Title
Open online communities as a type of knowledge infrastructure under threat

Author
Brown, Titus C.

Publication Date
2020-02-26

Open online communities as a type of knowledge infrastructure under threat

C. Titus Brown, ctbrown@ucdavis.edu

Jan 5, 2020

1. *What are the most urgent research questions to address about KI? Why?*
2. *Identify a KI whose survival is under threat.*
   a. What led to these threats? Over what time frame?
   b. What actions or changes in circumstances might lead to its survival?
   c. What will be gained or lost, by whom, if this KI fails to survive?
3. *How do KI spread information? Misinformation? Alone and in combination with other infrastructures?*

I'd like to suggest two things: one is that open source software development projects, and open online communities more generally, are knowledge infrastructures; and the other is that many of these knowledge infrastructures are under constant threat.

Open online communities are knowledge infrastructures

As specific examples to start, I'm going to use two projects with which I'm relatively familiar: one is the Python programming language, and the other is the Carpentries, a global computational skills training community. Both projects exist as open online communities, and both projects have tremendous reach and have been in existence for at least a decade.

Python and the Carpentries are both knowledge infrastructures because they encode specific knowledge about the human and natural worlds. Python is a widely used programming language that is frequently taught as an introductory language; the language encodes specific knowledge about how humans learn and use programming languages, while connecting those humans to deeply technical knowledge about how computation works, in the form of high performance libraries of data structures and algorithms. The Carpentries is a global computational skills training community that consists of approximately 2,000 volunteer instructors who give two-day hands on training workshops; the organization's lessons and instructors encode knowledge of how people learn, and the materials are targeted at (mostly) natural scientists and hence contain scientific knowledge. In both cases (and in most open online communities) the products are free and openly available for modification, distribution,, and reuse.

Both Python and the Carpentries rest on the backs of their communities. Most Python and Carpentries work is done by contributors who either volunteer for free or who include the work as part of their job. Python is governed by a board of five community members who are elected by the core developers, while the Carpentries Executive Committee is elected from the community of active instructors. Importantly, in both communities, those who do the work chart the future course of the project: both projects are self-governed. As such, I think another way in which both Python and the Carpentries (and open online projects in general) are knowledge infrastructures is that they encode a great deal of praxis about online community organization and governance.

And this leads to a research question that I find fascinating, and that I think is one valuable lens through which to view knowledge infrastructures. Zeynep Tufekci claims (and I agree) that "much of the web is an exception to the famed 'free rider' problem"[1]; and I think it is unclear how open online projects are created, expanded, maintained, and sustained. Given the increasing reliance of modern academia and industry on infrastructure built and maintained by open source projects[2], I can think of few questions that are more timely than how these projects work.

Open online communities are under threat.

Community based projects like Python and the Carpentries face a tremendous number of existential challenges. They have a considerable technical maintenance burden due to bit-rot, but volunteers who enjoy doing maintenance work are hard to find. There is no straightforward business model, since at its core their products are openly available. Community maintenance and community-based governance is another chore requiring significant effort, with little in the way of historical knowledge (given the newness of the Internet). And there are many extractive contributors who seek to take more than they give back, ranging from companies "strip-mining" openly available products for their own profit, to new community members who wish to direct the project's effort towards their own goals.

In practice, what these projects stand to lose if they don't survive is a (surprisingly effective) channeling of effort towards common goals - moreover, one that meets the needs of many people. The larger community of external "users" (academia, industry, government) will lose an effective product that iterates to meet their newly revealed needs while providing a stable platform for programming (Python), learning (Carpentries), or other products.

These projects also demonstrate and explore effective alternatives to brick-and-mortar, hierarchical, and otherwise "established" structures, often in ways that are reminiscent of a commons. Interestingly, they do so over long distances, with many community members rarely or even never meeting in person. They are, in a very real sense, communities of practice in

---

[1] https://www.wired.com/story/altruism-open-source-fuels-web-businesses-love-to-exploit-it/

[2] https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/

sustained long-distance collaboration. The conceptual innovation and practical implementation of these open online projects is worth studying in much more detail.