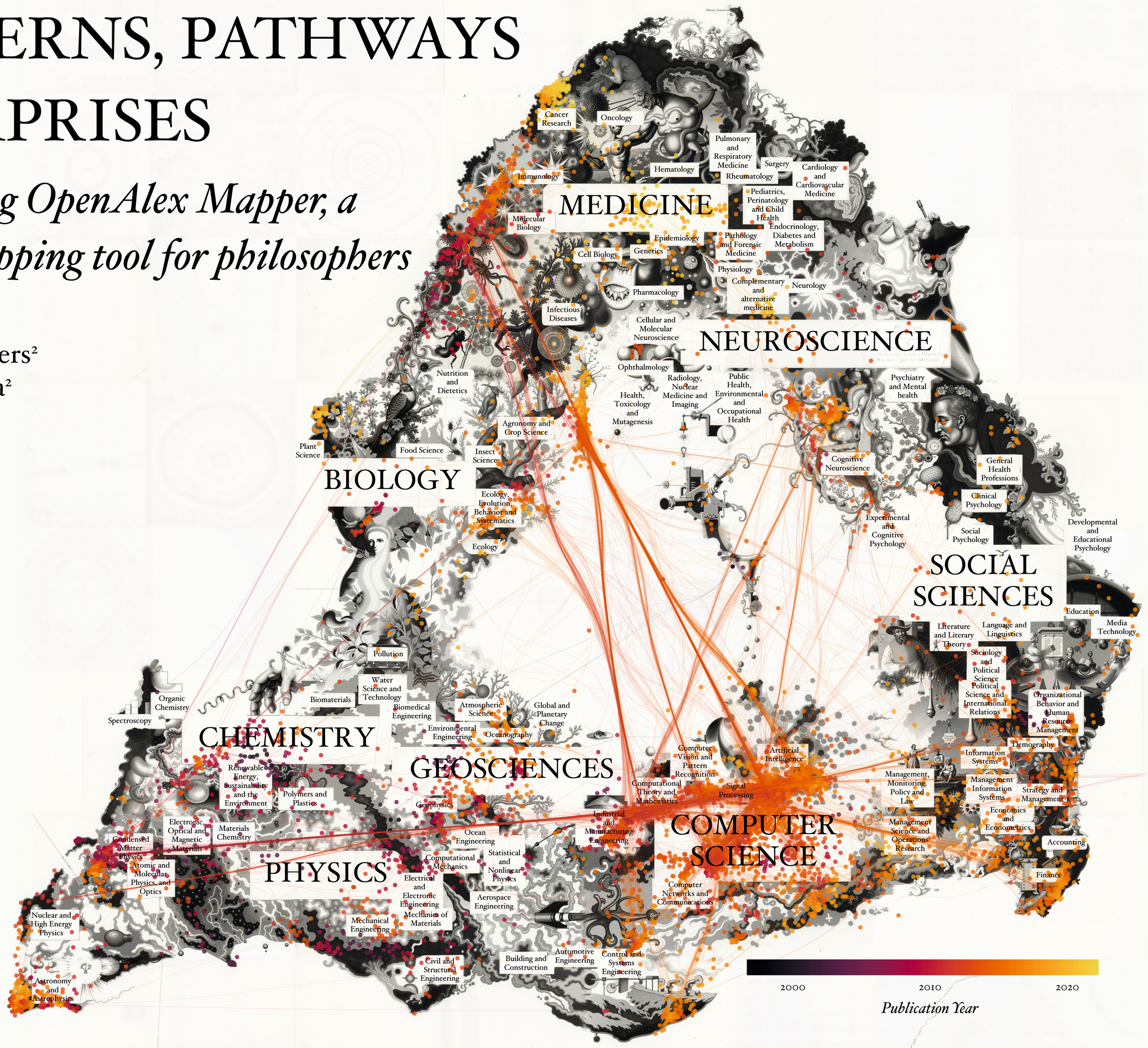
PATTERNS, PATHWAYS & SURPRISES

Introducing OpenAlex Mapper, a science-mapping tool for philosophers

Max Noichl[1]
Andrea Loettgers[2]
Tarja Knuuttila[2]



WHAT IS THIS? This is a map of science, onto which we project the history of scale-free networks (Fig. 1), an influential model template from network science (Barabási and Albert, 1999). Data points are mentions of the model, links are citation-relationships. We see how the model starts out in physics (dark red), and then settles in computer science (orange), from where it branches out into a variety of disciplines, including neuroscience, epidemiology and finance (yellow). We also find very early applications in molecular biology.

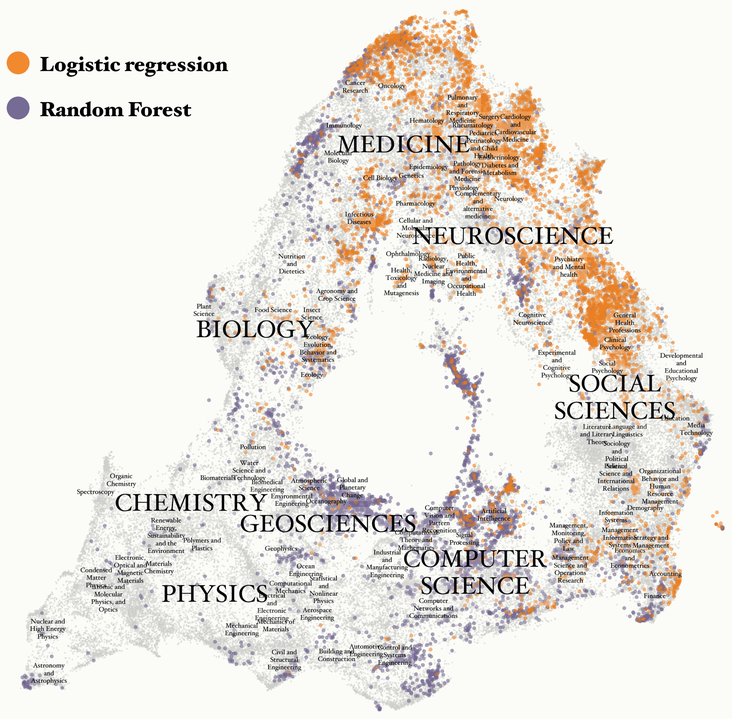


Fig. 2. Distribution of two analysis techniques that can be applied to similar problems. One (LR) is more strongly associated with statistics, the other (RF) with machine learning. We perceive a clear separation of modeling cultures across different areas of science.

HOW DOES IT WORK? We built a base map of the scientific literature by drawing a random sample of three hundred thousand English language articles from the OpenAlex database of science. We then encode their abstracts with an LLM custom fine-tuned for scientific texts and produce a two-dimensional map of them using UMAP (McInnes et al., 2018), which tries to preserve the semantic similarity between the abstracts. This map gets labelled using the categories provided by OpenAlex. We built an interactive, online tool – OpenAlex Mapper – that allows users to project arbitrary search queries, which can be for models, concepts, methods, institutions, citations, etc., onto this base map. This projection answers the question: where would the papers returned by the query end up, if they had been included in the construction of the base map. For this project we use full-text search for scale-free network as the query. The projection, together with a citation graph, lets us see the model's disciplinary spread and the local practices that come with it.

How can we trace the dissemination of models, the use of concepts, the uptake of methods across all of science?

WHY DID WE BUILD THIS? Philosophers have been interested in interdisciplinary model transfer. The frequency of exchange of formal methods between scientific disciplines has been taken as suggestive of a reorganisation of the scientific landscape (Humphreys, 2002), a prime mover of the unification of science, a puzzle for the analysis of scientific representation (Knuuttila & Loettgers, 2016; Houkes & Zwart, 2019) or a vehicle of scientific imperialism (Bradley & Thébault, 2017). But prior work on interdisciplinary modelling tends to make use of a handful of relatively narrow historical case studies, like the Ising- or Lotka-Volterra models. This makes it hard to judge the validity and generalisability of these philosophical conclusions – especially in view of the enormous size of contemporary, global, rapid discovery science. In the case of scale-free networks, we find that the model distribution is mediated by a central core of CS and applied mathematics. Less theoretical universality; instead, a diversity of modeling practices spreads the model across disciplines. Interactive computational methods like ours can help to fill in this picture. The method we introduce isn't limited to models, but allows easy investigation of all kinds of philosophically interseting entities found in the scientific literature – theories, methods (see Fig. 2) and concepts. Scan the QR code to explore it yourself!

www.maxnoichl.eu/openalexmapper

Fig. 1. Distribution of papers referring to scale-free networks, color-coded by publication year. The background of the map stems from an associated art project, in which the sciences were illustrated through AI image generation.

[1] *Utrecht University,* [2] *University of Vienna*