GitHub repository |
| | | Code | | |
| **Tools** | **Literature review**: Atlas.ti, Zotero | **Data collection/analysis**: Python, R (a little), Cloud computing  **Data storage**: MySQL, Elasticsearch and cloud storage | | **Visualisation**: D3.js, Leaflet.js, Svelt |

**User research/engagement**: Workshops, interviews, field visits, surveys, usability testing

**Coordination**: Slack, Jira, GitHub, Jupyter Notebooks

- **Scoping**: In this phase, we delve deep into the literature and engage closely with our stakeholders to determine where new data sources, methods and outputs will add the greatest value to policy-relevant questions. We uncover the 'pain points' in the existing evidence pipeline - areas which are particularly in need of fresh insights. From this preliminary work, we prioritise a set of exploratory pilots for the next phase. For example, in EURITO we drew on a policy workshop to identify eight high-priority policy questions, ranging from the lack of indicators for mission-driven innovation to the need for more timely data about emerging technologies.

- **PIloting**: In the piloting phase, we perform rapid, iterative explorations of data opportunities to see what avenues are promising and which are not viable. This stage requires an agile mindset and workflow, as well as an ability to quickly fail and move on. For example, in our project to develop new indicators for the European Innovation Scoreboard, we performed a pilot on the crowdfunding landscape using data scraped from web-based platforms. One important goal, at this stage, is to identify data limitations and risks before committing to large scale data collection and analysis.

- **Scaling**: After we have carried out our pilots, we have a better sense of which ideas are viable for larger-scale production and deployment. In the scaling phase, we polish and expand the pilots that showed most promise in the previous phase. Scaling can occur across multiple facets, for example geographic (eg moving from collecting data in one country to several) or thematic (eg moving beyond a particular technology or industry).

- **Reporting and dissemination**: We share outputs throughout the life cycle of a project, however the bulk of our output happens towards the end. This is when we have a broad range of visualisations, reports, datasets and code to share. For example, notable reports from 2018 include our *Map of the Immersive Economy in the UK* and the *Creative Nation* report.

- **User research/engagement**: User research and engagement is essential for us. We use a mix of techniques adapted from social science research, as well as design and web development, to involve stakeholders and users in our work. This ranges from workshops, interviews, field visits and surveys. For example, we developed a set of personas to help guide our work in our RWJF-funded project to map health innovation globally, following a series of in-depth qualitative interviews with key stakeholders and end users.

- **Coordination**: Strong coordination across project phases and team roles is essential to delivering complex, innovative outputs. We are continually experimenting with, and evolving, our management style to suit the changing needs of our projects. For example, some components of our work might require a more traditional, waterfall-style approach to project management with a linear sequence of activities, while other phases necessitate an agile, iterative working style to develop working prototypes that our users can give feedback on.

- **Ethics**: Many open data sources contain information that is sensitive - even if the information is already in the public domain. In the era before Cambridge Analytica, it was often reasoned that such data was 'fair game', but the ethical compass has since shifted with public opinion. Being legally compliant no longer equates to ethical accountability. We are pioneering a system for open data auditing which publicly catalogues our data sources, methodologies and reasons for applying our methodologies to data sources in each case. This is backed by an up-to-date data science ethics charter, to which we hold ourselves accountable.[43]

# 4

# What next

**We have outlined the urgent policy problems that innovation mapping seeks to solve, and how we at Nesta are trying to bring it into the mainstream of economic and innovation policy. There are many technical, analytical and organisational hurdles to overcome, but we are optimistic about the days ahead.**

We believe that the industrial and innovation policies of the future will be powered by innovation mapping. They will be designed, implemented, targeted and evaluated with much more precise and timely information than they are today, with a detailed understanding of the innovation systems they seek to transform and the complex interactions between science, the economy, society and policy.

Change will be visible not only in the types and qualities of the data and evidence used to inform decisions, but also in the types of policies that are carried out: these will be much more experimental, targeted, iterative and forward looking, using predictive analyses to understand what may happen in addition to what has already happened.

This does not mean that innovation policy will become purely data-driven, that the jobs of its policymakers will be automated, or that existing data sources and methodologies will be abandoned. We believe that innovation agencies across the world can strike the right balance between data and intuition, algorithms and creativity, and quantitative methods and qualitative ones, after much experimentation.

Ultimately, we believe that these policies will be more effective, impactful and inclusive, actively spurring the development of new technologies and industries, and their application, to tackle the big challenges of our time. This is the agenda that we are trying to advance through our innovation mapping work.

Get in touch with us at **innovation.mapping@nesta.org.uk** if you would like to learn more about our work and explore opportunities for collaboration.

# Glossary

| Methodology | Summary | Example application in innovation mapping |
|---|---|---|
| Natural language processing (NLP) | Analyse textual data to measure its sentiment and topics, quantify the semantic similarities between words based on their position in text and between documents based on their content | Analyse a company's website to identify the technologies it uses.<br><br>Measure the similarity between patents based on their text description.<br><br>Generate a technological taxonomy from the bottom-up based on what words appear together in a body of text. |
| Unsupervised machine learning (UML) | Identify similar groups (clusters) in data | Segment industrial ecosystems into a typology based on their shared characteristics.<br><br>Classify papers into different groups based on their abstracts. |
| Supervised machine learning (SML) | Train models on labelled data to predict labels in unlabelled data. Some of this involves artificial neural networks which some describe as 'artificial intelligence'. | Identify patterns that are highly predictive of an outcome of interest, like whether a patent becomes a breakthrough or a company attracts funding.<br><br>Enrich datasets with interesting labels coming from other sources, such as for example their industrial focus or the Sustainable Development Goal they are trying to achieve. |
| Network science | Analyse networks to understand their structure, and the position of nodes inside them. | Benchmark the connectivity of innovation systems<br><br>Identify communities of densely connected actors inside networks<br><br>Measure network attributes of actors (such as their centrality or influence) that could predict their outcomes. |

# Endnotes

1. Robert Solow, We'd better watch out, 'New York Times Book Review.' July 12, 1987, page 36.

2. OECD (2015). 'The Future of Productivity.' Paris: OECD

3. Brynjolfsson, E., & McAfee, A. (2014). 'The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.' New York: W. W. Norton & Company.

4. Carr, N. (2010). 'The shallows: How the internet is changing the way we think, read and remember.' New York: Atlantic Books Ltd.

5. Some key references on the productivity puzzle include Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2017). 'Are ideas getting harder to find?' (No. w23782). National Bureau of Economic Research (on the slowdown in R&D productivity); Brynjolfsson, E., Rock, D., & Syverson, C. (2018). 'The Productivity J-Curve: How Intangibles Complement General Purpose Technologies.' National Bureau of Economic Research (on productivity lags); Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2017). 'The fall of the labor share and the rise of superstar firms.' National Bureau of Economic Research (on superstar firms); Jones, R., & Wilsdon, J. R. (2018). 'The Biomedical Bubble: Why UK research and innovation needs a greater diversity of priorities, politics, places and people.' London: Nesta (on research misalignments). For a good summary, see Brynjolfsson, E., Rock, D., & Syverson, C. (2017). 'Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In Economics of Artificial Intelligence.' University of Chicago Press.

6. Arguably these policies have long been in vogue in Asia, where activist industrial policies drove industrialisation in the 1970s, as well as China. We however argue that the new wave is different from those because its focus on tackling social challenges and deploying technologies ethically instead of just achieving economic development and technological leadership.

7. Gustafsson, R., & Autio, E. (2011). 'A failure trichotomy in knowledge exploration and exploitation.' Research Policy, 40(6), 819–831, Schot, J., & Steinmueller, W. E. (2018). 'Three frames for innovation policy: R&D, systems of innovation and transformative change.' Research Policy, 47(9), 1554–1567, Frenken, K. (2017). 'A Complexity-Theoretic Perspective on Innovation Policy. Complexity, Governance & Networks, Complexity, Innovation and Policy.' https://doi.org/10.20377/cgn-41, Aghion, P., David, P. A., & Foray, D. (2009). 'Science, technology and innovation for economic growth: linking policy research and practice in 'STIG Systems.'' Research Policy, 38(4), 681–693.

8. Geels, F. W., & Schot, J. (2007). 'Typology of sociotechnical transition pathways.' Research policy, 36(3), 399-417.

9. Frenken, K. (2017). 'A Complexity-Theoretic Perspective on Innovation Policy. Complexity, Governance & Networks.' Complexity, Innovation and Policy. https://doi.org/10.20377/cgn-41

10. Cantner, U., & Vannuccini, S. (2018). 'Elements of a Schumpeterian catalytic research and innovation policy.' Industrial and Corporate Change, 27(5), 833–850.

11. Mazzucato, M. (2018b). 'Mission-oriented research & innovation in the European Union.' Brussels: European Commission.

12. HM Government. (2018). 'Industrial Strategy: Building a Britain fit for the future.' London: Department for Business, Innovation and Skills.

13. Goldfarb, A., & Trefler, D. (2018). 'AI and International Trade.' National Bureau of Economic Research.

14. Mateos-Garcia, J. C. (2018). 'The Complex Economics of Artificial Intelligence.' Available at SSRN 3294552.

15. Mazzucato, M. (2018b). 'Mission-oriented research & innovation in the European Union.' Brussels: European Commission. Mowery, D. C., Nelson, R. R., & Martin, B. R. (2010). 'Technology policy and global warming: Why new policy models are needed (or why putting new wine in old bottles won't work).' Research Policy, 39(8), 1011–1023.

16. Mazzucato, M. (2018b). 'Mission-oriented research & innovation in the European Union.' Brussels: European Commission.

17. Bakhshi, H. (2011). 'State of uncertainty: innovation policy through experimentation.' London: Nesta.

18. Bakhshi, H., & Mateos-Garcia, J. (2016). 'New data for innovation policy (Working Paper).' London: Nesta.

19. Some important examples of this include the Atlases of Science, Technology and Innovation developed by Katy Börner and her collaborators, the indicators for measuring complexity in the economy as well as the relations between different industries in the work of César Hidalgo, Ricardo Hausman, Ron Boschma and Pierre Alexandre Balland.

20. Einav, L. and Levin, J. (2014b). 'Economics in the age of big data.' Science, 346(6210):1243089.

21. Choi, H., & Varian, H. (2012). 'Predicting the present with Google Trends.' Economic Record, 88, 2-9.

22. Innovate UK (2018). 'Mapping the Immersive the Economy.' Swindon: Innovate UK.

23. Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2018). 'Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology.' arXiv preprint arXiv:1808.06355.

24. Abreu, M., Grinevich, V., Kitson, M., & Savona, M. (2010). 'Policies to enhance the 'hidden innovation'in services: evidence and lessons from the UK.' The Service Industries Journal, 30(1), 99-118.

25. Mateos-Garcia, J., Bakhshi, H., & Lenel, M. (2014). 'A map of the UK games industry.' London: Nesta.

26. Arthur, W. B. (2009). 'The nature of technology: What it is and how it evolves.' Simon and Schuster.

27. Hain, D. S., & Jurowetzki, R. (2018). 'Introduction to Predictive Modeling in Entrepreneurship and Innovation Studies.'

28. Mateos-Garcia, J., Stathoulopoulos, K., & Mohamed, S. B. (2017). 'An (increasingly) visible college: Mapping and strengthening research and innovation networks with open data.' SocArXiv Papers.

29. Mateos-Garcia, J,. Klinger, J. and Stathoulopoulos, K. (2018). 'Creative Nation.' London, Nesta.

30. Stilgoe, J., Owen, R., & Macnaghten, P. (2013). 'Developing a framework for responsible innovation.' Research Policy, 42(9), 1568–1580.

31. Korinek, A., & Stiglitz, J. E. (2017). 'Artificial intelligence and its implications for income distribution and unemployment (No. w24174).' National Bureau of Economic Research. Mateos-Garcia, J. C. (2018). 'The Complex Economics of Artificial Intelligence.' Available at SSRN 3294552.

32. Mateos-Garcia, J. (forthcoming). 'Mapping Missions. Using data science to build prototype indicators for mission-driven innovation policies (with an application to the UK mission to transform the treatment of chronic diseases with AI).'

33. Salganik, M. J. (2017). 'Bit by bit: social research in the digital age.' Princeton University Press.

34. Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). 'The parable of Google Flu: traps in big data analysis.' Science, 343(6176), 1203-1205.

35. Einav, L. and Levin, J. (2014a). 'The data revolution and economic analysis.' Innovation Policy and the Economy, 14(1):1–24.

36. Pasquale, F. (2015). 'The black box society.' Harvard University Press.

37. O'Neil, C. (2017). 'Weapons of math destruction: How big data increases inequality and threatens democracy.' Broadway Books.

38. http://www.eurito.eu/

39. Mateos-Garcia, J,. Klinger, J. and Stathoulopoulos, K. (2018). 'Creative Nation.' London, Nesta.

40. Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2018). 'Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology.' arXiv preprint arXiv:1808.06355.

41. https://github.com/nestauk/arxiv_ai

42. https://www.nesta.org.uk/blog/omnislash/

43. https://www.nesta.org.uk/blog/red-lines-grey-area/

nesta

58 Victoria Embankment
London EC4Y 0DS

+44 (0)20 7438 2500
information@nesta.org.uk
@nesta_uk
www.facebook.com/nesta.uk
www.nesta.org.uk