Tracing the Trajectories of Knowledge: A Systematic Analysis of Knowledge Mobility Patterns in Scientific and Cultural Texts

Marc Santolini (1, 2), Liubov Tupikina (1, 2, 3)
(1) Université Paris Cité, Inserm, System Engineering and Evolution Dynamics, F-75004 Paris, France
(2) Learning Planet Institute, F-75004 Paris, France
(3) Nokia Bell labs, France

In this work, we operationalize spatial movement metaphors commonly used to describe scientific discovery by systematically investigating knowledge mobility patterns. Utilizing low-dimensional embedding techniques, we create a knowledge space composed of 1.5 million articles from physics, computer science, and mathematics [1,2]. Analyzing individual researcher's textual contributions, our study uncovers striking patterns in knowledge mobility, which resemble physical mobility. We find that individual trajectories collectively form 'mobility flows.' These flows can be depicted via a gravity model, illustrating that knowledge leaps are more probable within high-density areas and less likely across extended distances. Our study also identifies a dichotomy of two primary categories of knowledge contributors: interdisciplinary 'explorers', who lead in new domains, and 'exploiters', who focus on their areas of expertise. These findings offer new ways to represent knowledge evolution, contrasting with methods used in tracking trends in art historical data [3]. Finally, we offer perspectives on how this approach can be applicable to texts that encapsulate broader cultural philosophies, including religious scriptures and wisdom traditions, exploring its applicability in representing collective trajectories in cultural evolution. In sum, this research contributes a robust and versatile analytical framework to the study of knowledge mobility and evolution.

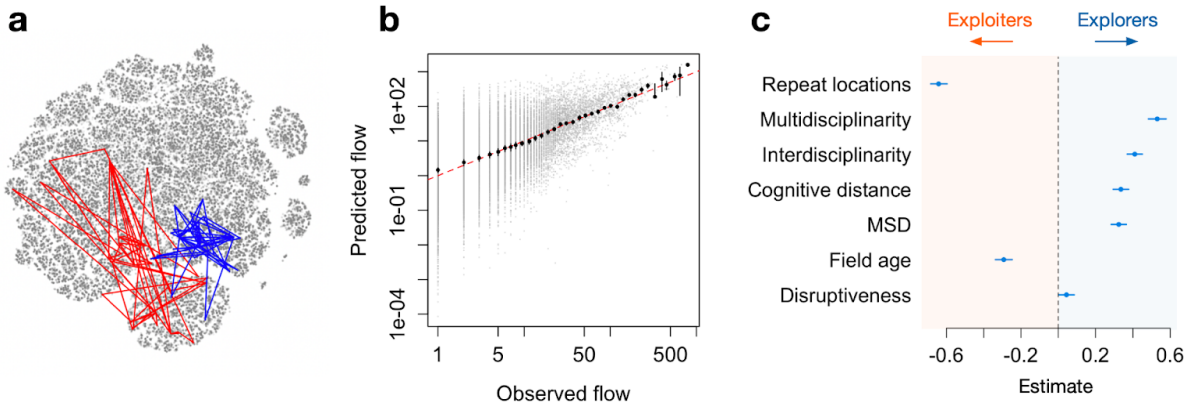

Figure 1. a. We use the metadata from 1.45 million articles posted on the arXiv, corresponding to the article field tags, authors, and timestamp. We build a high-dimensional 175 space where each article is uniquely mapped through field tags corresponding to orthogonal dimensions. This high-dimensional space is finally embedded within a 2-dimensional knowledge space using the tSNE algorithm. Each point represents an article. The sequence of arXiv pre-prints of each researcher identifies a unique scientific trajectory in the knowledge space. We show two example trajectories indicating different behaviors of researchers (in red and blue). **b.** By dividing the projected space into a grid, we can calculate aggregated densities and study their effect on flow patterns between grid elements. We show a comparison between predicted and observed mobility flows using a gravity model at a grid size resolution 10x10. **c.** We compute a logistic regression of a binary variable y indicating whether an individual is an explorer (y = 1) or an exploiter (y = 0), for different characteristics of the researchers. For each attribute (rows), we show the estimate and

95% confidence interval of the standardized coefficient of the regression, controlling for the number of articles and the main field of interest of the researcher.

References

[1] C. Singh, L. Tupikina, M. Starnini, M. Santolini “Charting mobility patterns in the scientific knowledge landscape” Nat.Comms arxiv.org/abs/2302.13054 arxiv, under rev. (2023)
[2] C. Singh, E. Barme, R. Ward, L. Tupikina, M. Santolini “Quantifying the rise and fall of scientific fields”, Plos One 17(6): (2022)
[3] A. Karjus, M. Canet Solà, T. Ohm, S. E. Ahnert, and M. Schich, "Compression ensembles quantify aesthetic complexity and the evolution of visual art," EPJ Data Science, 12, 21 (2023)