How to use the Economic Complexity Index to guide innovation plans

Mauro Pelucchi
Head of Global Data Science

Lightcast

Descriptions

In this talk we will present the Economic Complexity Index (ECI) and the Product Complexity Index (PCI), two network measures that provide unique insights into economic development patterns.

We will show how to compute these metrics and explore the network theory behind these indices (Hidalgo and Hausmann, 2009).

The measures are also related to various dimensionality reduction methods and can be used to determine distances between nodes based on their nodes based on their similarity.

Finally, we will discover how to interpret these metrics to compare countries, markets, products, and guide our plans in a data-driven context.

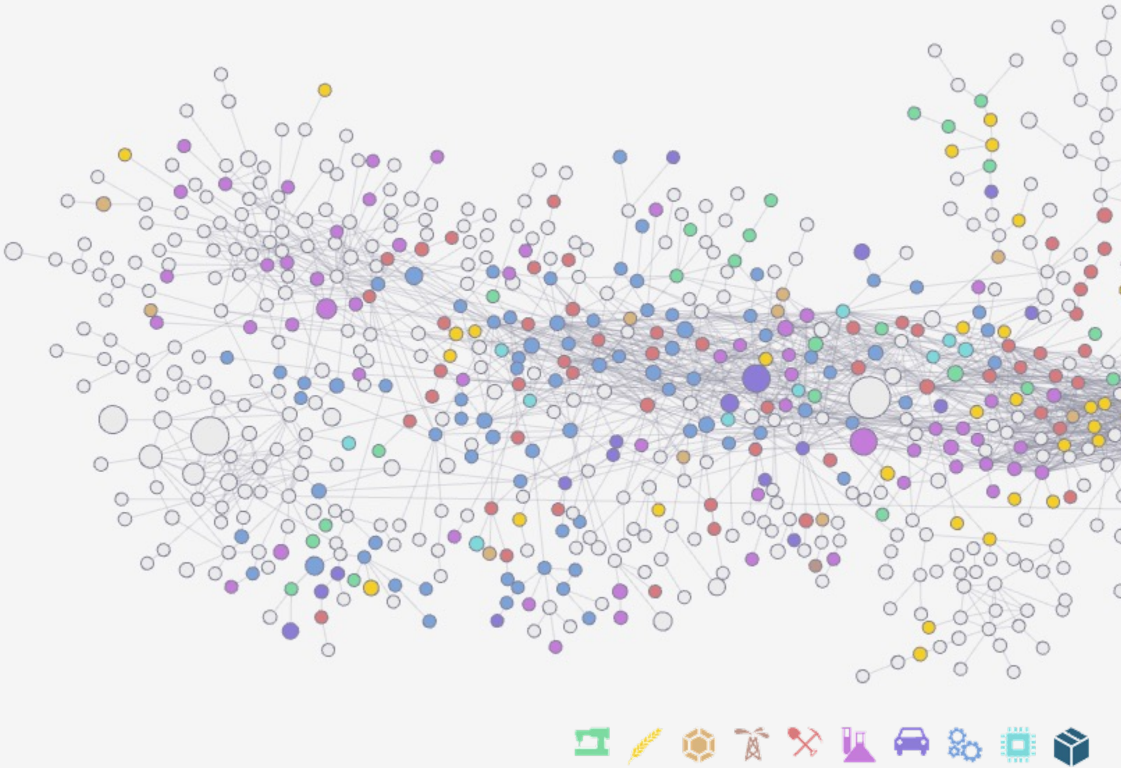
I am a senior data scientist and big data engineer responsible for the design of the "**Real-Time Labour Market Information System on Skill Requirements**" for

I currently works for

In Lightcast, I work as **Head of Global Data Science** with the goal to explore, design and delivery innovative solutions about Labour Market Data.
In collaboration with the **University of Milano-Bicocca**, I took part in many research projects related to the labour market intelligence systems.
I collaborate with the **University of Milano-Bicocca** as a lecturer at the Master Business Intelligence and Big Data Analytics and with the **University of Bergamo** as a lecturer in Computer Engineering.

Complexity

https://atlas.cid.harvard.edu/
Hidalgo (2020)
(Hidalgo and Hausmann, 2009).



Check for updates

Economic complexity theory and applications

César A. Hidalgo [1,2,3]

Abstract | Economic complexity methods have become popular tools in economic geography, international development and innovation studies. Here, I review economic complexity theory and applications, with a particular focus on two streams of literature: the literature on relatedness, which focuses on the evolution of specialization patterns, and the literature on metrics of economic complexity, which uses dimensionality reduction techniques to create metrics of economic sophistication that are predictive of variations in income, economic growth, emissions and income inequality.

Since Adam Smith's pin factory, wealth has been related to the division of knowledge and labour. Yet, even though scholars have long recognized economies as complex systems[1–6], the empirical study of economic complexity only accelerated in the past decade, with the emergence of new data and methods.

Like traditional approaches to economics, economic complexity focuses on the duality between economic inputs and outputs. But, unlike traditional approaches, which either aggregate output — as gross domestic product (GDP) does — or assume the nature of inputs — such as capital, labour and knowledge — economic complexity methods embrace fine-grained data on thousands of economic activities to learn both abstract factors of production and the way they combine into thousands of outputs. This is made possible by applying dimensionality reduction techniques to data on the geography of activities, such as product exports, employment by industry or patents by technology. These techniques — which are related to matrix factorization and are common in machine learning — provide a powerful way to summarize the geography of economic output and can be used to construct predictors of a location's diversification and development potential.

The study of economic complexity accelerated during the last decade thanks to two contributions. The first involved the introduction of metrics of relatedness[7,8], which measure the overall affinity between a specific activity and a location. Relatedness metrics explain path dependencies and predict which activities will grow or decline in a location. They help answer questions such as how 'far' Quito, Kiev or Guadalajara are from having a thriving pharmaceutical industry. The second contribution was the development of metrics of complexity[9]. These use data on the geography of activities (such as exports by country or region, or employment by city and industry) to estimate the availability, diversity and sophistication of the factors or inputs present in an economy. Metrics of complexity extract key information about an economy's capacity to generate[9] and distribute income[10].

Unlike previous approaches to economic growth and development, which attempt to identify individual factors, relatedness and complexity methods are agnostic about the nature of factors. Instead, they try to estimate their combined presence, without making strong assumptions about what these factors may be. For instance, relatedness metrics can be used to estimate the combined presence of inputs that are specific to an activity, no matter if these inputs involve specific forms of labour, capital or institutions. Relatedness can, then, be used to anticipate changes in specialization patterns[7,8,11–18], such as the probability that a location enters or exits an activity.

Complexity metrics apply dimensionality reduction techniques (related to singular value decomposition (SVD); BOX 1) to identify the combinations of factors that best explain the geography of multiple economic activities. Unlike in traditional growth models, which assume the nature of factors, dimensionality reduction techniques can be used to learn factors directly from the data. Economic complexity metrics are useful for predicting economic growth[9,19–25], income inequality[10,26–29] and greenhouse gas emissions[30–33].

Beyond data and methods, the study of economic complexity was motivated by other key trends: the revival of industrial policy, the growth of artificial intelligence (AI) and the development of endogenous growth theory.

Complexity methods grew together with a revival of industrial policy[34,35] and the realization that economic development requires upgrading. Complexity methods help characterize detailed economic structures and provide a quantitative base for industrial policy efforts. Today, these efforts are embodied in Europe's Smart

[1]Center for Collective Learning, Artificial and Natural Intelligence Toulouse Institute (ANITI), University of Toulouse, Toulouse, France.
[2]Alliance Manchester Business School, University of Manchester, Manchester, UK.
[3]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA.
e-mail: cesifoti@gmail.com
https://doi.org/10.1038/s42254-020-00275-1

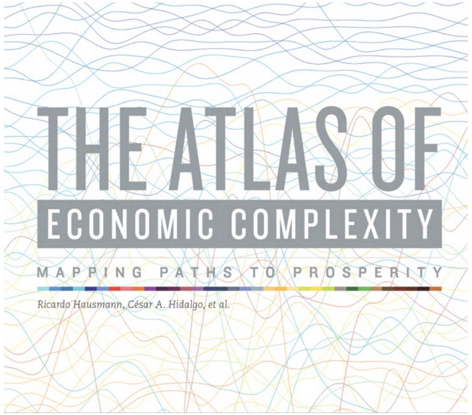
Moving beyond skills frequency using economic complexity theory

- Concept firstly developed by Hidalgo and Hausmann in 2009 to identify countries comparative advantage.

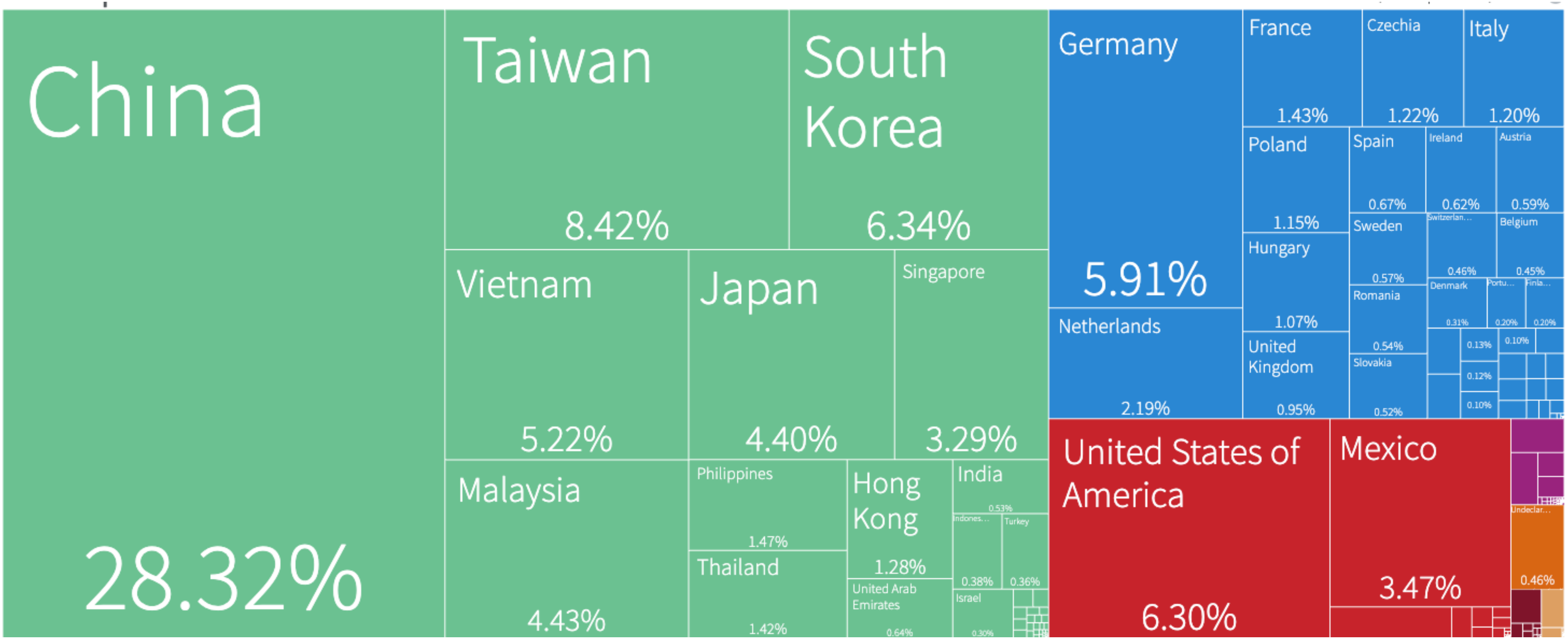
- Economic complexity is used to capture the process of knowledge creation in an economy by looking at the combination of two metrics:

  1. **Diversity**: how many different sectors does a place specialises in?
  2. **Ubiquity:** how many places specialise in that particular sector?

- On average, economies with a higher ECI tend to have a competitive advantage in several highly complex sectors.



How difficult is to product an activity?



Who exported Electronics in 2020?

Economy complexity theory and applications

- A measure of the knowledge in a society as expressed in the activities it makes. The economic complexity of an area is calculated based on the **diversity** of activities a area does and their **ubiquity**, or the number of the areas able to do them (and those area’s complexity).

- Areas that are able to sustain a diverse range of activity know-how, including sophisticated, unique know-how, are found to be able to do a wide diversity of activities, including complex one that few other areas can make.

Economy complexity theory and applications
Labour Market

- A measure of the knowledge in a occupation as expressed in the skills it requires. The complexity of an occupation is calculated based on the **diversity** of skills a occupation requires and their **ubiquity**, or the number of the occupations require them (and their complexity).

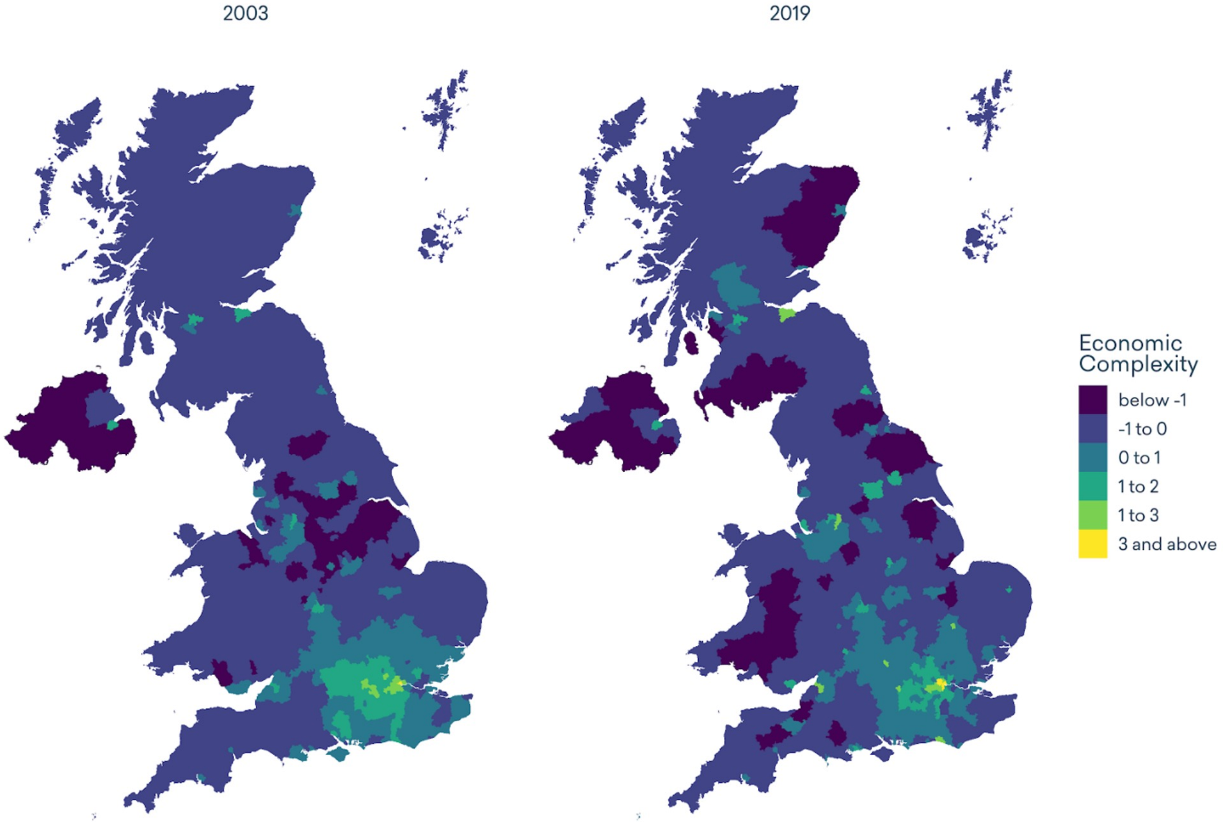
- Occupations that request to know a diverse range of skill know-how, including sophisticated, unique know-how, are found to be able to do a wide diversity of skills, including complex one that few other occupations can make.

Specialisation, ubiquity and diversity

- **Specialisation**: Significant presences or absences of the skills for the occupation
- **Diversity**: A measure of how many different types of skills an occupation requires. A skill requires a specific set of know-how; therefore, an occupation's total diversity is another way of expressing the amount of collective know-how held within that occupation
- **Ubiquity** measures the number of occupations that require a skill
  - More complex knowledge diffuses with more difficulty
- **Proximity**: Measures the probability that an occupation requires skill A given that it requires skill B, or vice versa. Given that an occupation requires one skill, proximity captures the ease of obtaining the know-how needed to move into another occupation.
  - Proximity formalises the intuitive idea that the intensity/ability required for a skill can be revealed by looking to the other skills required.

The changing map of economic complexity

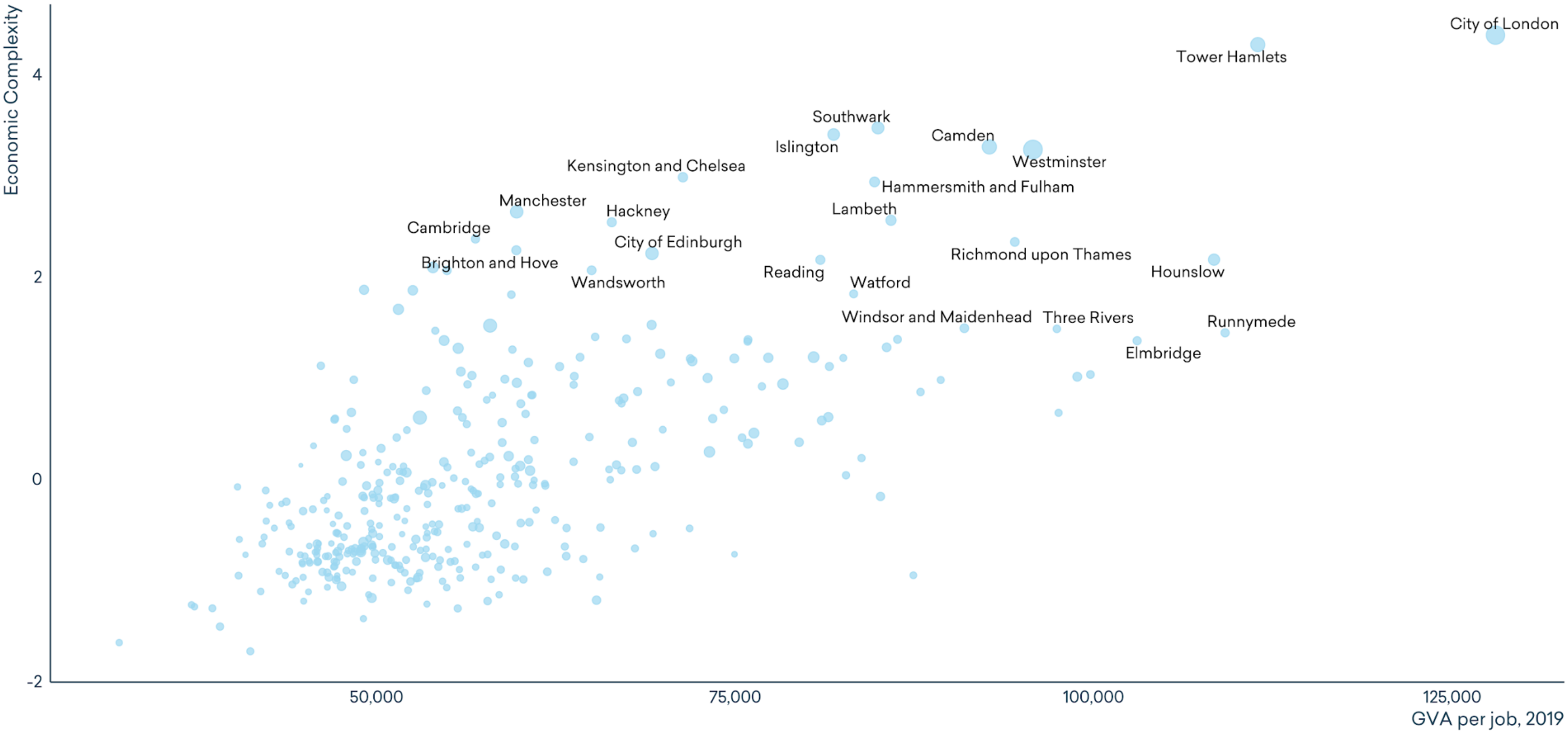
Local authorities by economic complexity, 2003 and 2019



Data: Economic Complexity Index calculated from SIC group jobs from Core LMI

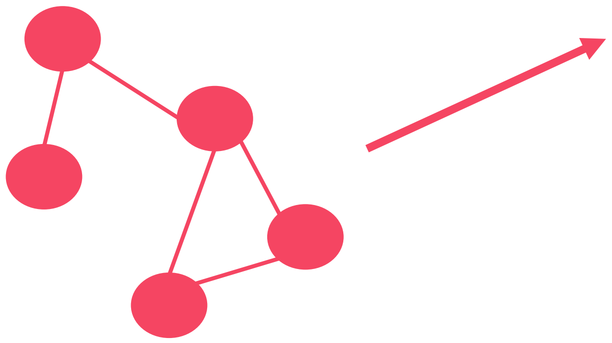
Economic Complexity from 2003 to 2019

Using the ECI to understand the knowledge networks driving economic difference



Data: Economic Complexity Index calculated from SIC group jobs from Core LMI

Flows of knowledge, natural resorces, goods, financial resorces, technologies, people, ….



Time

Define Complexity

$$K_c = f(M_{cp}, K_p),$$

$$K_p = g(M_{cp}, K_c).$$

(Hidalgo and Hausmann, 2009).

Define complexity

$$K_c = \widetilde{M}_{cc'} K_{c'}$$

$$K_p = \widetilde{M}_{pp'} K_{p'}$$

with:

$$\widetilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{M_c M_p}$$

and:

$$\widetilde{M}_{pp'} = \sum_c \frac{M_{cp} M_{cp'}}{M_c M_p}$$

Normalize

- Extensive-Extensive

$$K_c = \frac{1}{M_c} \sum_p M_{cp} K_p$$

$$K_p = \frac{1}{M_p} \sum_c M_{cp} K_c$$

- Insensive-Intesive

RCA

$$rca(o_i, s_l) = \frac{sf(o_i, s_l) / \sum_{j=1}^{p} sf(o_i, s_j)}{\sum_{k=1}^{m} sf(o_k, s_l) / \sum_{k=1}^{m} \sum_{j=1}^{p} sf(o_k, s_j)}$$

Specialisation, diversity and ubiquity

(Hidalgo and Hausmann, 2009).

$$R_{os}^{emp} = \frac{X_{os}E}{X_s E_o}$$

$$R_{co}^{emp} = \frac{X_{co}E}{X_o E_c}$$

Specialisation

$$M_{os} = \begin{cases} 1 & \text{if } R_{os}^{emp} \geqslant 0.25 \\ 0 & \text{if } R_{os}^{emp} < 0.25 \end{cases}$$

$$M_{co} = \begin{cases} 1 & \text{if } R_{co}^{emp} \geqslant 0.25 \\ 0 & \text{if } R_{co}^{emp} < 0.25 \end{cases}$$

Binary specialisation

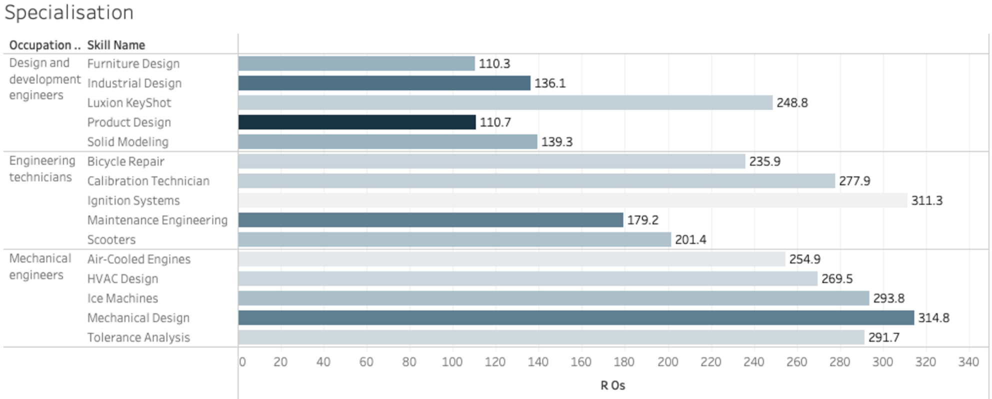
$$M_c = \sum_o M_{os} = diversity$$

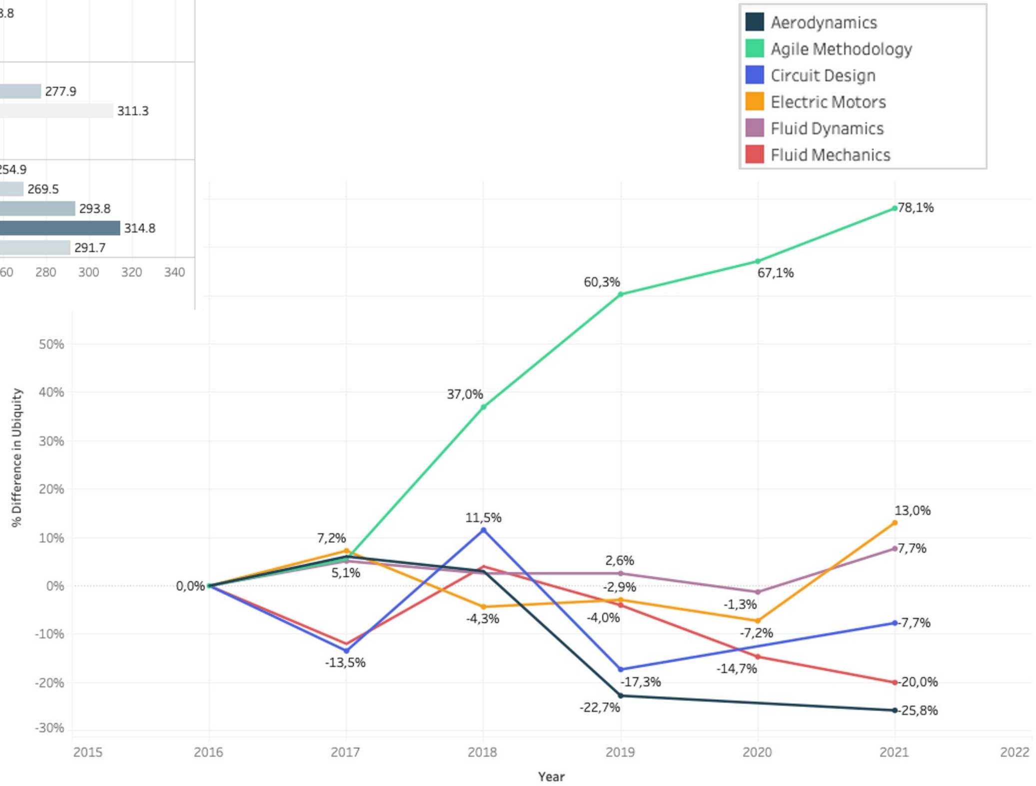
$$M_o = \sum_s M_{os} = diversity$$

$$M_o = \sum_c M_{co} = ubiquity$$

$$M_s = \sum_o M_{os} = ubiquity$$

Diversity, ubiquity





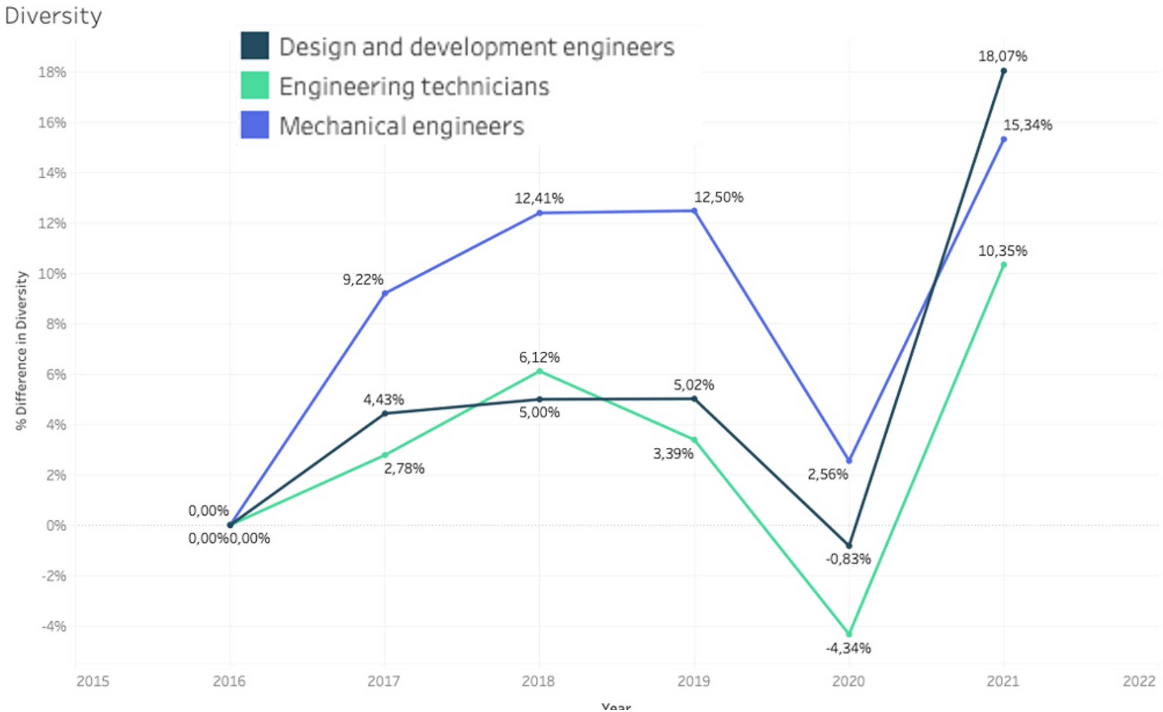
Ubiquity brings out which skills are more complex. A decrease in the ubiquity of a skill shows a decrease in the diffusion of that skill over time, and thus greater complexity, less diffusion of its knowledge.

From the graph we see that skills related to project management, Agile, are now widely diffused and therefore less complex. CFD, Fluid Dynamics, ... on the other hand, are rapidly increasing in complexity.

UK - Economy Complexity
Probability to increase (or decrease) employability of an occupation

The principle of relatedness states that the probability of an area will increase (or decrease) the employability of an occupation with the presence of related activities/occupations in that region.

Engineering and specialisation

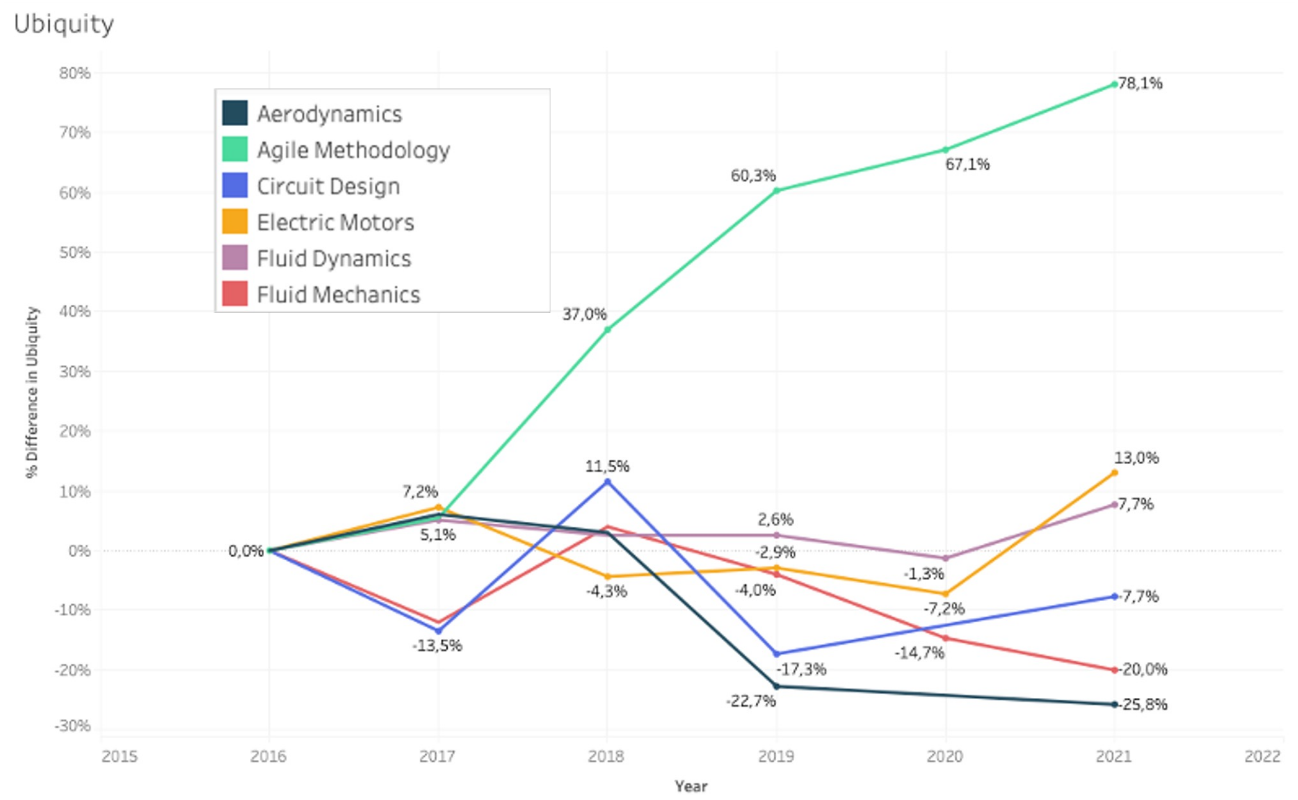


Diversity helps us measure the evolution of role complexity over time.

An increase in diversity is closely related to an increase in the value in the specialisation matrix: in other words, to the number of required specialised skills is increasing.

The graph shows that for engineering profiles we have an average diversity increase of 15% compared to 2016.

More complex knowledge diffuses with more difficulty



Ubiquity brings out which skills are more complex.

A decrease in the ubiquity of a skill shows a decrease
in the diffusion of that skill over time, and thus greater complexity, less diffusion of its knowledge.

From the graph we see that skills related to project management, Agile, are now widely diffused and therefore less complex.

CFD, Fluid Dynamics, ... on the other hand, are rapidly increasing in complexity.

Skills Graph: Microsoft Excel

Occupation	Skill	Base salary	N.	with Skills	Know-how	Skills Intensity	Cluster salary
Programmers and software development professionals	Microsoft Excel	£48,368.60	1	SQL (Programming Language)	High know-how	76,08%	£53,414.76
			2	Problem Solving	Middle know-how	48,09%	£47,999.21
			3	Management	Low know-how	45,95%	£47,585.53
Finance and investment analysts and advisers	Microsoft Excel	£38,435.30	1	Leadership	High know-how	109,41%	£47,569.77
			2	Financial Modeling	High know-how	95,75%	£45,469.41
			3	Planning	High know-how	85,78%	£43,936.74
			4	Investments	High know-how	85,47%	£43,888.45
			5	Budgeting	High know-how	82,74%	£43,469.40
Elementary storage occupations	Microsoft Excel	£22,374.85	1	Warehousing	Low know-how	47,91%	£22,187.40
						0.0 0.5 1.0	0K 20K 40K 60K

Thanks

Mauro Pelucchi
Head of Global Data Science

Mauro.pelucchi@lightcast.io
Mauro.pelucchi@gmail.com
Mauro.pelucchi@unimib.it