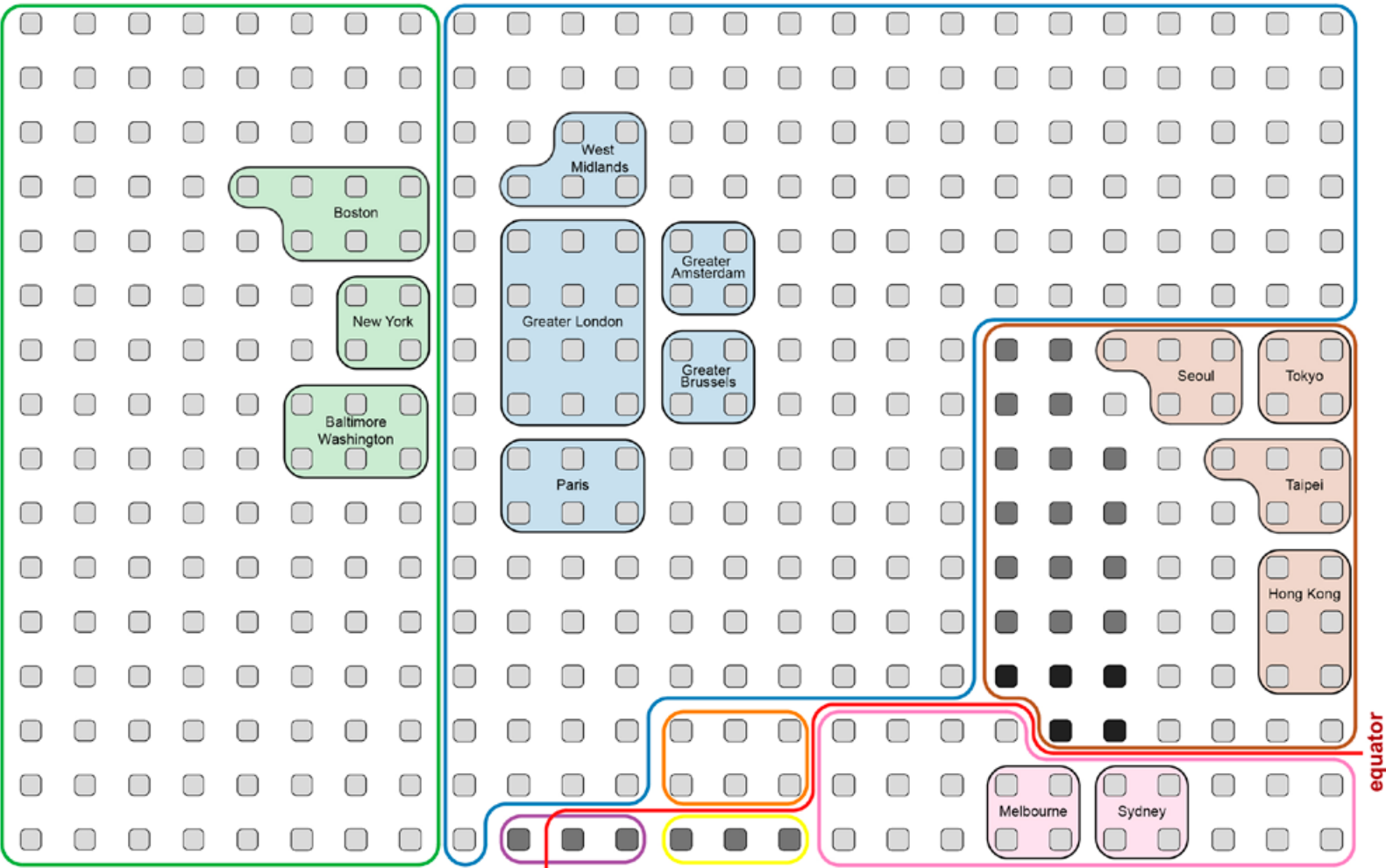


Uneven Digital Geographies

... and Why they Matter

Mark Graham, Stefano De Sabbata, Sanna Ojanperrä

markgraham.space



Top 400 Universities
according to
Times Higher Education
2013-2014

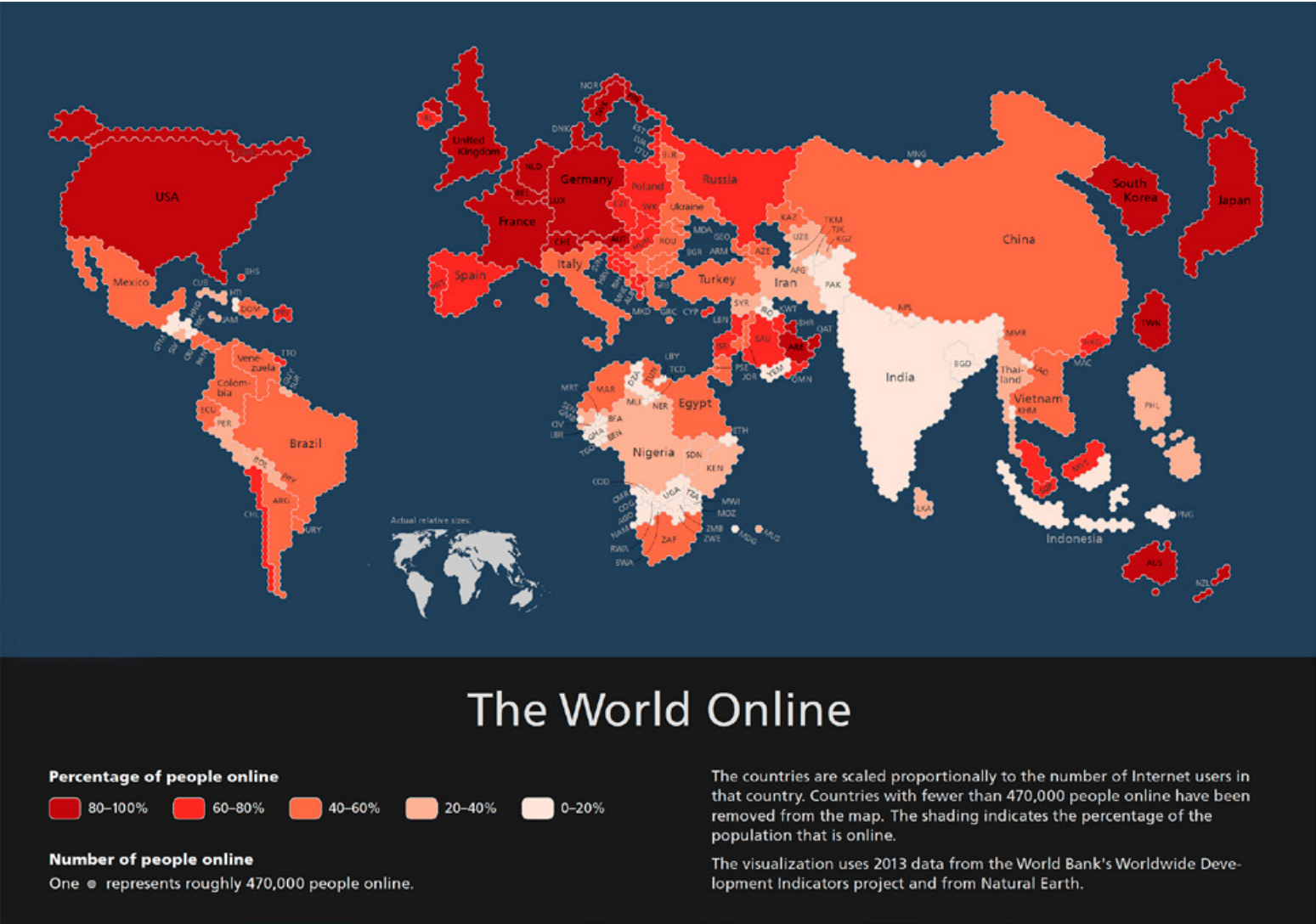
Information has always been spatial. It is produced somewhere; it is used somewhere; it moves between places (Graham et. al., 2015a). And the geographies of information have always been imbricated in power relationships. Some people have far more control over it than others, and some places are central in information ecosystems whilst others are peripheral.

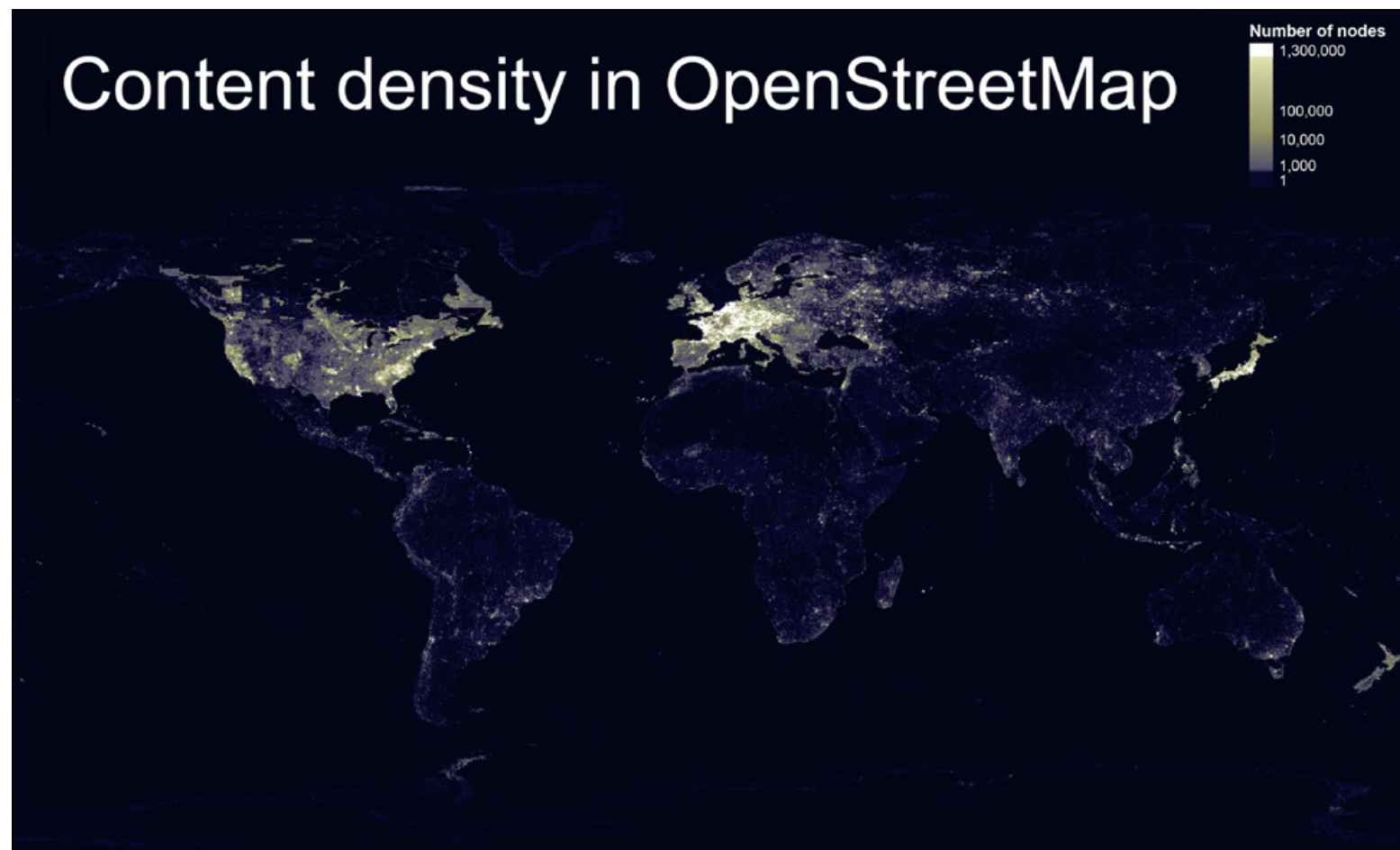
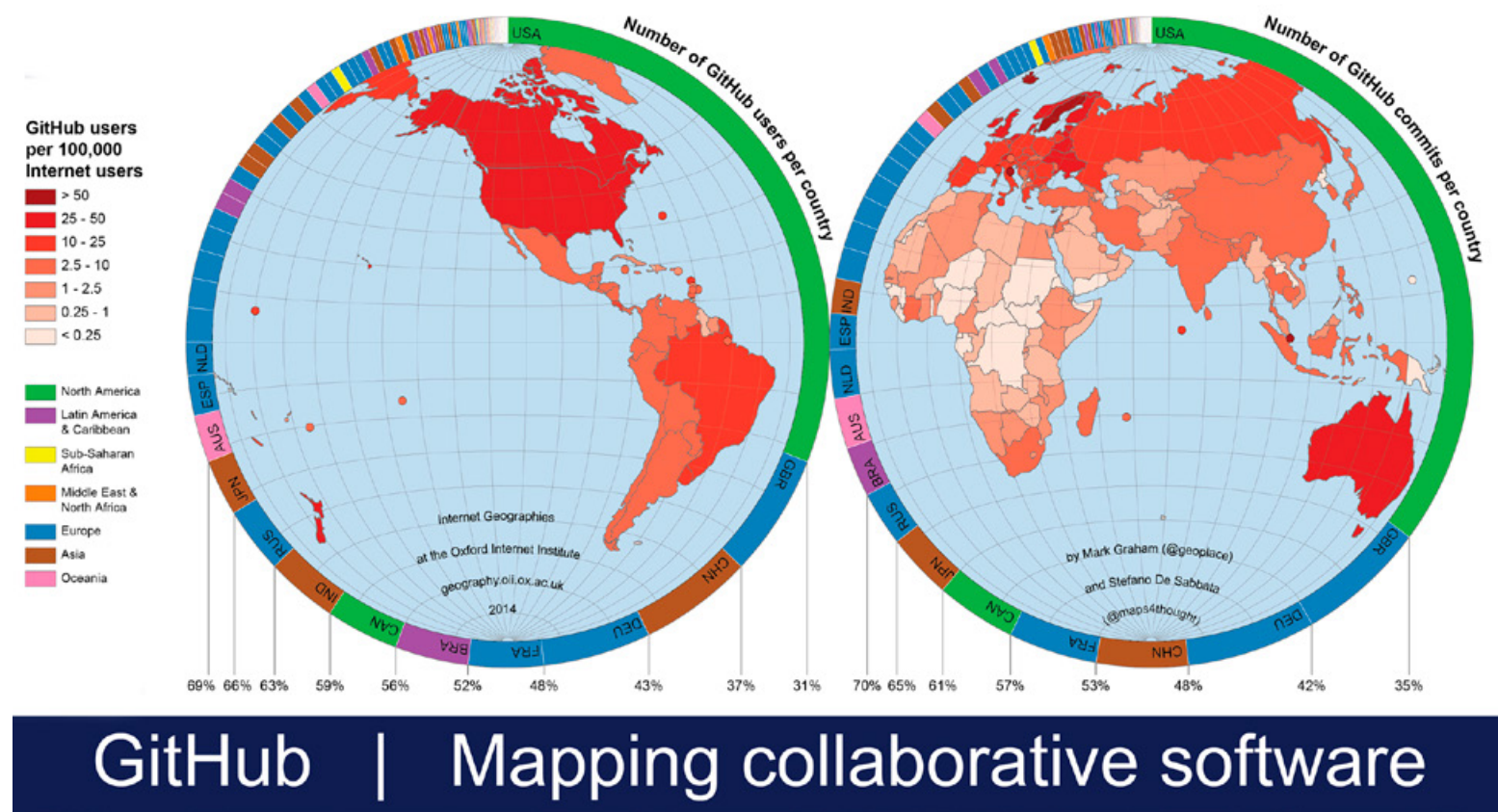
For instance, the map Top 400 Universities (see opposite page) shows the locations of the world's top 400 universities as ranked by the Times Higher Education. It also illustrates the relative wealth of the country that hosts each university. There are no universities from low-income countries present on the list, and India is also the only lower-middle income country represented, being home to five of the world's top-400 ranked universities. Most of the world's elite universities are in the Global North, most of the world's published academic knowledge is produced in the Global North, and even acceptance rates for most journals tend to be higher for authors from the Global North. Amazingly, the Greater London cluster alone, contains the same number of top-400 universities as all of Sub-Saharan Africa, the Middle East, and Latin America combined! We have a state of affairs where the Global North has tended to be a producer of knowledge and the Global South, a consumer of it.

Relatedly, the maps Submission by Country... and Acceptance Rate by Country ... (Graham, 2015a) were created from some submission data that SAGE journals shared with me. Amongst other things, the data tell us where authors of articles come from, and primary discipline of the journal they are submitting to.

We see much more academic content coming from the Global North than from the Global South. Africa in particular is notable for its absence. Most countries on the continent fail to register even a single journal article submission. Not only do a lot of countries in the South have particularly a low number of submissions, they also have very low acceptance rates for the small numbers of submissions that they do have: further deepening the geographic divides in knowledge production.

However, many have pointed to the Internet as a way of transcending some of these traditional constraints. Access to the Internet, in theory, allows users to access the sum of all codified human knowledge; it allows people to participate in a more level playing field. This is because there are relatively few geographic barriers to information flow over the internet. With a few exceptions (notably China and a few other authoritarian regimes), content like a Wikipedia page or Google Book is equally accessible to anyone on Earth.





But it is worth remembering that the internet, and information within it, is also characterised by real geographic inequalities. A majority of humanity has still never used the internet, and some parts of the world have very little representation in our digital world.

But, even with those imbalances in mind, it is worth remembering that there are about four billion internet users in the world. All of those people can potentially contribute to the wealth of information that we all share and use on the Internet. The problem is that they don't.

Digital Participation

We can look at one of the world's biggest and best-known hosting services for software development projects, for instance: GitHub. The shading of the map *GitHub | Mapping collaborative Software* (see opposite page bottom) illustrates the number of GitHub users as a proportion of each country's Internet population. The circular charts surrounding the two hemispheres depict the total number of GitHub users (left) and commits (right) per country.

North America and Europe each account for about one third of the total number of GitHub users. The Middle East, North Africa, and Sub-Saharan Africa together represent less than 1% of GitHub users, and just about 1% of commits. Switzerland alone counts almost as many GitHub users as the Middle East and North Africa region, and more than Sub-Saharan Africa. The geography of digital engagement in this facet of the "knowledge economy" is thus starkly uneven.

Wikipedia is another useful example of a platform that in theory allows anyone in the world to submit information to it. In practice, though, we also see massive inequalities in the amount of content submitted to Wikipedia from different parts of the world. The vast majority of Wikipedia is written by people in the Global North, and only a tiny amount of content comes from people in the South (see Graham et. al., 2016) for more on the topic). This matters because editors from the North can easily overpower editors from the South when writing about contested topics.

The Middle East is perhaps the part of the world where we see these divides most starkly manifested. There are almost as many edits that come from Israel as the entire rest of the region combined: from Morocco in the West to Iran in the East (Graham, 2012). Thus people in just a few parts of the world seem to have much more voice than everyone else.

Digital Representation

Not only are some parts of the world left out of practices of digital participation, some parts of the world are also covered by much thinner augmentations of information than others. These layers of information help to comprise and define a place: it is thus important to understand not just where they come from, but also where they represent.

The map *Content density in OpenStreetMap* (see opposite page top) shows the location of edited content in the world's largest collaborative mapping project: OpenStreetMap. In OpenStreetMap, high-income OEDC countries are home to about 80% of the submitted content. We thus end up with comparisons like the fact that Egypt accounts for as many nodes as Iceland, despite being 10 times as big and being home to 250 times the population.

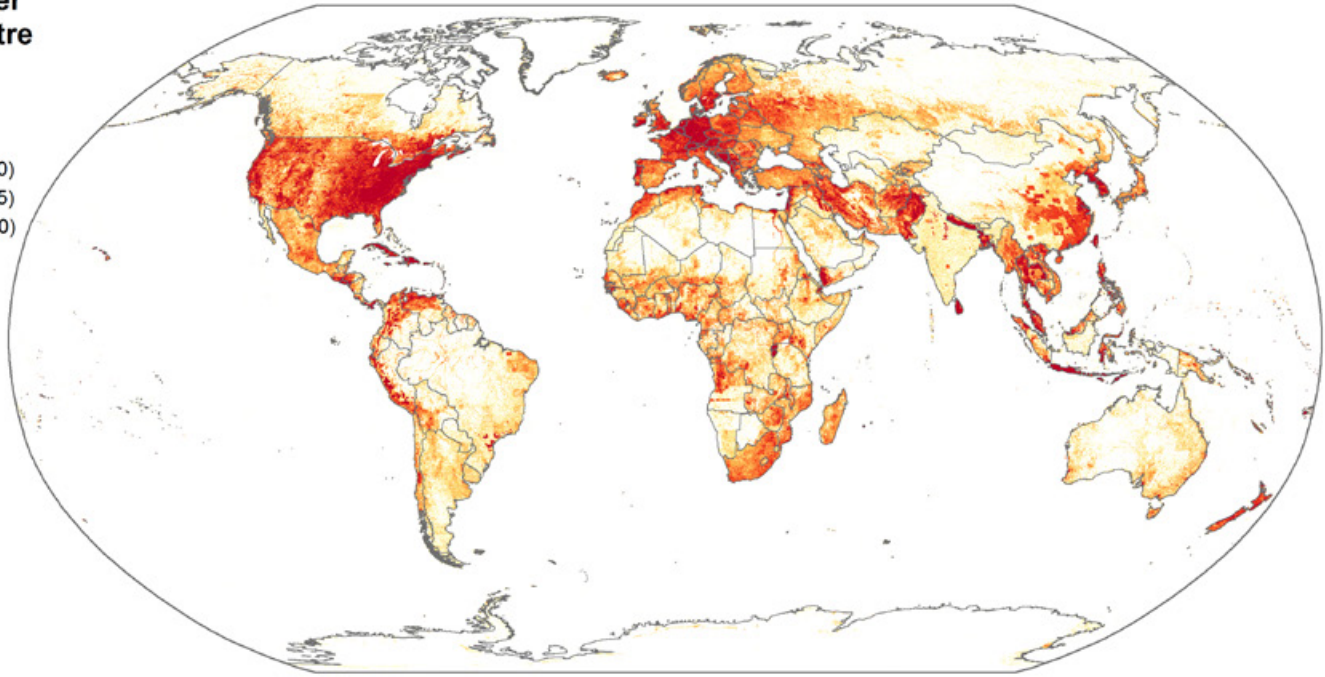
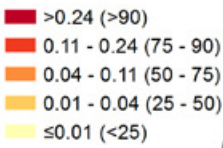
We can see similar uneven geographies of digital representation if we look at GeoNames (see following page), which is the world's largest freely available gazetteer (i.e., a dictionary of geographic place names). The pixel colours represent the number of names referring to a geographic place) per spatial unit: a square of one tenth degree of latitude and one tenth degree of longitude (see Graham and De Sabbata, 2015) for more on the method).

The US accounts for slightly more than a quarter of the database. There is actually more content created about the US than all of Asia combined (Asia accounts for only about 23% of geographic content, despite being home to over half the world's population). Interestingly, the information presences that we see are characterized by unusual patterns. Not only do we see the usual suspects of Western Europe and the United States with large amounts of geographic information, but we also see significant densities in places like Sri Lanka, Iran, and Nepal. By defining structured geographic information about the world, gazetteers ultimately have the power to shape and structure how geographic meaning is made. The presences and absences of data within shape how the world is digitally re-made.

Finally, it is worth exploring the geography of content in Wikipedia. The map *The Geographical Uneven Coverage of Wikipedia*, presented on the double page, points out the highly uneven spatial distribution of (geotagged) Wikipedia articles in 44 language versions of the encyclopaedia. Slightly more than half of the global total of 3,336,473 articles are about places, events and people inside the red circle on the map, occupying only about 2.5% of the world's land area (see Graham et. al., 2014) for more on this research). Much like GeoNames or OpenStreetMap, Wikipedia plays an important role in shaping how we understand the world. The geography of content within it, therefore matters immensely.

Placenames per square kilometre in GeoNames

(Percentiles)



Information Geographies and the Geographies of Information

It is worth remembering that the geography of information production has always been characterised by significant geographic biases. But the fact that billions of people are now connected to the internet was seen by many as a moment of change. The internet has been described as a "leveller" and a "democratiser" (c.f. Graham et. al., 2015b) – allowing anyone to access what Wikipedia refers to as "the sum of all human knowledge"; allowing anyone to contribute. If users in Manchester, Mombasa, and Mumbai are connected, there should be no difference in their propensity to access and created digital knowledge, right?

In practice, we see a very different world from that vision. A world in which some places are far more visible than others. A world in which people in some parts of the world have a much bigger say in how our digital environments are constructed (see also Graham, 2015b).

Let's remember that the places that we live in are increasingly digital. Our cities are no longer just made of bricks, mortar, glass, and steel. They are also made of data (Graham, 2013; Graham et. al., 2013).

As such, it will continue to be hugely important to interrogate the digital layers of places. Where are they? What are they? What do, and don't, they exclude? Who constructs them, and who is sidelined? And who controls them (Shaw & Graham, 2017)? These are the questions that we need to be asking if we ultimately want to work towards less uneven and more just information geographies.

References

Graham, M (2012) 'Mapping Edits to Wikipedia from the Middle East and North Africa', Online: [http://www.markgraham.space/blog/mapping-edits-to-wikipedia-from-the-middle-east].

Graham, M. (2013) 'Geography/Internet: Ethereal Alternate Dimensions of Cyberspace or Grounded Augmented Realities?', The Geographical Journal, vol. 179, no. 2, pp. 177-182.

Graham, M., Zook, M., and Boulton, A. (2013) 'Augmented Reality in the Urban Environment', Transactions of the Institute of British Geographers, vol. 38, no. 3, pp. 464-479.

Graham, M., Hogan, B., Straumann, R. K. and Medhat, A. (2014) 'Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty', Annals of the Association of American Geographers, vol. 104, no. 4, pp. 746-764.

Graham, M. (2015a) 'The geography of academic knowledge', Online: [http://geonet.oi.ox.ac.uk/blog/the-geography-of-academic-knowledge/].

Graham, M. (2015b) 'Information Geographies and Geographies of Information', New Geographies vol. 7, pp. 159-166.

Graham, M., De Sabbata, S. and Zook, M. (2015a) 'Towards a study of information geographies:(im)mutable augmentations and a mapping of the geographies of information', Geo:Geography and Environment, vol. 2, no. 1, pp. 88-105.

Graham, M., Andersen, C., Mann, L. (2015b) 'Geographical Imagination and Technological Connectivity in East Africa', Transactions of the Institute of British Geographers vol. 40, no. 3, pp. 334-349.

Graham, M., De Sabbata, S. (2015) 'Mapping Information Wealth and Poverty: The Geography of Gazetteers', Environment and Planning A, vol. 47, pp. 1254-1264.

Graham, M., Straumann, R. K., Hogan, B. (2016) 'Digital Divisions of Labor and Informational Magnetism: Mapping Participation in Wikipedia', Annals of the Association of American Geographers, vol. 105, no. 6, pp. 1158-1178.

Shaw, J., Graham M. (2017) 'An Informational Right to the City? Code, Content, Control, and the Urbanization of Information' Antipode, doi:10.1111/anti.12312.

Illustrations

Map first spread left by Mark Graham (@geoplace) and Stefan De Sabbata (@maps4thoguht), Internet Geographies at the Oxford Internet Institute, 2013, geography.oi.ox.ac.uk
data source: Times Higher Education, timeshighereducation.co.uk, Oxford Internet Institute, University of Oxford

Map first spread right by Ralph Straumann, Mark Graham, geonet.oi.ox.ac.uk
Map previous spread top: Oxford Internet Institute, University of Oxford
data sources: GitHub.com; World Bank, data.worldbank.org

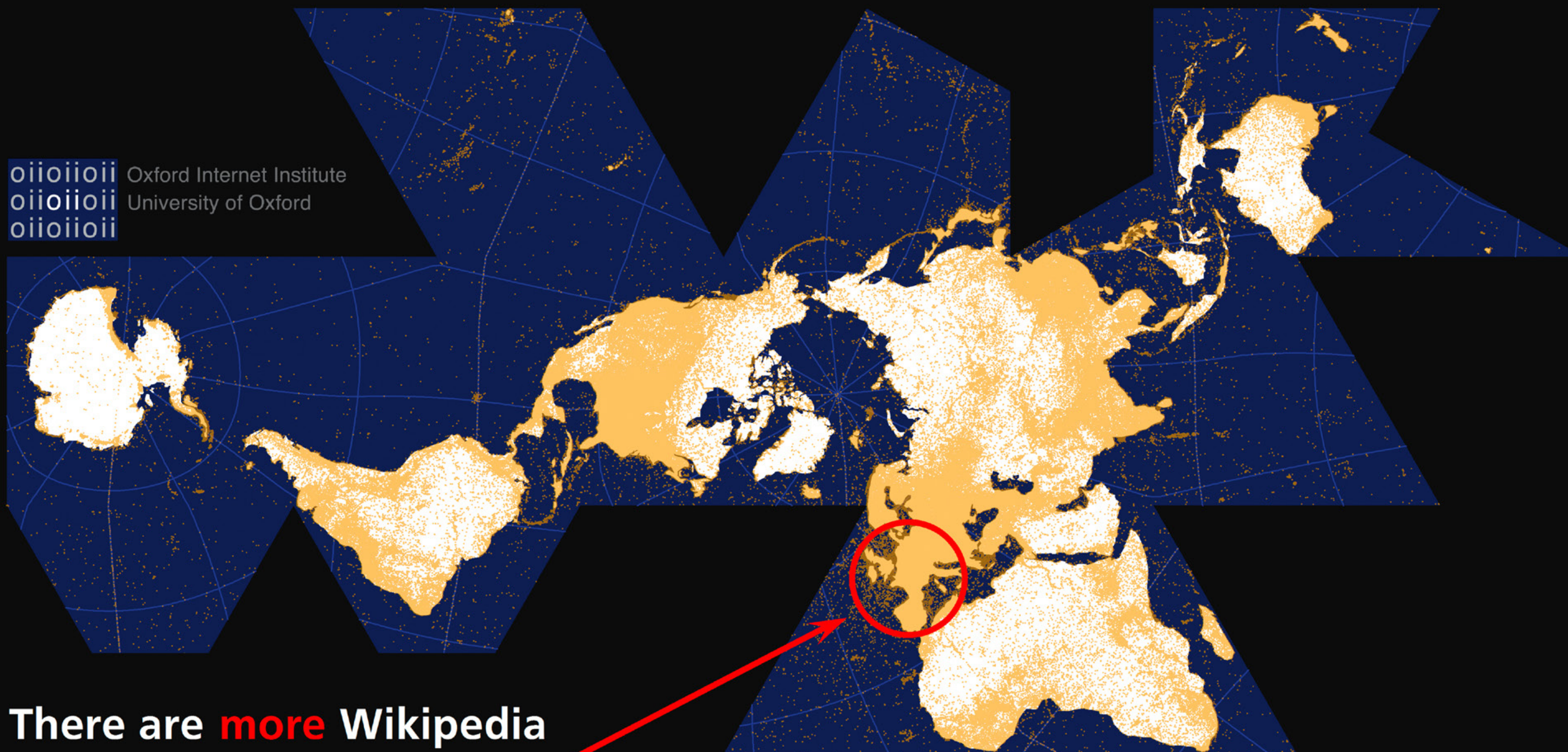
Map previous spread bottom by Mark Graham (@geoplace) and Stefan De Sabbata (@maps4thoguht), Internet Geographies at the Oxford Internet Institute, 2014, geography.oi.ox.ac.uk
data sources: openstreetmap.org; geofabrik.de
Data as from ".*-latest.osm.bz2" files available per world region on download, geofabrik.de on December 12, 2013



The Geographically Uneven Coverage of Wikipedia

While it is an invaluable resource of knowledge for numerous users, Wikipedia's articles have a strong bias in their geographic distribution. Below map is based on 3,336,473 geotagged articles in November 2012 data dumps of 44 language versions. Each article is represented by an orange dot.

oioioioioio
oioioioioio
oioioioioio
Oxford Internet Institute
University of Oxford



There are **more** Wikipedia
articles **inside** this circle
than outside of it

This map is part of the Information Geographies
project at <http://geography.oii.ox.ac.uk>.

CC-BY-NC
Ralph Straumann, Mark Graham
Data sources: Wikipedia, Natural Earth