

Twin Transition Technologies in the UK through Attributed Spatial Networks

Presenting author: Emerald Dilworth, University of Bristol

Co-authors: Emmanouil Tranos and Daniel Lawson, University of Bristol

When individuals and companies struggle to keep up with new technologies, there can be a wide range of barriers stifling their access, and in turn, many consequences from failing to keep up with the cutting edge. As we find ourselves in the midst of the fourth industrial revolution, we look to firms involved in green and digital technologies in the UK to better understand how and where twin transition technologies are developing. Where there are developments in technologies to digitise and green the economy, this twin transition (TT) leads to sustainable economic growth and multiple environmental benefits.

Traditionally researchers have relied on surveys, patent data, and place attribute-based data to investigate such problems. Instead we utilise the “digital breadcrumbs” created simply from the use and adoption of the Web. From vast quantities of web crawl data, we construct directed, weighted networks of hyperlink connections at various levels of granularity, in addition to utilising the webtext to inform if a website has an interest in green and/or digital technologies. We do this by exploiting existing Large Language Model (LLM) semantics information to apply term frequency–inverse document frequency (tf-idf) in order to explore if firms are involved with twin transition technologies.

Our paper illustrates that digital technologies provide a vital data resource for understanding the evolution of economic relationships [?, ?]. Exploiting such data as a way of observing at a very high level of granularity can reveal the spatial structure of such economic relationships, and importantly allow these relationships to be contextualised. We use these web data to create meaningful geographic knowledge at the postcode level about urban and regional inter-dependencies within the UK.

Every 4 to 6 months the Common Crawl makes a dump of ‘webgraph’ data from its crawls – that is all the hyperlinks between archived websites. We analyse one such webgraph that spans the crawls from May 2022 to August 2022. We interpret this webgraph as a directed network of websites and their hyperlinks, which spans millions of websites. We only consider websites anchored in the United Kingdom (UK), i.e. those with the ‘.uk’ top level domain (TLD) and, therefore, our network is a directed network between all the archived ‘.uk’ websites.

Importantly, our network is a spatial network as we are able to geolocate the archived websites – i.e. the nodes of the networks – using postcode references found in the HTML text. Websites often have more than one postcode found among their many webpages crawled. There are various ways we can interpret what the number of postcodes we find for a website implies. Not all websites use any schema.org vocabulary or reference any postcodes in their text, or either.

All in all, we are able to create a *spatial* and *attributed network* containing all the hyperlinks where the websites at both ends of the link have a user defined number of postcodes

associated with them. This is a directed network, which is a subgraph of the whole webgraph, where each website is geolocated to one or more places in the UK.

This work combines using both network data and attribute data to examine the digital connectivity landscape of the twin transition in the UK. We model digital and green clusters and relate this to geography to understand the role space plays and the power of existing business agglomeration forces in shaping these spaces.

Such data can reveal the spatial structure of those involved with TT technologies, and importantly allow these relationships to be contextualised. We use these web data to create meaningful geographic knowledge at the postcode level, and highlight where and why there is big growth for twin transition technologies, as well as where there is a struggle. Our empirical findings have the potential to support and inform city and regional policies.