University of Exeter coursework header sheet

Extension

138156

BEMM466 Business Project (A, TERM3 2021/2)

1060264

700069637

Coursework: Report

Submission Deadline: Fri 15th Oct 2021 12:00

Personal tutor: Dr Rebecca Mitchell

Marker name:	N/A
Word count:	10000

2021

Analysing Public Funded Research Project for AI via GtR

BUSINESS PROJECT REPORT

Analysing Public Funded Research Project for AI via GtR

Table of Contents

Introduction

"*Success in creating AI would be the biggest event in human history, unfortunately, it might also be the last, unless we learn how to avoid the risks*". – Steven Howkins.

Thanks to raise of PC, smart phones, and big data, the computational science today is finally strong enough to host once was a dream, to replicate human like intelligence. The 21st has never been so close to give birth of artificial intelligence.

While, in HBO series and Hollywood films, AI often fall victims of paranoia and fear, the UK visioned a different future. In BBC's science fiction, *the Humans* (2015), AI are walking, human-like pacifist who were bullied by angry laid-off British working-class. The series ended in 2018, with arriving of human-robot hybrid infant, a metaphor of human and robot in harmony. The exact same year, UKRI announced over 38 million Strategic Priority Fund were awarded the *Alan Turing Institute*, home to artificial intelligence science research.

What science projects are funded today shapes the future of humanity. This dissertation explores public funding for artificial intelligence using data provided by GtR gateway to research. It seeks to answer the "who", "where", and "what" questions:

- Objective1: Who are the key players in Artificial Intelligence research?
- Objective 2: What is the UK really investing?
- Objective 3: Where is AI related research happening?

These research questions are no strangers in the Innovation Cluster and the Triple Helix literatures. However, little has been done from quantitative perspectives.

This dissertation seeks to answer these questions by exploring GtR data, which kept full recoded research projects that is public funded in the UK. Data exploration is deployed via a range of different techniques from as simple as descriptive statistics to linear regression, ANOVA, and network analysis.

A key feature of this dissertation is structured topic modelling (Blei & Jordan, 2003) which could be helpful in answering the "what" questions.

Literature Review

A Brief History of AI (see if it works like helix)

Development for modern artificial intelligence or any significant scientific innovation is, in a famous quote from Isaac Newton, "Standing on the Shoulder of the Giant".

Hoenlein and Kaplan, (2019) **review the history of AI**:

Two decades post the second ward war (1950s) witnessed a bloom of computational science and enthusiasm for artificial intelligence. This may be attributed to two persons, Issac Asimov and Alan Turnings. Asimov is an American fictionist who originated the "Three Rule of Robotics"; Alan Turning, from whom Turning test were named after. The two inspired a generation of researchers seeking creation of machines that are capable of human intelligence. By 1970, Mavin Minsky, confidently stated, while gave interview to Life Magazine, that "*a general intelligence of an average human being could be developed within three to eight years"*. However, investors, governments and even scientists themselves, sharply fall short of their high expectations to replicate true human like intelligence. Two reports criticise high spending on AI research with little progress: ALPC () and Lighthill (), one influenced the U.S. and one the UK. By 1973, the British Government ended all support AI research leaving only three universities, and soon follows the U.S. government. This put progression towards AI on halt for over a decade, known as the "first AI winter".

It was not until the 1980s did the interests for AI reignites, thanks to commercial application of the "expert system". As Fatima Hameed Khan, Muhammad Adeel Pasha & Shahid Masud (2021) explains, a program called XCON saved the company USD 40 million a year. Enthusiasm for AI rejuvenate, and investment were pulled. Japan, for instance, established a dedicated amount of funds for developing a machine that could communicate, translate language and recognising pattern. XCON is an "expert system", a fruit of research from enthusiastic era of AI, one that survived the first AI winter. An expert system is:

> "*a collection of rules which assume that human intelligence can be formalized and reconstructed in a top-down approach as a series of "if-then" statement*". (*Hoenlein & Kaplan, 2019*)

An "expert system" can make judgement. An expert cannot receive "perceptions". This leads to research in Artificial Neural networks. However, the work stagnated earlier in the 1960s when Minsky and Papert demonstrated that the computer at that time did not have sufficient processing power. Fatima et al. (2021) detailed the technical aspect of the computational constraints and how "expert system" retrieve and research of AI develops under the hood of "machine learning".

From 1990s to early 2000s, raise of personal computers market helped finance research and development for microchips. As Fatima et al. (2021) noted, the processing power has helped to host neural networks, which is the key constrains limits progression of AI research (Haenlein & Kaplan, 2019). In addition, data generated from personal computers may have aided development of AI in ways that images were used to train for visual recognition. One achievement is AlphaGo, developed by google.

what drives the innovation

Technology development of AI have gone through two bottlenecks. One is computational processing power; the other may have been data. Microprocessor to mature first to host artificial neuro networks for development of AI. But often, AI does not evolve in linear way. It evolves almost simultaneously. When AI concept were introduced (1950s), it was same time when foundation of microprocessor, semi-conductor transistor, were invented (1940s by Bell Lab). At the time, the two may not seems necessarily connected. This highlights the importance of diverse findings for diverse research.

The force, "invisible hands" of the economic market has played vital role in pushing (and hinging) the development of modern AI: Research in AI has gone ups and down because the research investment went ups and down. Research investment went ups and down because the stakeholders, the government, the investors, the people's expectations for AI went up and down.

Gartner's hype cycle of technology adoption (Gatner, 1979; Steinert, & Leifer. 2010) descopes this pattern of innovation: investor, scientist tends to hold inflected expectations of certain technology (i. Innovation Trigger; ii Peak of inflated expectation);The enthusiasm of investors seems recess when science meet a bottleneck of a certain era, as result research were halted (iii. Thought of

Disillusionment). It revived when a private firm can commercialise scientific research into scientific product (iv. Slope of enlightenment). In the case AI, both expert system and neuro-network were matured in the house of private sector. For instance, XCON in __ and AlphaGo in Google. Science gets to further develop within those sectors because the commercial value it makes can circulates back into research and development for that field.

The Triple Helix and Innovation Cluster theory

The western innovation model frequently refers to the Triple Helix model (Etzkowitz and Leydesdorff, 2000) or Innovation cluster (Engel, 2015). Both have strong Silicon Valley origin. The core component of both is in fact the similar. Both emphasis the collaborative role of government, universities, and the industry.

The word "Helixes" took reference from the shape of DNA (think of the word helicopter). The government, business, and research institutes are three "helixes" in the model, which implies that the government, business, research institute entangled together. Such a collaborative approach highly resonates with *political pluralism*. Further there is a growing recognition of adding the general public to the helix. This is termed as *Quadruple Extension* in the literature *(Cai & Etzkowitz, 2020; Monteiro & Carayannis, 2017).* Social agents such as consumers, users, non-for profit have been unformidable forces in driving innovation *(Ivanovo, 2014; Arnkil et al. 2010; Miller et al., 2018; as cited in Cai & Etzkowitz)*. Again, those novel notions highly resonate with political pluralism and democratic sentiment, which congruent with values in western society. Such value congruence may be why the helix model highly recognised in Europe: The European commission (2016) officially recognise the triple helix model so does *the World Bank* and *OECD* (as cited in *Cai & Etzkowitz*, *2020*).

However, Pluralism, and democracy is not always necessity of innovation. In different cultural context, the process of science discovery unfolds differently. For example, in China, central government took control of innovation process as innovation ecosystem mature. The UK once deploy similar innovation model during World War II. One famous outcome is Alan Turing's decrypt German's secret communication. Innovation them were organised and led by the military. Etzkowitz & Leydesdorff (2000) refer this as the statist-model of innovation, similar model has once been deployed in the former Soviet Union, France, and in the US during World War II.

"Statist" model contrasts a more libera model which is called "Laissez-faire model". In this model, the three spheres are independent from each other, each maintains their own autonomy in research. Cai & Etzkowitz (2000) note the contemporary Silicon Valley is a reflection such model, with large corporate such as Apple, Google and Facebook overshadow universities and government. Social ideology is very libera, as large company compete intellectual resources with research universitas. Because those corporates are highly successful, they often have sufficient resources to attract and acquire the brightest talent and is sufficient in funding their own R&D.

Working in industrial environment is different from academia environment. Industrial is more value driven, fast phase, which in a way counterproductive towards innovation as research shows relaxed environment tends to encourage people to be more creative (Shin & Grant, 2020). Such tension between academic and industry is illustrated in Sweden (Fogelberg & Thorpenberg, 2012).

Countries endorse triple helix model of innovation often looks up to Silicon Valley model without context. What is called "mimetic isomorphism" (DiMaggio & Powell, 1983; cited by Cai, 2015) occurs when a structural element is copied without context. Science discovery rarely emerge in a linear way. No single entities: no one government, business, or researcher can govern the direction of science innovation themselves. As Cai & Etzkowitz (2020) explains:

> "*The innovation itself is not a self-organised evolution but a pre-structured process*" *(in Cai & Etzkowitz, 2020).*

That is, innovation is a result of casual determinism of collective action. One cannot control the process of innovation, but it is possible facilitate its emergence. Hence the one shared goal among three spheres is to create innovation ecosystems.

In Europe, a common strategy is to concentrate funding in area that has high potentials. As Cai (2015) summarised:

> *some public funding has been used to provide incentives and to promote university and industry cooperation for knowledge generation and knowledge transfer. Nowadays one common governmental strategy is* ***to concentrate*** *funding on supporting research and development*

(R&D) ***in fast-growing or high-potential areas*** *in the expectation that the investment will eventually pay off through economic growth led by the companies that benefit from cutting-edge knowledge.*

Parallel Research

Few peer-reviewed journal article analysis public funding related to artificial intelligence from quantitative perspective. In terms of what are UK investing, Dunham Melot & Murdick (2020) derived four area for fundamental AI research by analysis scientific publication: Natural Language Processing, Computer Visioning, Robotics, Machine Learning, the most heavily researched field is said to be computer visioning according to Rahkovsky et al (2021).

UK claims to be world leading in AI research and innovation, mathematical and computational science and host a third of Europe's AI start-ups (UKRI, 2021) but there is a need to transform research into economic values. UKRI took reference from McKinsey & Company (2018) who estimate that utilising AI can gain 20 – 25% economic growth in the next decade. Since, UK has dedicated 38.8 million *Strategic Priorities Fund* to Alan Turing Institute (reflected in GtR data). One of research program is *AI for Science and Government*. *AI for Science and Government* is delivered in six themes: Digital Twins (Urban Analytics and Complex Engineering Systems); Health; Criminal Justice System; AI for Science; Tools, Practices and Systems (Alan Turing Institute, 2020).

Data city[1] (Forth, Connell, & Laflin, 2018) draws data from open-source data and published interactive map for artificial intelligence and data. The map explores innovation cluster in the UK. Innovation score is measured by business activities, events and papers published in each region. Summarising findings from Data City (2018):

- Bigger cities in the UK tends be rank better in their matrix
- London is scored highest aggregation all three activities.
- Birmingham, Manchester and Reading ranks follows but compares to London they only count less than half of the score.

However, it is not clear if the data had considered value awarded to researchers when counting paper published in each region. Further, student dissertation can be considered counted as "paper published". It is questionable if they can account of research activities for each region. Further Although rankings differentiate from others, the true statistical means may not be different.

Data and Methodology

Data

Data - GtR

UKRI GtR host information regarding project and funding in both downloadable csv format and json which can be accessed via API (java notation format). Only projects with classification tagged "Artificial Intelligence" were selected.

There is total 836 projects in the downloadable csv data, 35 pages in 25 projects per page. No project abstract or project outcome is included in this the original dataset. This information is, however, can be accessed via GtR API schema.

Project abstract (content for topic modelling) are extract via GtR api by writing code in Python. Although the GtR's API schema restrict user from access more than 100 per page, this project has been managed to acquire all data page by page. Research funding usually takes into consideration of research duration. Shows that research duration seems to be left tilted. The mean of funding for AI project is about 600 pound per day, but this is not a very good representation considering extreme values in this distribution. Compares mean, the median daily award for one project is only 300 pounds per day. That is, 50% of the time AI funding has daily value that is below 300 pounds.

[1] Data city's research interest is very similar to this dissertation. However, their report is not peer reviewed jornal article. Their finding aviable:
http://thedatacity.com/products/uk-tech-innovation-index-2/?options=true&datagroup=All%20five%20technologies&location=null

The project dataset contains some extreme outliners. Those outliners are extreme cases that should not have been taken into considerations. The approach here is to cap *price* at 10000. Calculation of "price" is explained later in the same chapter.

Additional Data

One conquest of this business project is to find out where is UK's AI innovation cluster. The approach used here is to pin university where they are on a map. R-library `gg-map` (Cambon et al., 2021) provide functions that automatically find coordinates of a given address. There is a verity of options regarding what geocode search engine the algorithm should use to extract geocode. This report uses ArcGIS as an alternative to Google.

The free-to-used geocode service does come with its limitation. Several UK research institutes/Universities were mislocated. One reason is because algorithm confuses location names that were shares same name. For instance, "The National Archives" in London (the UK) were confused with one in Maryland (the U.S.). "Leeds Beckett University", previously known as "Leeds Metropolitan", were confused with their oversea brunches. A total of 9 entries had to be amended manually to correctly pinpoint their location on the map. Those locations are assumed to be within the UK. However, not all leading research institute funded by UK government is in the UK. Two leading research organisations are located overseas. Those two universities are "*BC3 Basque Centre for Climate Change*" (Biscay, Spain) and *Delft University of Technology* (Amsterdam, Netherland).

Methodology

Exploratory Descriptive Statistics

The research council will determine how much fund is allocated to a project. Hence one goal of this business project/dissertation is to find out the area of projects that are at the interests of research council. As known from literature review. The specialist council took consideration of research durations of a project. Arguably, the total amount of fund allocated to a project could hardly reflect true value of the project. For instance, one research project could be funded heavily only because it has a long duration. Hence our approach is computing the unit "*price*" of a funded project, that is dividing total amount of award by its duration:

$$price = \frac{pounds\ awarded}{duration\ of\ the\ project}$$

Such indicator can reflect true value of the project. This assumes a relationship between project duration and award value. To verify such assumption, linear regression is used to evaluate use of this matrix.

Analysis of Variance & Effect Estimate

Objective 1 asks key players in AI research. Analysis of Variance (Chamber, Freeny & Heiberger, 1992) have been used to verify if price differences between each research institutes are truly significant. This is established by comparing group statistically parameters with statistics of university parameters. A clear advantage of ANOVA instead of just comparing university research fund rankings is when certain research organisations have larger variations. In *R-programme*, this is established by linear regressions.

The method cannot be used to compare in-between individual universities. To compare if one university is truly different from the other one can apply Tukey Procedures. However, as outlined in discussion, the data is not suited for Tukey method to achieve research objectives in this dissertation.

Topic Modelling

The abstract part of data is modelled via **Structured Topic Modelling** in R package `stm` by Roberts, Stewart and Tingley (2019). The `stm` package in R provides method to join topics matrix with meta data, hence facilitated the possibility for this project to explore which topics are associated with higher cost or longer durations.

The method chosen here is what is called Latent Dirichlet Allocation (LDA). There are a few alternatives to LDA. Such as, the LSA (Latent Semantic Analysis), pLSI (Probabilistic Latent Semantic Index), CTM (Correlated Topic Model). Chang et al. (2009) conducted human-based experiments to evaluate interpretability amongst those four methods. LDA outperform the other three in both interpreting words and interpreting topics.

There are several limitations regarding LDA. The most prominent here is that LDA could yields different results. Technically, this is because LDA is "*multimodal*" thus "i*nitialization*" is different each time repeated (Roberts, Stewart & Tingley, 2016; pp. 51 - 97). To put this less technical terms, the algorithm requires randomly sampling, and each time the algorithm is repeated, different sample will be draw. The sampling process, for ease of understanding, hereby is so-called "*initialization*". In computational science, there is a spectral of from run an algorithm repeatedly several time, to run it once and right. For LDA it is the case of the later. This is because LDA consumes a lot of computational power. For this project, as recommended by Roberts et al. (2016), *spectral initiation*, is set so save results can be generated, hence makes this project repeatable.

20 topics has been extracted. A few words have been removed from document matrix other than normal English stop words, punctuation, and numbers. Those includes "develop", "research", "will". Due to the fact they appear in document too frequently.

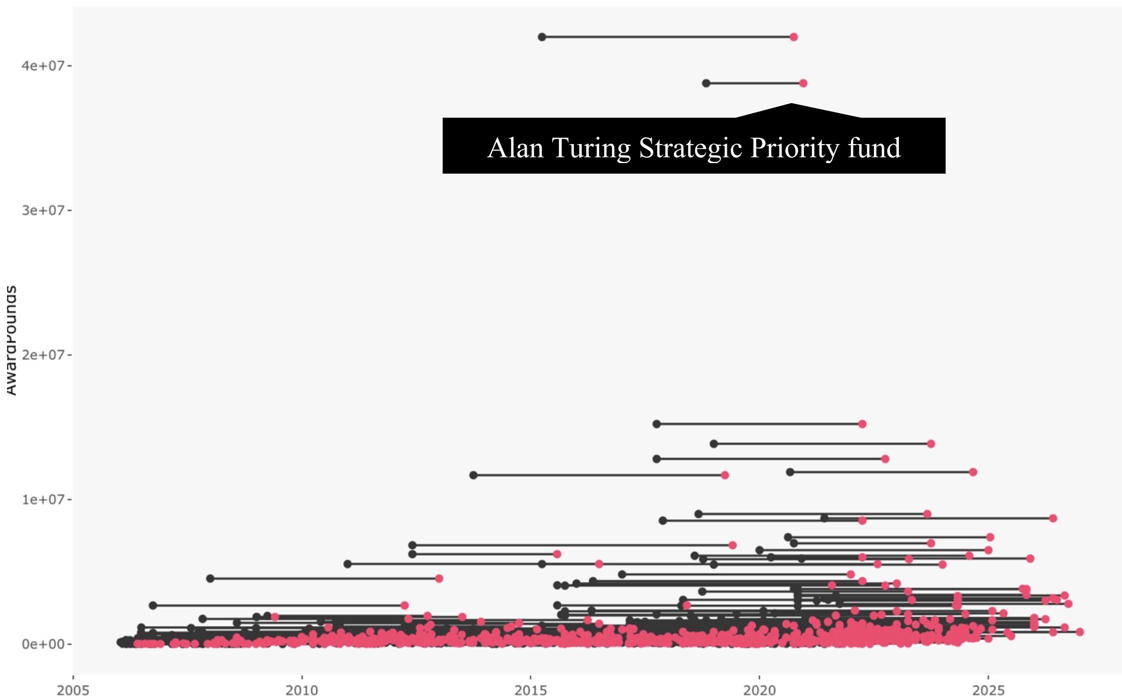
Network Anlayis/Visualisation

Because we can pair transferred project together assembles into an edge data table, with project ending with "/1" being the source, "/2" being the sink. Node of those network are leading research organisations, linked by projects that was transferred.

Basic network analysis is conducted by looking leading organisations with highest edge linked to them, and by looking net inflow, that is, for each node, total number of flows into node deducted by total number of flows out of the same node. The findings are reported in further findings.

Findings

Overview



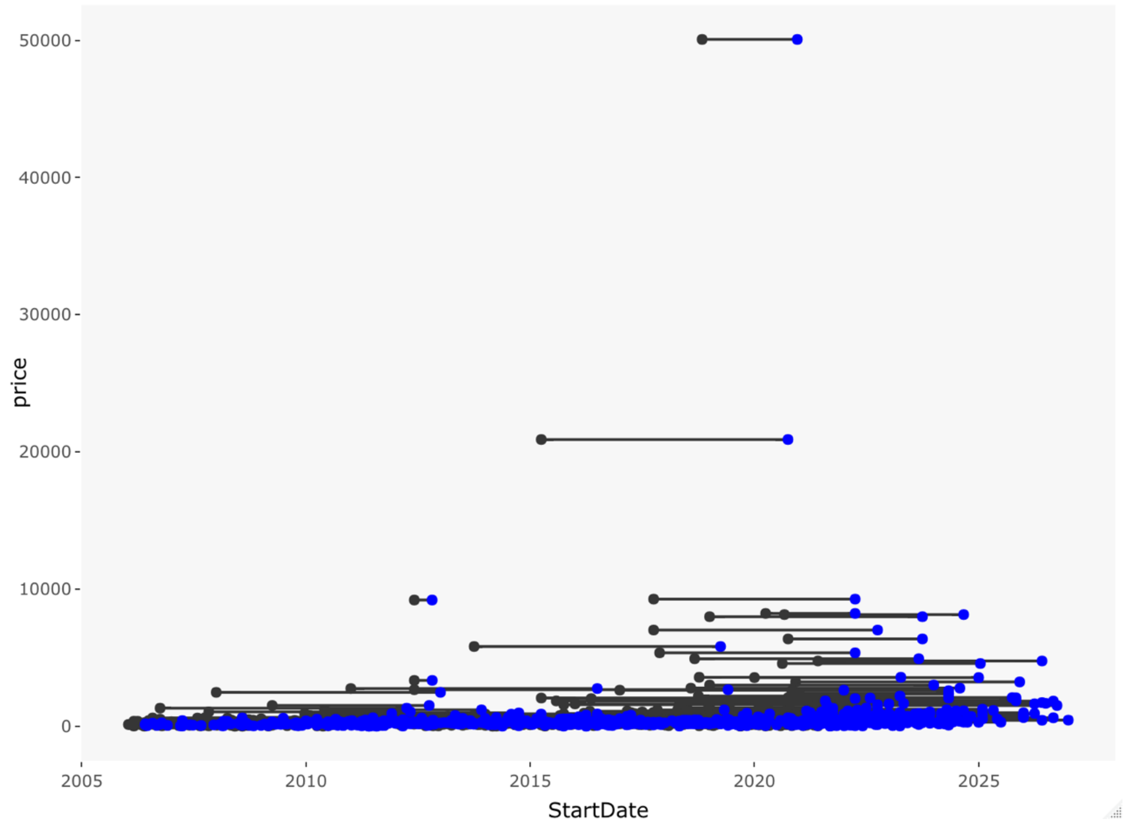


Figure1.1 (top): an overview of project by total award. Figure 1.2 (left): by price.

It is easy to see one or two research projects are more favourable than the other. Top the two highest lines are respectively funding towards Alan Turing Institute, and *the strategic priority fund*. Alan Turing also funds its research partners. It is possible that some of the research projects are sub-funded. However, by looking at GtR data itself, this is not the case.

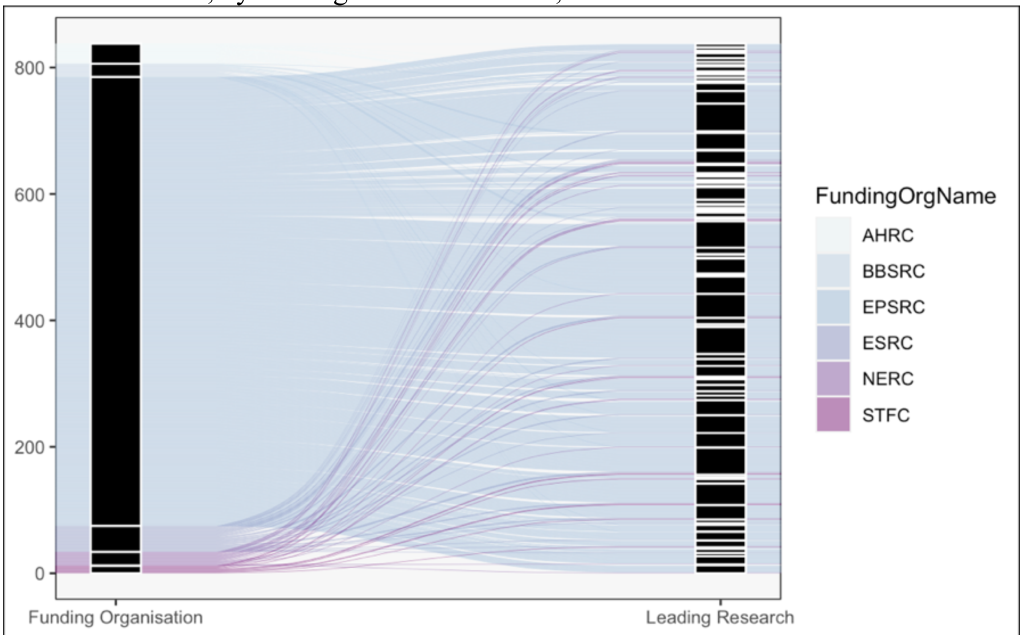


Figure 2 Flow chart of funding visualise by paring funding organisations (left) to leading research organisations (right). Height of a single block indicate value of

All funding comes from the six specialist research councils. No funding comes from any specific organisation.

All funding is almost exclusively coming from EPSRC. Organisations that are most funded are almost exclusively funded by EPSRC. It makes sense as artificially intelligence is a discipline within engineer degree.

Validate duration-award relationships

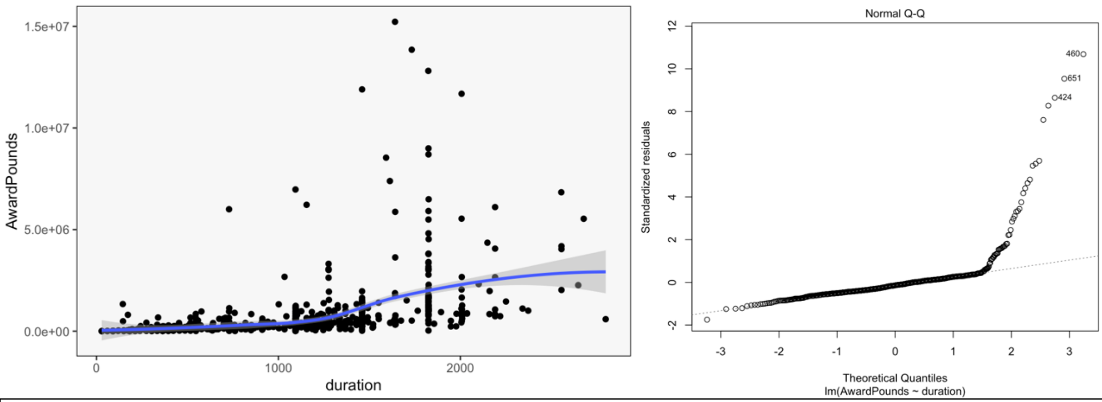


Figure 3. The relationships between project duration and amount that has been funded is linear within certain range. The variability increases reach a certain range.

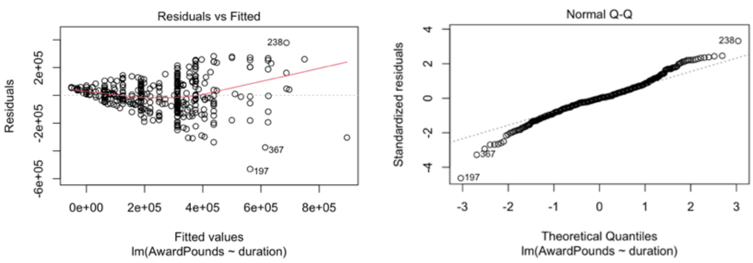
One primes of using price as a Metrix to filter out the effect of project duration is price is linear associated with duration of the project.

Linear model has been used to test this assumption. Average project duration is about more than three years (5% trimmed mean = 1053 days, 1053/365*12/11 = 3.14 years, taking consideration of 1 month holiday per year) with almost two-year variation (sd = 517.2869 days) across the data. Establishing linear model, project duration explains about 20% standard deviation of amount awarded (Adjusted R-squared = 0.2219, p < < 2.2e-16). Suggesting that duration of project affecting amount awarded.

Quantile-to-quantile plot shows that linear model performs well only at first few quantiles but poorly at lower quantile. By trimming top 30% quantile of data, adjust R-square improves by another 40% (adjusted R-square 0.66). Corresponding the fact that data are heavily tailed.

Judging by R-square, best model amongst logarithm scale, square-foot, lambda transformation, is square-root transformation (adjusted R-square = 0.6869), closely followed by lambda transformation (Adjusted R-squared = 0.6833). Log transformation yields lowest R-square (Adjusted R-squared = 0.5999).

Regarding the four assumptions of linear model, heterogeneity assumption is slightly violated. However residual significantly improves after transform into normal ().



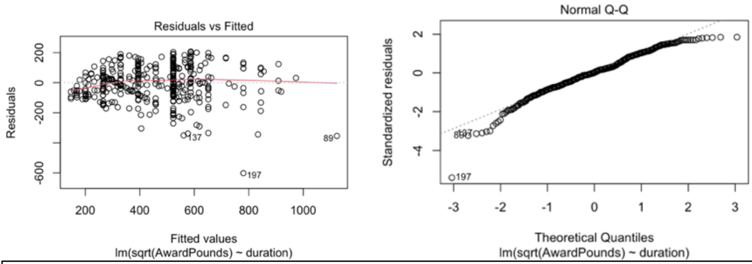


Figure 4. Homogeneity assumption is violated in both models. That is value subjects to larger variance when fitted values are higher. The case is more severe in simple non-transformed linear model than in transformed via square-foot transformation.

Ranking

The following chart explores the most heavily funded research institute by looking at total funding received and average award value per project.

Figure

```
# A tibble: 93 x 6
   LeadROName                       sum  median     mean        sd     n
   <fct>                          <dbl>   <dbl>    <dbl>     <dbl> <int>
 1 The Alan Turing Institute  109808696 6694380 13726087 16704452.     8
 2 University College London  47023890  434113  1119616.  2428192.    42
 3 University of Bristol      36911989  307814.  1318285.  2529423.    28
 4 University of Southampton  36563933  489242  1589736.  2755087.    23
 5 University of Oxford       32309089  480411   751374.   975533.    43
 6 University of Manchester   27961959  403168.  998641.  2360549.    28
 7 Imperial College London    26835128  396784   609889.   539080.    44
 8 Heriot-Watt University     24747214  287812.  1237361.  3362350.    20
 9 University of Surrey       23038077  668298.  1645577.  2275181.    14
10 University of Sheffield    20962724  249266   616551.  1094684.    34
# … with 83 more rows
> t %>%
+   mutate(ranking = 1:length(LeadROName)) %>%
+   mutate(ranking = 1:length(LeadROName)) %>%
+   filter(str_detect(t$LeadROName, "(?i)(Exeter)"))
# A tibble: 1 x 7
  LeadROName              sum median    mean      sd     n ranking
  <fct>                 <dbl>  <dbl>   <dbl>   <dbl> <int>   <int>
1 University of Exeter 4828503 269363 438955. 354561.    11      27
```

```
# A tibble: 93 x 6
   LeadROName                        sum  median     mean        sd     n
   <fct>                           <dbl>   <dbl>    <dbl>     <dbl> <int>
 1 The Alan Turing Institute   109808696 6694380 13726087 16704452.     8
 2 Newcastle University         12699908 1537424. 2116651.  1608326.     6
 3 University of Surrey         23038077  668298. 1645577.  2275181.    14
 4 University of Southampton    36563933  489242  1589736.  2755087.    23
 5 University of Bath            7599238  794580  1519848.  1726764.     5
 6 Royal Veterinary College      1433975 1433975  1433975        NA      1
 7 University of Bristol        36911989  307814. 1318285.  2529423.    28
 8 University of York           16639808  379941  1279985.  1850970.    13
 9 Heriot-Watt University       24747214  287812. 1237361.  3362350.    20
10 Queen's University of Belfast 7919495  403635  1131356.  1954771.     7
# … with 83 more rows
> t %>%
+   mutate(ranking = 1:length(LeadROName)) %>%
+   filter(str_detect(t$LeadROName, "(?i)(Exeter)"))
# A tibble: 1 x 7
  LeadROName              sum median    mean      sd     n ranking
  <fct>                 <dbl>  <dbl>   <dbl>   <dbl> <int>   <int>
1 University of Exeter 4828503 269363 438955. 354561.    11      31
```

Figure 5.1 (left), by looking at total amount of awarded funded to the organisation; Figure 5.1 (right), by looking at the average amount of award funded per project.

The most heavily funded research institute is the Alan Turing Institute. The Alan Turing Institute, situated in central London, received over fifty thousand per day's Strategic Priority Fund in 2018 (). University of London appears absorb large amount of funding. The reason is because large amount of project that were taken in London (42 from recorded time). A few new universities emerge from the ranking list. For example, University of Bath and Newcastle University, both are not part of the Russell group but on average, both performs secure funding good funding. Surrey and Southampton, both outside close to London, stable both in terms of total funding secured and average, suggesting London may be a potential AI innovation cluster of UKS. This finding collision with the fact London being the finance hub of UK but our findings suggest AI clusters to be outside London. Part of Research Funding allocation process is to justify London pricing to the funding. This does not seem to affect the analysis very much by just looking at the top three rankings. UCL for instance, the project award went down the ranks by looking at average amount. This could potentially reflect competition. Outside London may not be the only AI innovation clusters in the UK. Marks that university that are close to city regions top the list, for instance, Manchester, Belfast. Following section map's location of research organisations in an attempt to reveal spatial patterns of AI-related public funded project.

Spatial visualisation

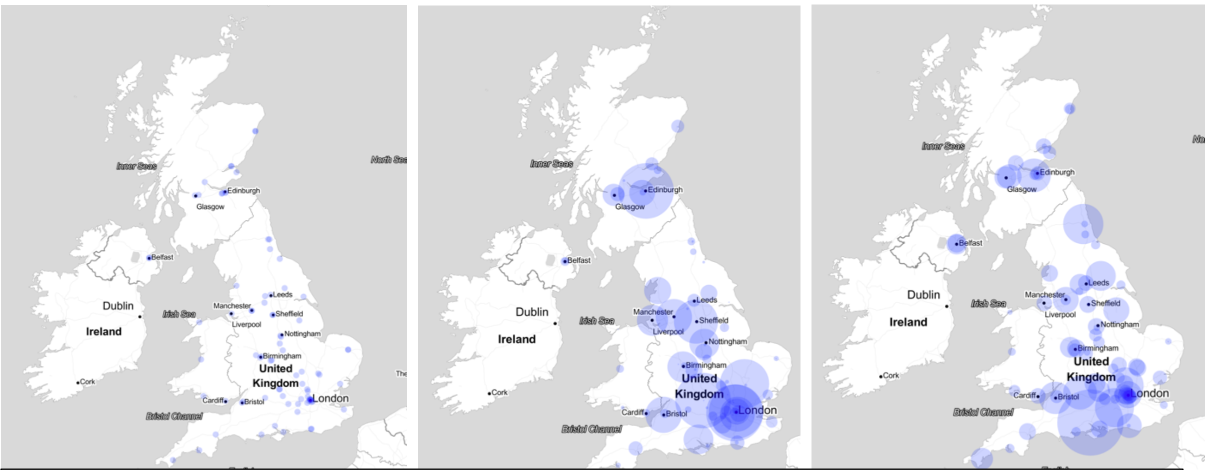


Figure 6.1 (Left): all research institute that conduct Ai-related research; figure 6.2 (middle): by looking at number of funded projects; figure 6.3 (right): by looking at "price" of funded project. The radius of the circle indicates the number of projects that has been led by a research institute. Note that all following analysis use data that has caped at a "price" of 10000 because of the extreme values.

As it shows, the golden triangle, London – Oxford – Cambridge is still solid in terms of amount of research output. But London advantages in terms of number of research provided. This is because in London there are more universities were being funded. Interestingly, "price" in London is actual cheaper by just ignoring the duration of funded project even without adjusting London Pricing. Potentially, this could be attributed to market competition. Research councils may be forced to balance the funds allocation in different regions but because there are many universities in London, the actual funds the universities receive decreases. It is unknown if those London research institute had received further private funding from companies as a compensation.

One university that stands out is University of Southampton. The university ranks top in terms of number of AI-related funding project. On average, each project could be awarded over 1800 k per day, stands out from the rest of the UK.

There are only two UK-funded projects that are led by foreign organisations. They are respectively led by *Basque Centre for Climate Change bc3*[2] and *Delft University of Technology*[3]. Both have UK research institute participance. Both are amongst lower quantile in terms of price. Both are research project in social disciplines. The one led by *Basque* is valued at only 80.1 pounds per day and the one led by *Delft* is only 97 pounds per day. According to GtR, the *Basque* project aimed towards poverty alleviation. The *Delft* project studies how AI can integrated into workplace.

ANOVA and Linear Effect Estimate Model on organisation and price

Although ANOVA is robust against abnormal data, ANOVA assumes tested sub-group have equal variance (homogeneity assumptions). This pre-requisite can be tested using Levene test. Levene test shows that variance between each group in significantly unequal (p = 1.714e-07). Hence normalise data by subject them to boxcox transformation and squar-root transformation. This is informed in previous section that square root transformation works best in normalise price. Again, use levene test on transformed data:

Normalise via boxcox transformation (lambda = -0.1) result in p-value = 0.09889; via square root transformation result in p-value = 0.5364. Suggesting square root transformation works best in reducing heterogeneity of data.

---

[2] Basque Centre for Climate Change: Which Ecosystem Service Models Best Capture the Needs of the Rural Poor. GtR link: https://gtr.ukri.org/projects?ref=NE/L001195/1#/tabOverview

[3] Delft University of Technology: Learning to Efficiently Plan in Flexible Distributed Organizations. GtR link: https://gtr.ukri.org/projects?ref=EP/R001227/2#

***Table 7.1**: statistically significant lead research organisations by price*

term	estimate	p.value	n
Newcastle University**	22.23578	0.003177928	5
University of Surrey**	19.14079	0.002906627	11
University of Southampton**	18.76404	0.001852319	18
University of York**	18.25945	0.008821391	7
University of Bath*	18.10131	0.023325551	4
Heriot-Watt University*	13.84464	0.024226219	15
University College London*	12.84829	0.027474806	25
University of Bristol*	12.79507	0.028664720	24
University of Manchester*	12.73202	0.035356460	17
(Intercept)	*12.49189*	*0.018905095*	

** at 99% confidence interval; *At 95% confidence interval table shows via ANOVA after square root transformation.

ANOVA and Linear Effect Estimate Model on organisation and duration

At 99% confidence interval, only Newcastle, Surrey, Southampton, and York stand out from the rest. The rest only significant at 95% confidence interval. Note Southampton was ranked amongst top by project price and third by lead total award received but went down after normalise data. This implies Southampton went up the list may be a result of a few project that receive high award.

***Table 7.2**: statistically significant lead research organisations by project durations*

term	estimate	p.value	N
Newcastle University**	1106.6000	0.0009982306	5
Queen's University of Belfast*	847.0000	0.0115875271	5
University of Strathclyde*	743.2857	0.0166840591	7
University of York*	613.0000	0.0481777904	7
University of Surrey*	565.9091	0.0476842763	11
University of Manchester*	563.7647	0.0365361395	17
University of Glasgow*	550.8462	0.0482050766	13
(Intercept)	747.0000	0.0016708770	*NA*

Newcastle is the only research organisation significantly longer than others at significance level of 99%. Queen’s University of Belfast, Strathclyde, York, Surrey, Manchester, and Glasgow is significant at 95% confidence interval. The other research organisations are not distinguishable from group mean.

Note that projects led by Alan Turing institute has been removed from this analysis. The reasons are disclosed in discussion section.

Project Topics

20 topics are extracted. For full list of topics refer to odd page table. Label were given for each topic number. Note this is entire based on human judgement. For reference, estimation results are reported in topic numbers so if reader do not agree with a specific label, they can refer to this page.

Topics include robotic automation (), neuro-networking (), and natural language processing () three of four major research frontier for Artificial intelligence (). The fourth visual image is harder to identify among those topics as no topics has explicated includes “visual”. Highest proportion is topic 6, which maybe about digitising industries, which is in line with strategic agender of the government. The lowest topic is topic 13.

Estimate Effect

Only topic 6 (Financial Hub) is significantly associated with both price (p-value = 0.00384) and total amount of award (p-value = 0.000112, at global uncertainty). Further, the significance increases when using sqrt transformation. Topic 6 is also positive significantly associated with project duration (0-valye = 0.0018). That is project with topic 6 content tend to be longer.

***Table 8.1*:** variables that significantly associated with topics

Label	topic	Variable	estimate	statistic	p.value
Industrial collaboration	6	AwardPounds	1.461785e-08	3.903355	1.057551e-04
Industrial collaboration	6	duration	5.261786e-05	3.041227	2.460212e-03
Art and Recreation	7	duration	-6.906311e-05	-3.982341	7.671726e-05
Industrial collaboration	6	price	1.161988e-05	3.040449	2.466483e-03
Industrial collaboration	6	sqrt(price)	3.245849e-03	5.545333	4.422384e-08

At global uncertainly. All variables are at 0.99% confidence interval

For technique reasons, the algorithm is unable to test association between topic proportion and leading organisation. As the unstable result were yield.

Associate time with topic proportion, the algorithm can test topics tends. Proportion of topic 6, 8, 19 positively associated with time, while topic 20, 2 and 11 is negatively associated with time.

***Table 8.2*:** variables that significantly associated with topics

label	topic	Variable	estimate	statistic	p.value
Industrial Collaboration +	6	date	1.701374e-05	3.370655	0.0007986458
Urban Twin +	8	date	1.288501e-05	3.130839	0.0018288558
Chemistry +	19	date	1.070715e-05	2.677932	0.0076130905
Ontology -	20	date	-9.474094e-06	-2.707223	0.0069804161
[4]*Data Science -*	2	date	-1.588167e-05	-3.480557	0.0005371376
Evolutionary Computation -	11	date	-1.729653e-05	-3.225799	0.0013253718

[4] It is difficult to distinguish what topic 2 really are. As can be seen from figure 10, topic 2 and topic 11 are very close. Data science is label due to more

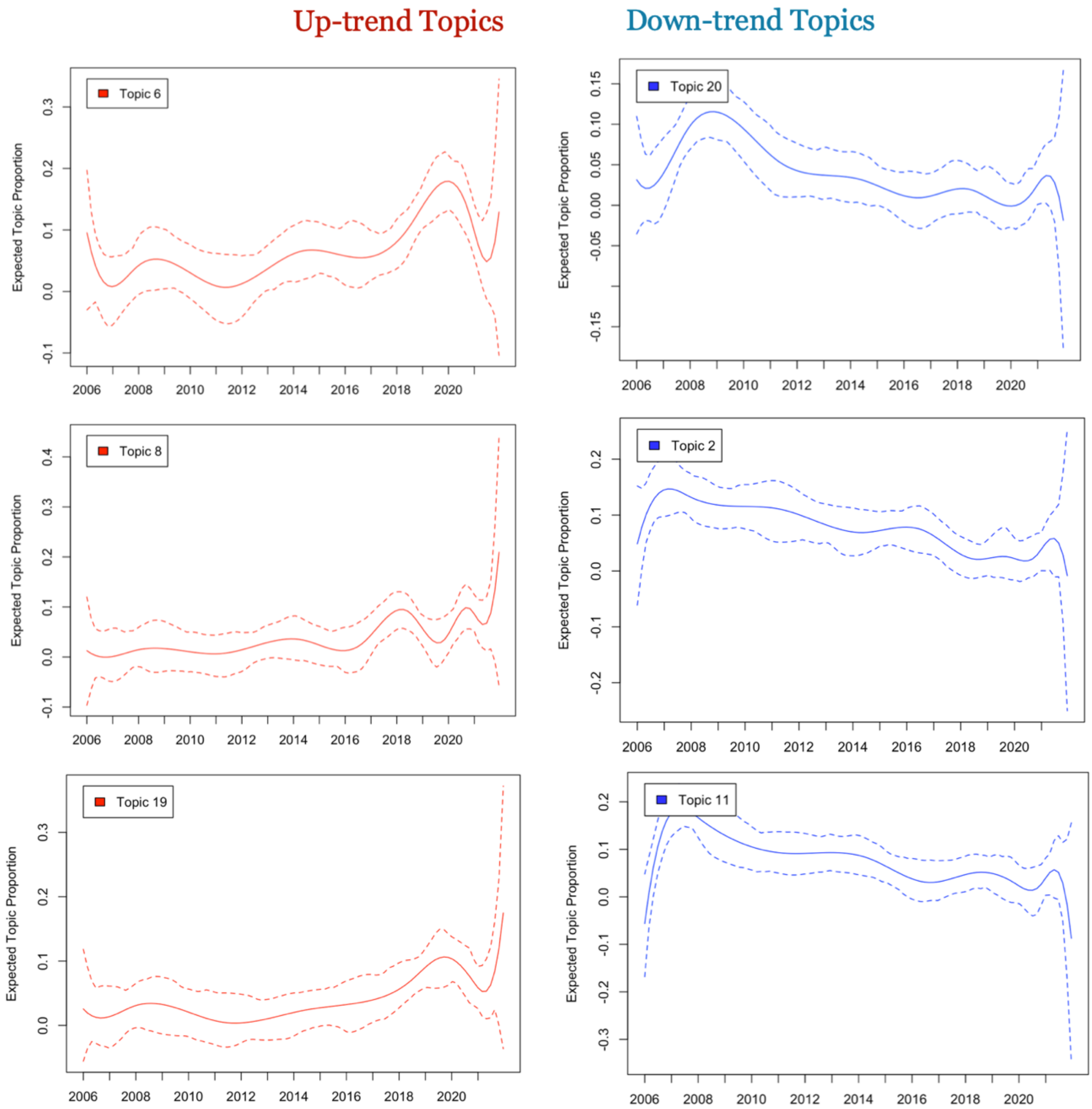


Figure 9: uptrend topic and down trend topic. The method took reference from Robert et al (2019), which use spine functions to render the line.

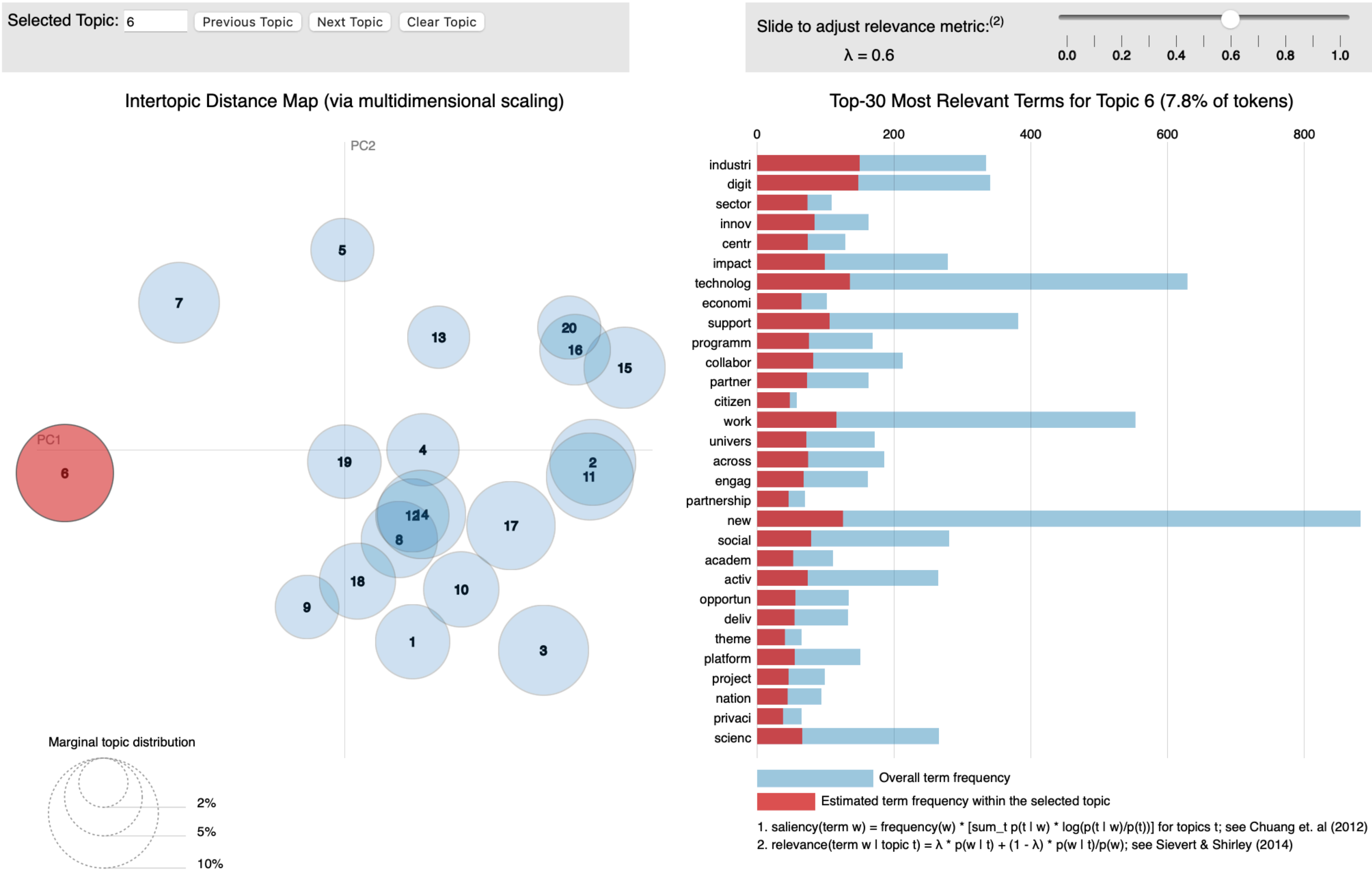
Figure 10. 1: Topic Modelling with k = 20. * is used to indicate it is not clear to interpreted topic.

label	topic	terms	examples
AI in decision Making	1	system, decis, algorithm, trust, use, legal, human, data, need, respons	Transparent Rational Decisions by Argumentation (TRaDAr), UKRI Trustworthy Autonomous Systems Node in Governance and Regulation, EthicalML: Injecting Ethical and Legal Constraints into Machine Learning Models
Data Science*	2	problem, model, data, applic, learn, method, algorithm, use, domain, propos	Northwest European Seasonal Weather Prediction from Complex Systems Modelling, Generative Kernels and Score Spaces for Classification of Speech, Warm starting Techniques for Stochastic Programming Problems solved by Interior Point Methods
Robotic and Automation	3	robot, system, autonom, environ, object, human, learn, task, oper, control	Sustained Autonomy through Plan-based Control and World Modelling with Uncertainty, REcoVER: Learning algorithms for REsilient and VErsatile Robots, Robot In-hand Dexterous manipulation by extracting data from human manipulation of objects to improve robotic autonomy and dexterity - InDex
Others*	4	inform, user, use, technolog, cancer, can, person, process, requir, record	Framework for Computational Persuasion, Novel optimization framework for real-time automated radiation therapy, EPSRC-SFI: SpheryStream
Sound Recognition	5	sound, imag, collect, new, use, audio, archiv, learn, sourc, machin	Making Sense of Sounds, From Lima to Canton and Beyond: An AI-aided heritage materials research platform for studying globalisation through art, Computational wszapproaches to cognition: the origins of social and causal reasoning in children and primates
Industrial Collaboration	6	industri, digit, technolog, new, work, support, develop, impact, innov, collabor	Centre for Digital Citizens - Next Stage Digital Economy Centre, Creative Media Labs: Innovations in Screen Storytelling in the Age of Interactivity and Immersion, The Digital Creativity Hub
AI for Recreation	7	comput, music, creativ, workshop, new, research, game, intellig, network, support	Engaging three user communities with applications and outcomes of computational music creativity, AEOLIAN (Artificial intelligence for cultural organisations), Computer-Human Interactive Performance Symposium (CHIPS)
Urban Twin[5]	8	data, (Urban, Twin), model, use, learn, can, transport, digit, machin, new, physic,	deeP redUced oRder predIctive Fluid dYnamics model (PURIFY), Probing for New Physics at the LHC: Unraveling the Higgs Mechanism through Polarisation and Hadronic Decays, Learning an urban grammar from satellite data through AI
Business Analytics	9	servic, busi, see, market, cost, technolog, map, provid, product, requir	Data Science for the Detection of Emerging Music Styles, Bilateral ESRC/FNR: Experimental Assessment of the Societal Impact of Algorithmic Traders in Asset Markets, Unlocking the Potential of AI for English Law
AI as Management Agent	10	agent, human, can, plan, system, intellig, use, argument, decis, comput	Designing Mechanisms for Automated Resource Allocation: A Case for Support, AI Social Agents, Intelligent Agents for Home Energy Management
Evolutionary Computation	11	model, problem, comput, process, use, biolog, can, system, new, design	Network Coding via Evolutionary Algorithms, Rigorous Runtime Analysis of Nature Inspired Meta-heuristics, Watched Literals and Learning for Constraint Programming
UI System-User Interactions	12	system, interact, user, can, use, natur, dialogu, develop, languag, interfac	Unmute: Opening Spoken Language Interaction to the Currently Unheard, MaDrIgAL: MultiDimensional Interaction management and Adaptive Learning, An Infrastructure for Adaptive System Development
Knowledge Generation	13	data, use, method, work, approach, new, report, provid, model, techniqu	Analysing Narrative Aspects of UK Preliminary Earnings Announcements and Annual Reports: Tools and Insights for Researchers and Regulators, Fast Generalised Rule Induction, Matheuristics for multi-criterion data clustering: towards multi-criterion big data analytics
Healthcare	14	patient, diseas, data, use, can, clinic, health, healthcar, care, risk	Integrated Technology Platform to Support Optimal Management of Ageing with Diabetes, SPHERE - A Sensor Platform for Healthcare in a Residential Environment (IRC Next Steps), Feasibility Study: Integrating Games-Based Learning and Computational Modelling to Control MRSA.
Natural Language Processing	15	model, languag, learn, use, translat, data, machin, human, can, system	Non-Parametric Models of Phrase-based Machine Translation, Bayesian Synchronous Grammar Induction, Modelling Discourse in Statistical Machine Translation
Programming	16	use, comput, softwar, techniqu, develop, secur, can, automat, human, theori	Fast Runtime Verification via Machine Learning, Automatic repair of natural source code, Statistical Natural Language Processing Methods for Computer Program Source Code

[5] The package `toLDAviz` provide alternative way to evaluate term relevance. Term "urban" and "twin" were amongst exclusive to topic 8, hence topic 8's label.

Neural Networking	17	learn, brain, network, comput, neural, use, new, design, process, can	Brain-inspired non-stationary learning., MOA: High Efficiency Deep Learning for Embedded and Mobile Platforms (Full EPSRC Fellowship Submission), Automating electron microscopy: machine learning for cluster identification
Data Analytic	18	data, develop, monitor, use, detect, system, sensor, analysi, inform, cloud	Integrated software solution for the 3-dimensional capture and analysis of footwear evidence, CRITiCaL - Combatting cRiminals In The CLoud, SENSUM: Smart SENSing of landscapes Undergoing hazardous hydrogeological Movement
Chemistry	19	new, use, drug, materi, develop, equip, process, cell, manufactur, make	A Robot Chemist, High-throughput Differential Expression Proteomics, Capital award for Core Equipment
Ontology	20	knowledg, ontolog, reason, use, concept, can, system, model, exampl, web	ConDOR: Consequence-Driven Ontology Reasoning, HermiT: Reasoning with Large Ontologies, LogMap: Logic-based Methods for Ontology Mapping

Figure 10.2: LDA visualisation of topics



Further Analysis

Project Transfer due to Academia Mobility

We took advantage of transferred projects and conducted basic network analysis in attempt to reveal pattern of knowledge transfer amongst research organisations. However, academia moves in different directions.

For example, in project "*detecting and preventing mass marketing fraud*", Principal Investigator (PI) Whitty were in University of Leicester, COI Investigator Sorell were at University of Warwick. However, for unknown reason, PI, Whitty withdrawn from the project, so COI Sorel were assigned as lead investigator. As a result, project "transferred" from Leicester to "Warwick". The project was transferred from Leicester to Warwick not because the principles investigator has transferred but was out. The project has not been transferred to Warwick. Warwick has always been part of the project. Academia transfers does not always explain project transfer. Each case may be different.

Project transfer due to academia attrition might share same principal investigators. Based on this assumption, we can filter project transfers mainly that solely academia attrition.

Originally 37 pair of transferred projects now reduced to 35. Only two projects are removed, including the projects demonstrated as an example. The other project "Computational Creativity Theory" having one project missing (reference: EP/J004049/3). This changes project flow of Warwick from 3 to 2.

The centre of the network appears to be University Collage, London, which has 8 connections with other universities. Six projects are transferred into University College London and only two projects are transferred out of University College London, and both are transferred to University of Warwick.

Major destiny is University College London (4 net flows); Queen Mary, University of London (3 net flows), University of Warwick (3 net flows).

Major sources may be Lancaster University, which has four projects flows out. Note Lancaster is amongst highly funded universities.

Most of the projects being transferred tends to end up in London, with Cambridge and Warwick being two distinct exceptions.

Topic 6: Industrial Collaboration

Topic 6 is the only projects that is the only project that related with assemble innovations hubs that collaborate with diverse experts and stakeholders. Only topic 6 is significantly associated with amount of funds allocated to research organisations.

One output of topic is theta, which is a number reflecting topic proportion for each document. Theta can indicate how pure is a topic. Selecting documents with topic 6 at theta above 0.75, one is able to filter out projects most related to topic 6. Alan Turning institute has been removed from as statical outliner. As a results 19 projects are yield. Among those projects, proportion of topic 6 can reach as highest as reach 0.98, as lowest as 0.83. As can be seem the even though has topic proportion has been caped at 0.75, none of the of value is below 0.8.

Topic 6: Sensitivity to Topic Purity

Above analysis is established on "topic purity" of above 0.75. Lower this standard may reach different conclusions. Further, using different measurement (for example use either average award or average price) also might yield different results. One way to exam the sensitivity is to sample research universities and includes Totally three pair of comparation have been made, listed in a table of six charts.

Figure 11.1: visualise network on a map

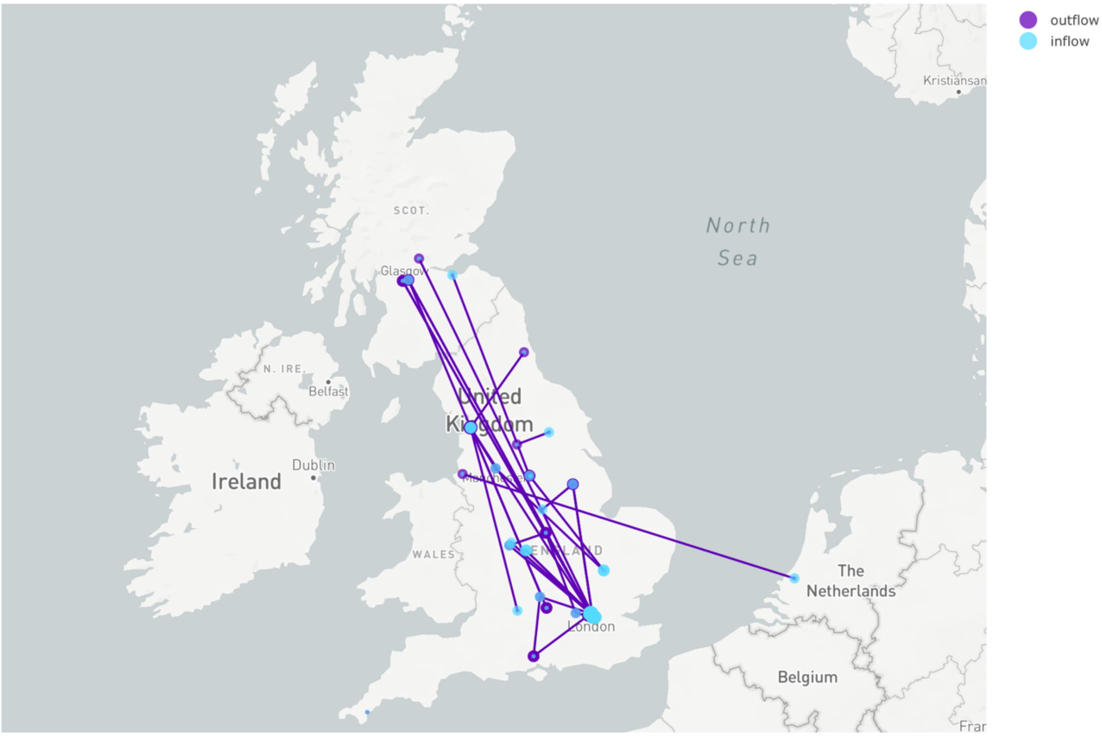
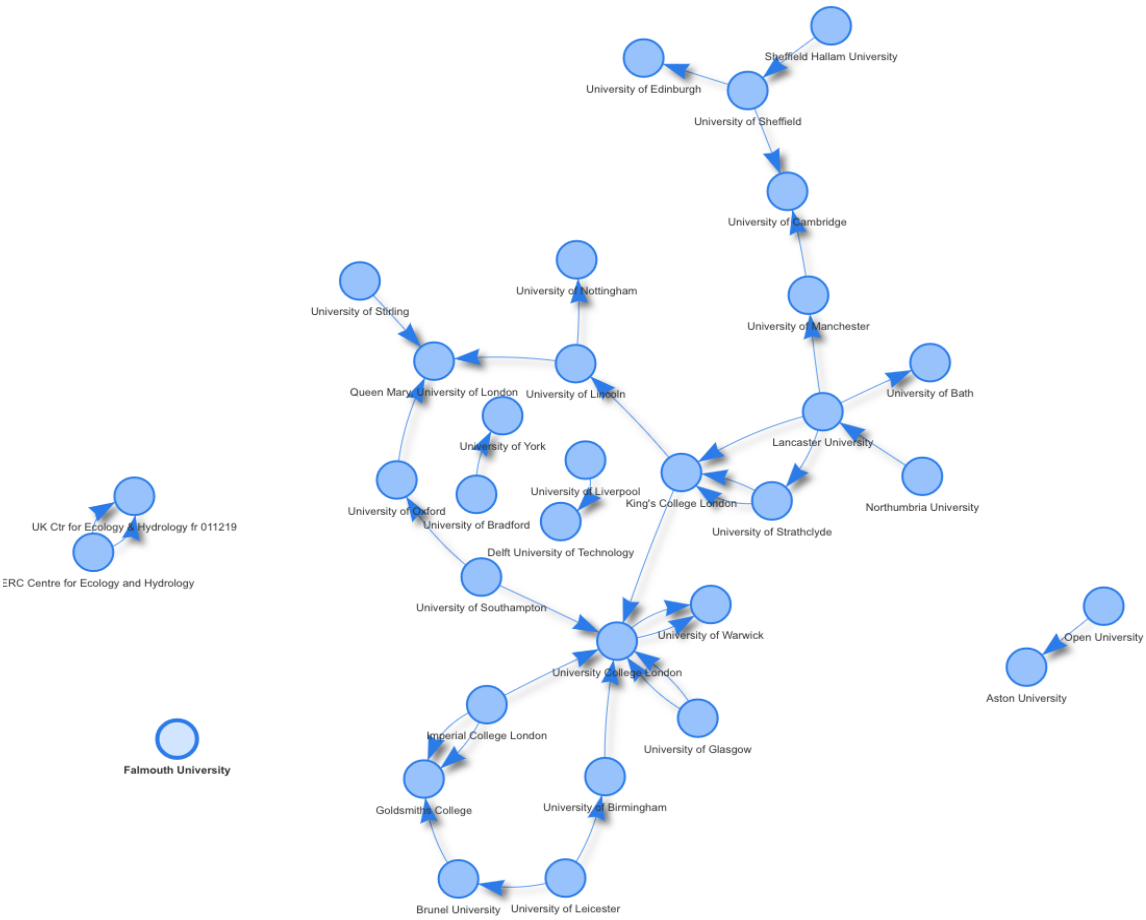
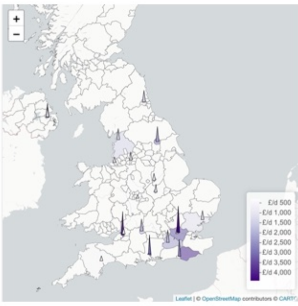
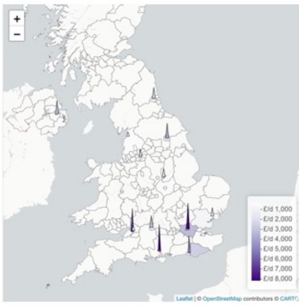
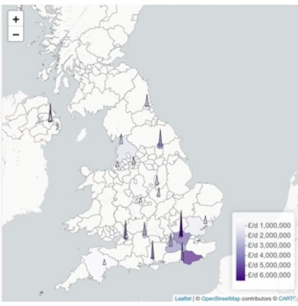
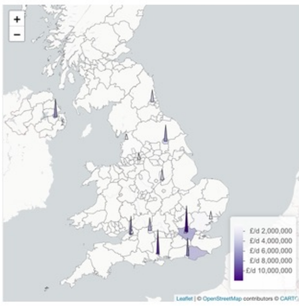
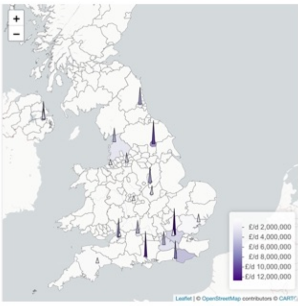
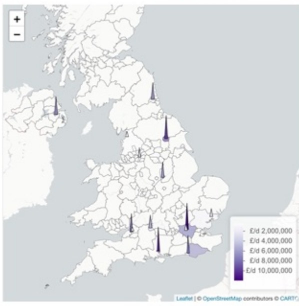


Figure 11.2: underlaying connections



Sensitivity To Purity

Indicator	Purity > 0.5	Top 5 > 0.5	Purity > 0.75	Top 5 at > 0.75
Average Price		1. UCL, 2. Bristol, 3. Southampton, 4. Sussex, 5. York		1. Southampton, 2. UCL, 3. Bristol, 4. Sussex, 5. York
Average Award		1. UCL, 2. Bristol, 3. Sussex, 4. Southampton, 5. Queen's University of Belfast		1. UCL, 2. Southampton, 3. Bristol, 4. Sussex, 5. Queen's University of Belfast
Total Award		1. UCL, 2. Southampton, 3. York, 4. Bristol, 5. Sussex		1. UCL, 2. Southampton, 3. York, 4. Bristol, 5. Sussex

Highlights represent key changes

Figure 12: how result is sensitive to measurements and topic purity.

Highlighted part measures key differences using difference via different measurement, topic purity cap. As expected, using more strict measurements number of topics would be reduces. Resulting in lower number of research organisation shows up on map.

For Southampton, UCL, Bristol, Sussex are relatively stable, using either average ranking or average award. When looking at average award, University of Belfast made it into top 5. This is in line with previous findings that University of Belfast tends to have longer research project (section).

A key change is University York will squeeze into top 5 rankings by looking at total amount of award. This may be attributed to the fact that University of York has two related projects while the others (universities that has been looked at in this example) have only one.

Based on analysis above, it is reasonable to conclude that UCL, Southampton, Bristol, Sussex, are heavily funded places other than Alan Turing Institute. Because by modify purity changes number of projects included, average method is chosen for executive summary.

Discussions and Limitation

Duplications dues to project transferred and sub-projects

Several projects in GtR shares same title same contents. Some of those duplication projects are essentially the same project. This is because when an academia transfer from one organisation to another, they can take research grant with them. When this happens, the organisation associated with the transferred academia changes. This is explained by GtR data dictionary (GtR b, 2017). Each time a project is transferred, the unsend budget is also transferred. In most cases transfers are resulted from principal investigator's career change.

For a project that has been transferred, duration of the project is sorter. Hence results in a "price" figure that is higher than it is. Those transferred projects over-estimate "price" is used to indicate true value of research projects. Hence, transferred projects are taken out for ANOVA.

Number of people working on a project unaccounted

This report uses "price" as an indicator of true value of a potential research project. Linear models validated such use. Project duration could account as high as 68% of variance. Further the higher the duration the higher the "price". Arguably, it is easy to see this as a case of longer the research duration the higher the investment risks. However, it could also because longer project may have more people working on it.

Specific rankings are not important

It has been demonstrated in further findings, although the big trend does not always change, adopting different techniques results in specifics of university ranking to be different. The reason that when adopting relatively stricter criteria for topic purity, results does become relatively stable is because topic proportion for topic 6 clustered around 0.8 – 0.99.

ANOVA and effect estimation has been used to test if true price or project duration difference with each organisation. However, ANOVA only compares mean with each organisation. To compare between each individual organisations, the best approach is *Tukey Procedure* as introduced by Chambers (Statistic Books). This would have been a long process as there are 82 organisations to compare with. That would have been, 3321 (81*(81 + 1)/2) comparations to make. Because the specifics of university rankings is not necessary in drawing conclusion for the purpose of this research, Tukey procedure is not performed.

University – Industry Linkages

Despite being the feature of Triple Helix Innovation theory, industry university linkage is not sufficiently examined in this project. GtR only record public funded research project. As a results, any research projects, even if the project has business partners participation, will be labelled as "being led by a public organisation. Although GtR do keep a record of partner organisations, those data are stored in highly structured json format, which make it very difficult to extract and cleaning within time limitation.

Research councils seems to reward research organisation who consult multi-stakeholders. One can still take a glimpse of how industrial collaboration is becoming hot spot of public investment. Topic 6 (industrial collaboration) is the only topic proportion that are significantly associated with amount of award and is the only project that is significantly associated, the topic proportion has been increasing. This implies growing effort of UK to increase industry-university linkages.

Sweden has undertaken similar effort to fund innovation hubs hoping to stimulate innovation. Recently, Sweden has been ranked the second world most innovative country by world intellectual property, (WIPO, 2021). However, less than a decade years ago, Fogelberg & Thorpenberg (2012) who studied the innovation hubs in Sweden found tensions between industries and research institutes. While industry complained research were unable to meet market demand, academia questioned competencies and knowledge of industries in assuming prudency of science innovation.

Academia and businessmen live in very different environment. There are challenges leading two parties to work together. For universities restraint in research capacities, there are opportunities to work towards university industry linkage. Although, whether those innovation hubs would contribute economic value may be a very different research question.

Alan Turing

During the ANOVA analysis, Alan Turing was picked out suspecting that Alan Turing may sub fund projects.

Alan Turing is the most well-funded AI innovation hub in the UK. Alan Turing Institute has respectively received 40 million *strategic priority funds,* and about another 38 million standard funds. Those two is 131 times more than any 50% of any standard research fund. Considering the significance of value, Alan Turing by itself is worthy of investigation.

Alan Turing's 2021 finical report has included a list of institutions receiving grant. Those includes University of Exeter and Southampton, classified under category Grant Payable, which is code name for cost incurred in research. In 2021, research related costs accumulate up to 12 million. However, it is not certain if any of those research activities were accounted under Alan Turing or if under any of partner universitas in GtR database. Further, Alan Turning Institute has subsidiary Turing Innovations Limited, a company, which supplement additional information. Those financial statement contains valuable information of innovation- researchers. Ultimately, tapping into these areas will challenge researcher's financial literacy and knowledge for funding policies, which is beyond the expertise and scope of this dissertation.

Regarding this dissertation, there is opportunity to deploy further analysis. GtR api scheme allows query data based on research organisations. Since Alan Turing is a research institute dedicated for AI, this data would have been appreciated towards this dissertation's research objectives.

Academia Mobility

In this report, academia mobility is measured intermediately via project that has been transferred. There totally 36 projects have been transferred (37 -1, because one project has been transferred three times). In further analysis this number were reduced to 35 on the assumption that academia moving is indicated by principal investigator being same person. Arguably, there is no way for sure this is the case for all 35 projects, with each reasons different case by case. However, so far as explanation GtR data dictionary (2020) goes, project transfer is only explained as principal investigator move. So, sensibly, the methodology adopted by this dissertation is likely to be highly representative of tracing academia's career movement.

The results have showed that academia tends to flow into central region, and in specific London. It seems that academia's favoured destination tends to be either London or world leading university, such as University of Cambridge, Warwick, UCL. Ironically, none of those are the best funded places in realm of AI.

Average project pricing in London region is lower compared to surrounding regions. These are likely to coincide the fact that London hosts a high dense of universities. As a result, each university compete for grant resulting in price for research in London to be lower.

Within the topic of AI, there are only 35 cases of project transfer, hence not making this a strong case to analysis academia mobility quantitatively. However, it is a stronger case to apply same method to more generalised topics. Innovation-researchers or analyst can trace what scientists are moving to where, which helps mapping the novel innovation clusters in UK. As there has been a growing interest for UK research councils to invest in building science parks and innovation hubs. Governments and investors can potentially leverage GtR data to detect raising innovation clusters and determine funding priorities.

Conclusions and Implications

This dissertation set off to explore public funded investment for Artificial Intelligence in the UK. It seeks to answer "who", "what" and "where" questions: "Who are the key players in research artificial intelligence?"; "What are UK really invested in?"; "Where is been invested?".

The Alan Turing Institute is undoubtfully play significant role in shaping the future of Artificial Intelligence, although this is not necessarily known through this dissertation. By using normalised ANOVA, Newcastle receives most award. Another key player is Southampton. Southampton is list as top by looking at average price. However, this is misleading because Southampton happens to be hosted a few expensive projects.

Regarding what are UK really invested in, the answers are as much as the triple helix theory can forecast. Research councils incentivise universities to collaborate with industries. This dissertation has showed this incentive is significantly high (topic 6, *figure 10.2*) across all research topics what is more, the topic proportion has significantly increased with time. Further, there is increase in “Urban Twin” and “Chemistry”, which coincide the Alan Turing project, AI for Government and Science.

Cai (2015) comments that some common public funding strategy in the West is to incentivising industrial - government linkage or invest in hot notch projects. Evidence was found for the former but not the later. No evidence suggest UK is investing in a particular AI sub area as across 20 topics, only one topic was funded to be signify in award value.

Regarding the locus of innovation cluster, London may still be considered as UK’s biggest AI hub. This conclusion agrees with Data City (2013). Although, average funding for each organisation diminishes in competition for limited grant. As a result, surrounding universities overshadow London.

Reference List

Abadi, H. H. N., He, Z., & Pecht, M. (2020). Artificial Intelligence-Related Research Funding by the US National Science Foundation and the National Natural Science Foundation of China. IEEE Access, 8, 183448-183459.

Alan Tuning Official Website

Benner, C. (2003). Learning communities in a learning region: the soft infrastructure of cross-firm learning networks in Silicon Valley. Environment and planning A, 35(10), 1809-1830.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022.

Cai, Y. (2015). What contextual factors shape 'innovation in innovation'? Integration of insights from the Triple Helix and the institutional logics perspective. Social Science Information, 54(3), 299-326.

Cai, Y., & Etzkowitz, H. (2020). Theorizing the Triple Helix model: Past, present, and future. Triple Helix, 7(2-3), 189-226.

Cambon J, Hernangómez D, Belanger C, Possenriede D (2021). "tidygeocoder: Geocoding Made Easy." doi: 10.5281/zenodo.4686074 <https://doi.org/10.5281/zenodo.4686074>, R package version 1.0.3, https://CRAN.R-project.org/package=tidygeocoder <https://cran.r-project.org/package=tidygeocoder>.

Chambers, J. M., Freeny, A and Heiberger, R. M. (1992) Analysis of variance; designed experiments. Chapter 5 of Statistical Models in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In Advances in neural information processing systems (pp. 288-296).

Dunham, J., Melot, J., & Murdick, D. (2020). Identifying the Development and Application of Artificial Intelligence in Scientific Text. arXiv preprint arXiv:2002.07143.

Engel, S. J. (2015) Global Clusters of Innovation: lessons from Silicon Valley, California Management Review, winter 2015, Vol. 57 Issue 2, pp. 36-65

Fatima Hameed Khan, Muhammad Adeel Pasha, & Shahid Masud. (2021). Advancements in Microprocessor Architecture for Ubiquitous AI—An Overview on History, Evolution, and Upcoming Challenges in AI Implementation. Micromachines, 12(665), 665. https://doi-org.uoelibrary.idm.oclc.org/10.3390/mi12060665

Fogelberg, H., & Thorpenberg, S. (2012). Regional innovation policy and public–private partnership: The case of Triple Helix Arenas in Western Sweden. Science and Public Policy, 39(3), 347-356.

Forth, T., Connell, P., Laflin, P., The Data City (2018) UK Tech Innovation Index 2 and The Data City Report. https://thedatacity.com/products/uk-tech-innovation-index-report/

Geuna, A. (1997). Allocation of Funds and Research Output: The Case of UK Universities. Revue d'Economie Industrielle, 79, 143–162. https://doiorg.uoelibrary.idm.oclc.org/http://www.cairn.info/revue-d-economie-industrielle.htm

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. California management review, 61(4), 5-14.

McKinsey and Company (2018), Growth and Productivity Estimates. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy

Monteiro, S. P. D. O., & Carayannis, E. G. (Eds.). (2017). The Quadruple Innovation Helix Nexus: A Smart Growth Model, Quantitative Empirical Validation and Operationalization for OECD Countries. Springer.

Philip Cooke, & Philip Cooke. (2004). Life Sciences Clusters and Regional Science Policy. Urban Studies, 41(5/6), 1113–1131.

Rahkovsky, I., Toney, A., Boyack, K. W., Klavans, R., & Murdick, D. A. (2021). AI Research Funding Portfolios and Extreme Growth. Frontiers in Research Metrics and Analytics, 6, 11.

Robert M., Stewart B., Tingley D. (2016). Navigating the local models of big data: the case to topic model. In R. M. Alvarez Computational Social Science: Discovery and Prediction (pp. 51-97). doi: 10.1017/CBO9781316257340 <https://doi-org.uoelibrary.idm.oclc.org/10.1017/CBO9781316257340>

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. Journal of Statistical Software, 91(1), 1-40.

Steinert, M., & Leifer, L. (2010, July). Scrutinizing Gartner's hype cycle approach. In Picmet 2010 technology management for global economic growth (pp. 1-13). IEEE.

Shin, J., & Grant, A. M. (2020). When putting work off pays off: the curvilinear relationship between procrastination and creativity. Academy of Management Journal, (ja).

Sievert, C., & Shirley, K. (2014). LDAvis: A Method for Visualizing and Interpreting Topics. Paper presented at the Association for Computational Linguistics Workshop on Interactive Language Learning, Visualization, and Interfaces (ACL2014), Baltimore, MA. USA.

Smith, H. L., & Waters, R. (2015). Regional Synergies in Triple Helix Regions: The Case of Local Economic Development Policies in Oxfordshire, UK. Industry and Higher Education, 29(1), 25–35. https://doi.org/10.5367/ihe.2015.0238

Stuart, D., & Thelwall, M. (2006). Investigating triple helix relationships using URL citations: a case study of the UK West Midlands automobile industry. RESEARCH EVALUATION, 2, 97.

The Data City (2020) UK Tech Innovation Index — 2.0 Beta: Artificial Intelligence and Data: http://thedatacity.com/products/uk-tech-innovation-index-2/?options=true&datagroup=Artificial%20intelligence%20and%20data&location=null

UK Research and Innovation (2021) Transforming out world with AI. https://www.ukri.org/wp-content/uploads/2021/02/UKRI-120221-TransformingOurWorldWithAI.pdf

UK Research and Innovation (2021) Gate Way to Research (GtR) a. https://gtr.ukri.org/resources/data.html

UK Research and Innovation (2021) Gate Way to Research (GtR) b. https://gtr.ukri.org

World Intellectual property Organisation (WIPO) (2021) Global Innovation Index 2020. https://www.wipo.int/global_innovation_index/en/2020/