Understanding the 'Capacity' of the Digital Humanities: The Canadian Experience, Generalised

Cyberinfrastructure for Research in the Humanities: Expectations and Capacity

Geoffrey Rockwell, SDH/SEMI and DH2010[1]

In 1997 the Canada Foundation for Innovation (CFI) was established by Act of Parliament to fund the development of research infrastructure. Since then it has committed over $5 billion for more than 6,600 projects across Canada.

The legislation that set up the Foundation defines infrastructure thus,

> "research infrastructure" means equipment, specimens, scientific collections, computer software, information databases, communications linkages and other intangible property used or to be used primarily for carrying on research, including housing and installations essential for the use and servicing of those things." (From the Budget Implementation Act, 1997, c. 26)

This massive investment in research infrastructure was intended to build capacity for innovation, attract and retain top researchers (often from the USA), train graduate students and research staff, foster collaboration and make sure Canadian institutions made good use of research infrastructure.[2]

---

[1] A variant of this was presented at SDH/SEMI 2010 in Montreal and later at DH 2010 in London.

[2] "The CFI Story" on the CFI web site describes the intentions thus:
"CFI support is intended to:

- strengthen Canada's capacity for innovation;
- attract and retain highly skilled research personnel in Canada;
- stimulate the training of young Canadians through research;
- promote networking, collaboration, and multidisciplinarity among researchers;
- ensure the optimal use of research infrastructure within and among Canadian institutions."

CFI was a welcome and new approach to funding research after the cutbacks of the early 1990s. What was different was that this funding was for something few of us had thought about, let alone applied for, namely "research infrastructure."[3] We were supposed to apply not to do research itself, but to set up facilities that would attract researchers, train graduate students and transform our research.

The establishment of CFI anticipated what I call an infrastructure turn elsewhere in the research world, notably in the United States where the 2003 "Atkins Report" of the National Science Foundation which had the modest title, *Revolutionizing Science and Engineering Through Cyberinfrastructure*. The Atkins Report was followed by *Our Cultural Commonwealth* in 2006 that advocated for innovative cyberinfrastructure in the humanities and interpretative social sciences.

The rhetoric around infrastructure, both that in reports and in our applications, promised extraordinary transformations in research environments, and presumably also research. As David Green put it in an introductory article on the issue of cyberinfrastructure for the liberal arts, "This is going to be big."[4] He goes on to quote Arden Bement of the National Science Foundation who wrote that the Cyberinfrastructure Revolution "is expected to usher in a technological age that dwarfs everything we have yet experienced in its sheer scope and power."[5]

The centerpiece for this revolutions, at least for us in the humanities and social sciences was to be the Bamboo project that, with funding from the Mellon

---

[3] According to a Treasury Board *Evaluation of Foundations*, "CFI was described, at that time, as an entirely new approach by the government to the support of research and development. From this starting point, involving a once-off investment of $800 million, the federal government went on to create a variety of foundations that either receive conditional grants for disbursement over a finite number of years or to create perpetual endowments that use the income generated by the endowment to fund their disbursement programs and operations."

[4] This is how Green starts his article "Cyberinfrastructure For Us All: An Introduction to Cyberinfrastructure and the Liberal Arts" which introduces a special issue on the subject. Other articles in the issue are also worth reading. See <http://www.academiccommons.org/issue/december-2007-cyberinfrastructure-and-liberal-arts>

[5] The Bement quote is from remarks he gave on "Shaping the Cyberinfrastructure Revolution: Designing Cyberinfrastructure for Collaboration and Innovation" which have been published in *First Monday.*

foundation, started a large international consultation with a view to creating a consortially maintained infrastructure layer in the cloud for humanities research.

Then the Great Recession came along and many of our dreams, including those for the Bamboo project, have had to be scaled back. Foundatins and governments that we were hoping could be convinced to provide long-term infrastructure funding are cutting budgets to areas considered higher priority.

Perhaps cyberinfrastructure will eventually be big, and it is likely that here in Canada we will see more rounds of CFI funding, but none the less the time for some realistic reassessment of infrastructure and research has come. Let us take this moment of recession to reassess. I therefore want to make a series of points about cyberinfrastructure:

1. First, we don't really know what should be cyberinfrastructure because, unlike traditional infrastructure, there isn't a tradition that defines where the line is between project and general technology. Most of us would agree that the Internet should be treated as infrastructure, but after that what? There has been remarkably little interest in learning from the histories and the economics of infrastructure. When does it make sense to stop funding something as equipment needed for a particular project and start supporting it widely and properly as infrastructure with professional staff and continuous funding? Who is asking those questions?
2. Second, and related to this is a warning. We should not be too quick to turn into infrastructure those technological services that are still a site for research and negotiation. To turn something into infrastructure typically means removing it from control by researchers so that it can be centrally and professionally run, which, when it is something that is an indisputable standardized need, is exactly what we want. But many infrastructure projects fail because it isn't actually clear what the research community will use and projects by powerful constituencies get dressed up as everyones' needs.

Edwards et al., in a must-read report titled *Understanding Infrastructure: Dynamics, Tensions, and Design,* that came out of a NSF workshop bringing historians and social scientists to bear on cyberinfrastructure, argue the importance of not turning too quickly to infrastructure,

> Given its relative immaturity and the rapidly changing technological backdrop against which cyberinfrastructure is unfolding, efforts not to prematurely "sink" or "fix" the form and vision of cyberinfrastructure (or distinct cyberinfrastructure projects) should be supported. (p. 42)

They also point out something that is especially true in our community, namely that infrastructure is social,

> It is also possible that a tech-centered approach to the challenge of data sharing inclines us toward failure from the beginning, because it leaves untouched underlying questions of incentives, organization, and culture that have in fact always structured the nature and viability of distributed scientific work. Questions of trust loom large here, and run both ways. (*Understanding Infrastructure: Dynamics, Tensions, and Design*, p. 32)

I am inclined to think the NINES experiment, because it focuses on the whole social political shift, is far more likely to succeed than technology projects like Bamboo and have argued the importance of the social to the Bamboo folk.

3. My third point is that infrastructure programs create an expectation of permanent funding. After all that's what infrastructure is supposed to be – that layer of technology, service and organization which is so indisputably useful to all that we sacrifice independence for efficient and ongoing support. The problem is that everyone wants their pet projects to be funded forever. What project wouldn't want to be turned into national infrastructure? To be fair to CFI, what they have really been funding is not national infrastructure, but project infrastructure. The National Platform fund is the exception. But they are now finding that organizations funded like Compute Canada are arguing for renewed funding. I suspect CFI is going to have to make a decision about how they handle such expectations.

4. The previous points notwithstanding, I think the experience of TAPoR is that we do now know what basic research infrastructure should look like in the humanities. I think it fair to say that a university that is serious about digital humanities research should be able to make the following available,
    a. Access to lab with specialized workstations, digitizing equipment and software. Labs with lots of computers will be underutilized (unless you use them for training) as most of us have our own laptop; what is needed is the specialized stations like video editing, book scanning and so on.
    b. Access to digitization facilities to able to acquire evidence for research.
    c. Access to support that can quickly set up basic off-the-shelf web research tools from distribution lists, blogs to wikis.
    d. Access to a server or virtual machine where projects can install the tools they need for specialized projects. Providing humanists with a locked-down CMS which you can only use to publish static pages does not allow us to use the wealth of open source tools and languages out there to create innovative research environments. Neither should security or standardizaiton rule any longer. Humanists should be able to get a Virtual Machine set up with sufficient storage for any project that has the programming support needed.
    e. Finally, and most importantly, access to good advising and technical support so as to be able to develop projects, apply for funding, and get project management support without being a humanities computing expert.

    I could go into greater detail, but most of you at larger institutions have something approaching this basic level. Something we should think about is whether SDH/SEMI should articulate something basic level of expectation so as to encourage universities to meet an achievable standard.
    Ironically, many of the resources are becoming inexpensively available from commercial vendors. If you know what you are doing you can get well

managed ISP accounts for under $10 a month. With Amazon and others offering cloud computing at rates that universities can't beat, the time may come when all we need at the university is the people to explain to us how to use the negotiated commercial services.

5. My final point has to do with alliances. The time may come where we don't need research computing infrastructure exclusive to the humanities just as we don't need libraries only for the humanities. We need to start reaching out to other constituencies that are either better organized or constituencies that, like us, are trying to define their needs. One national organization comes to mind that we should be talking to, and that is Compute Canada. They have, until recently had a narrow view of the types of support they should provide, but that has been changing as they reach out to us and vice versa. Perhaps the time has come to engage them in a conversation about providing national basic research computing support to the humanities.