# The Global Digital Technology Space and Its Relations to Digital Inequalities

*Keywords: Digital Development, Digital Inequalities, Technology Evolution, Network Analysis, Stack Overflow*

## Extended Abstract

Our research develops a new theoretical and methodological framework to explain spatial inequalities in the global digital economy and their changes over time. Many experts and pundits believed that the open internet would lower the barriers for individuals and small businesses in low, middle, and high-income countries to participate in the global economy (e.g., Greene, 2021). Yet, digital innovation centers remain spatially highly concentrated in a few core regions. Global digital inequalities are declining for some technologies and locations but are widening in others (e.g., Hilbert, 2016; Bauer, 2018). Digital divide research points to social and material conditions as well as to individual motivations and skills to explain these observations (e.g., Helsper, 2010; van Dijk, 2020). Our research chooses an alternative, complementary route that combines insights from digital inequality research with computational methods.

To better understand the emergent, global dynamics of digital inequalities, we propose to examine digital technologies as networks of knowledge, capabilities, and practices that evolve over time. We label this network the "technology space", a term inspired by recent work in development economics that successfully used models from network and complexity theory to explain the process of economic development (Hidalgo et al., 2007). We apply this framework to digital technologies by conceptualizing those technologies and the relations between them as a network of knowledge and capabilities. This space forms a nested system. Constituent digital technologies and their uses are of varying complexity and linked in a multifaceted system of relations. Global inequalities in digital development and their changes over time can then be analyzed by examining the positions of countries in the digital technology space.

Because the internet uses transparent, open protocols, its architecture and protocols are well understood and knowledge about them is widely diffused. Capabilities to design applications and services are also widely available, facilitated by the modular architecture of the internet. The same is not true for emerging programming languages, tools, and applications. In these areas, part of the relevant knowledge is not codified but tacit, private to the individuals and groups who have obtained it (Bauer & Prado, 2023). Although it may eventually become codified, during the development phase of software and applications it is communicated and shared in networked communities of practice. These networks contribute to a heterogeneous structuring of the technology space that conditions and shapes global digital inequalities.

To construct the network representing the digital development of nations, we scraped questions and answers posted on Stack Overflow between 2008 and 2018, their timestamp, as well as the geolocation of the authors of each of the 24 million observations in our data. We treat the tags associated with questions as indicators of the presence of a particular technological skill in a country and the number of occurrences as indicative of the prevalence of that respective skill. The network itself was constructed by using the co-occurrences of the most frequent tags. To separate signal from noise, we used a network backboning approach suggested by Coscia and Neffke (2017). To find technologies which are overexpressed in one country compared to world averages, we followed an approach suggested by Balassa (1965),

applied by Alabdulkareem et al. (2018), and calculated the revealed comparative advantages for every technology. This analysis enabled us to trace a country's development in the technology network over time and compare it to competitors/other countries. By calculating the distances between different countries' individual "fingerprints" (Coscia et al., 2020), we were able to identify countries of similar characteristics. The results are visualized in Figure 1 for 2010 and 2018.

The positions of countries in the technology network can serve as indicators of global digital inequalities. Figure 1 shows, in a highly condensed way, that membership in the groups varies, with some countries improving their position over time (e.g., China, Argentina, Egypt) and others losing ground (e.g., South Africa). This approach has advantages over traditional measures of digital inequality that are based on frequency distributions parsed by sociodemographic and other factors. In this framework, we will then analyze digital policy measures to understand whether these have helped countries to reach other parts of the technology space. This will allow to develop nuanced policy implications for strategies that are appropriate for countries in different regions of the technology space. It will also allow to to explore in more detail how different socio-economic, political, and cultural contexts interact with the conditions of the technology space.

# References

Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking the polarization of workplace skills. *Science Advances*, *4*(7), aa00630.

Bauer, J. M. (2018). The Internet and income inequality: Socio-economic challenges in a hyperconnected society. *Telecommunications Policy*, *42*(2), 333-343.

Bauer, J. M., & Prado, T. S. (2023). Digital innovation: An information-economic perspective. In D. R. Raban & J. Włodarczyk (Eds.), *The Elgar companion to information economics*. Edward Elgar.

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage. *The Manchester School*, *33*(2), 99–123.

Coscia, M., & Neffke, F. M. (2017). Network backboning with noisy data. *33rd International Conference on Data Engineering (ICDE)*, pp. 425-446, IEEE.

Coscia, M., Gomez-Lievano, A., Mcnerney, J., & Neffke, F. (2020). The node vector distance problem in complex networks. *ACM Computing Surveys*, *53*(6), 1–27.

Greene, D. (2021). *The promise of access: Technology, inequality, and the political economy of hope*. MIT Press.

Helsper, E. J. (2012). A corresponding fields model for the links between social and digital exclusion. *Communication Theory*, 22(4), 403-426.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482-487.

van Dijk, J. (2020). *The digital divide*. Polity Press.

Hilbert, M. (2016). The bad news is that the digital access divide is here to stay: Domestically installed bandwidths among 172 countries for 1986–2014. *Telecommunications Policy*, 40(6), 567-581.

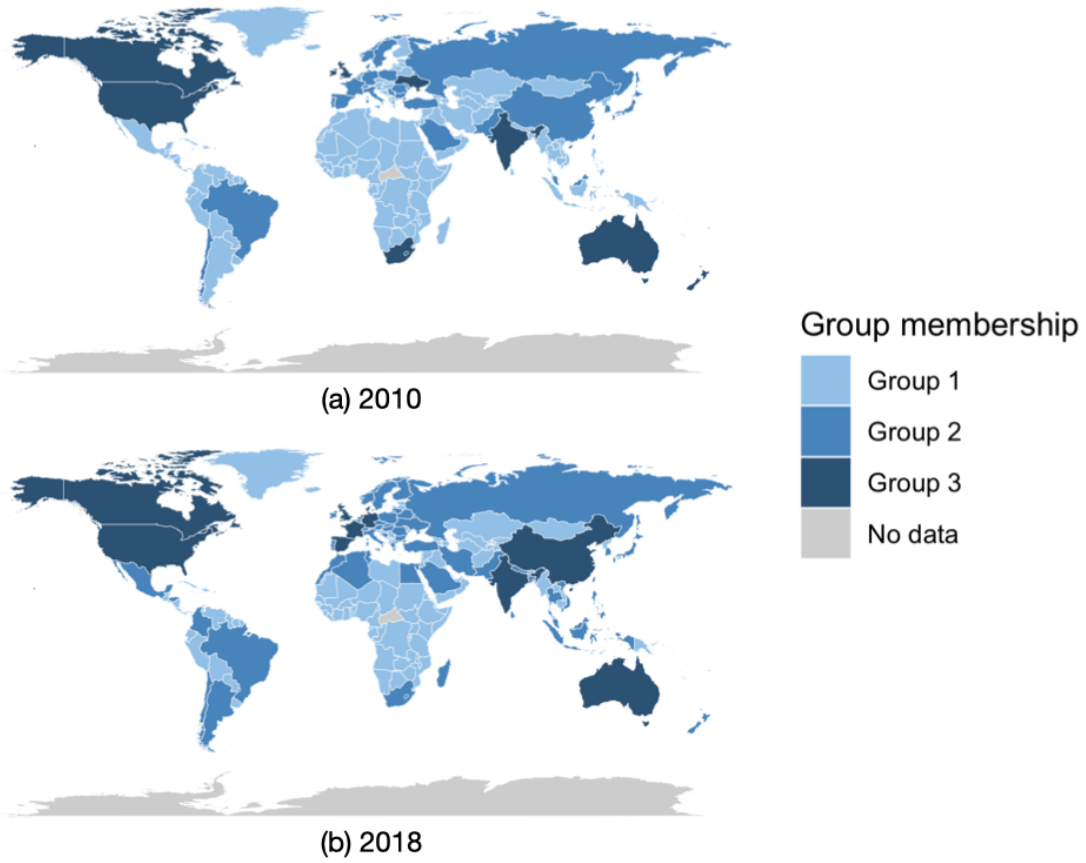


**Figure 1.** *For this visualization, we analyzed Stack Overflow tags related to 'Machine Learning' (ML) between 2008 – 2018. Countries are grouped by similar technological fingerprints. Group membership was obtained by applying hierarchical clustering on the distances between different countries' technological 'fingerprints', i.e., their respective individual occupations of the technology space. For our data, we observed a clear clustering into three groups.*