GENI Opt-In First Workshop

This section contains 15 white papers submitted by the workshop's participants.

Andrew Odlyzko, School of Mathematics, University of Minnesota
Yan Chen, School of Information, University of Michigan
Joseph Konstan, Department of Computer Science and Engineering, University of Minnesota
Michael Schwartz, Yahoo
Donna Hoffman, Chancellor's Chair and Professor of Marketing at the A. Gary Anderson Graduate School of Management, University of California – Riverside
Thomas P. Novak, Albert O. Steffey Professor of Marketing at the A. Gary Anderson Graduate School of Management, University of California – Riverside
Bill Lehr, Computer Science and Artificial Intelligence Laboratory, MIT
David Waterman, Department of Telecommunication, Indiana University
Barbara van Schewick, Stanford Law School, Stanford University
Jeff MacKie‐Mason, School of Information, University of Michigan
Christopher Yoo, Penn Law, University of Pennsylvania Law School
David Parkes, School of Engineering and Applied Science, Harvard University
Alessandro Acquisti, Information Technology and Public Policy, Carnegie Mellon University
Rahul Telang, Information Systems, The Heinz School, Carnegie Mellon University

Thoughts on network evolution

Andrew Odlyzko

School of Mathematics
University of Minnesota

odlyzko@umn.edu
http://www.dtc.umn.edu/~odlyzko

July 7, 2008

Abstract: This brief note suggests three key factors to keep in mind when planning new network architectures:

(1) Need for flexibility, since predicting technology developments and users needs is notoriously hard

(2) Continuation of a heterogeneous collection of networks

(3) Key role of application developers

(4) Possibility of inflection points, where the relative values of different technologies change suddenly

1. Need for flexibility

First, a few cautionary words about any kind of economic or technological predictions. It is rare for them to be accurate. This is true in general, and in networking in particular.

As just one example,
consider the influential book [Tanenbaum1989]. It was published in 1989, at a time when there were many incompatible data networks. This had been widely seen as undesirable, and a cooperative effort resulted in the creation of the OSI Reference Model. And Tanenbaum, representing the consensus of the time, wrote (p. xiii of [Tanenbaum1989]) that "[i]n the near future, almost all other network architectures will disappear."
The Internet was around then, it preceded the OSI Reference Model, and it is described in Tanenbaum's book. Yet there is no hint that this architecture was going to dominate within a few years. And unfortunately there does not seem to be any careful study of what made the Internet dominant.
The mistake in ignoring the Internet's potential does not seem to have been an obvious one. The logic behind the OSI model seemed impeccable, yet it failed.

The reasons for the success of the Internet do not seem to have been explored carefully. After all, it was quite an old technology, struggling with many competitors (X.25, ATM, SNA, DECNet, AppleTalk, ...). And its start was not all that auspicious, in that it was designed with many misconceptions in mind. (For example, email, the original, and many would say still dominant "killer app," was specifically excluded from the original design criteria for the Arpanet. And the designers did not anticipate the importance of local traffic, which was referred to as "incestuous traffic," and not reported to ARPA out of embarrassment.) So why did the Internet emerge as the dominant networking technology? That question has not attracted the detailed attention that it deserves. Perhaps the key factor seems to have been its flexibility, and its ease of use by application developers, what David Isenberg [Isenberg1997] has called being a "stupid network." (There were certainly other factors as well, one of them the hiding of the true costs of operations [Odlyzko1998].) At AT&T, where top management was declaring ATM to be the future well into 1998, internal groups were adopting IP for their applications en masse by the early 1990s. It was certainly not the efficiency of the Internet that made it the preferred choice. But its ease of use (even when that was achieved by pushing complexity and costs onto others) and ability to interconnect a variety of systems collected a devoted cadre of users (and, even more important, tool developers) at universities and research institutes, who were making it more and more attractive (in a form of the disruptive innovation that Christensen has popularized, where a new technology starts out serving a niche market). So by the time the need for general data connectivity was becoming really pressing in the late 1980s and early 1990s, the Internet was much more functional than initially, and could be used for many of the purposes that the OSI Reference Model was aimed at.
But the success of the Internet does not seem like something that should have been obvious two decades ago.

On the other hand, technological and economic forecasting is also full of examples of mistakes that are obvious, but are almost universally overlooked. Tanenbaum's book can be cited again. He wrote that a "computer network as a sophisticated communication system may reduce the amount of traveling done, thus saving energy" (p. 6 of [Tanenbaum1989]).
The assumption that transportation and communication are substitutes for each other is a couple of centuries old (and is cited prominently today in connection with proposals to ameliorate global warming). Yet it has a couple of centuries of evidence that uniformly refutes it; while the precise dynamics are complicated, and not completely understood, transportation and communication tend to grow in parallel, and one stimulates the other. And surely if one starts with obviously false assumptions, one's chances of coming up with correct models and predictions are not likely to increase. Yet such false assumptions are quite common in telecom. Often they are much more destructive than the example of transportation vs. communication cited above, which has limited practical impact. Perhaps the most damaging of the false dogmas is the "content is king" thesis that has led to a series of misguided decisions over the centuries [Odlyzko2001]. Yet much of the commercial network development is aimed at architectures designed for delivering content. (Yet another example of how even networking researchers have

difficulty facing reality is shown by Fig. 1 in [Kilkki2005], which shows huge preponderance of papers on ATM compared to those on Ethernet throughout the 1990s.) So the conclusion is that the chances that any network architecture, especially one chosen by a committee, is likely to succeed more by luck than by design. Hence the need for maximal flexibility.

2. Continuation of a heterogeneous collection of networks

In spite of the success of the Internet, we do not have a single network, and we should not expect to have one, ever. To make this point, recall the famous phrase "Never underestimate the bandwidth of a station wagon full of tapes hurtling down the highway" (p. 57 of [Tanenbaum1989]).
Similar comments are said to have been made already in the 1970s, and continue to be made today. Since transmission and storage technologies are progressing at roughly the same pace, we should expect that physical transport will continue to be the preferred method of moving large volumes of data when latency is not a major concern [Odlyzko2003].

And this is just one example. We have a large variety of networking technologies that coexist with the Internet, in various relationships, and at various levels of the networking hierarchy. In particular, in wireless we have a veritable zoo, with iDEN, GSM, CDMA, GPRS, WCDMA, HSDPA, HSUPA, 1xEV-DO, ... And let's not forget that the Internet achieved its role largely because it could ride on top of the voice network, so can be regarded as an overlay of that network. And the Internet at various times in its life depended strongly on being an overlay of the ATM network (in the core in the late 1990s, and even today in much of the access network, where DSL still often uses ATM).
But now voice is often an overlay on the Internet, carried as VoIP (whether end users knowingly use VoIP services, or whether their service providers simply use VoIP inside their networks). And ATM traffic is often carried over the Internet (over MPLS).

So where does that leave us? A multiplicity of networking technologies, interwoven in complicated ways, and all constantly evolving. What makes the Internet (taken to mean IPv4 in most cases) the most talked-about technology is that it provides an interoperability level.
Application developers can, if they don't have to worry about major performance limitations, treat the IP level as the "stupid network" of [Isenberg1997], and just toss bits at it. In situations such as wireless, where previous transmission technologies did face major performance limitations, there was a necessity for vertical integration (and in many wireless communication niches, there will continue to be a necessity for such vertical integration). But now, with evolution towards 3G and 4G, there is more transmission capacity, and so system designers can afford to be less efficient in use of that capacity, and more efficient with their time, can use the modular approach with IP as a key layer.

3. Key role of application developers

The main point about networks is that users don't care about the technology. What users care about is what they can do with applications, which in turn use networks. A good writeup of the issues is the paper of Kalevi Kilkki [Kilkki2008]. He considers something even a bit more complicated, namely where there is a user, who actually takes advantage of an application, and what he calls a customer (perhaps not the best name), who pays for that application. Separating these two roles is perceptive, often they are separate. For a fuller picture, it might even be appropriate to separate out the roles of network operators and their suppliers, since there is often a complicated dynamic there (especially in the wireless area today).
But for our purposes I think we can merge network operators and their suppliers on one hand, and users and customers on the other.
Sometimes it might be appropriate to aggregate users and application providers, as a large heterogeneous body that reacts to different networking technologies that are offered.
What is very likely to be very misleading is separating users into content providers and residential users.

So we have three main parties:

* users

* application providers

* network operators

And of course there are complications. Application providers may provide a standalone package, such as the Firefox browser, or an interactive service, like Google, or a package or service that provides access to some other service (think iTunes). And some users are big ones (like CNN, Google, Amazon, or Microsoft, the prototypical "content providers" who dominate mainstream thinking about broadband networks), who interact with many others, and generate large traffic, while most are at the "long tail," usually with little traffic and little value individually, but who provide most of the value in aggregate.

The division above suggests that one should look not at two-sided markets, but at three-sided ones. And some of the most fruitful analogies might be with the computer industry, at least over the last two decades or so, after the PC dethroned the mainframe.
(Note that the early days of the computer industry, when the mainframe makers provided the hardware, the operating system, and the application software, can be likened to the case of the "walled garden" approach that is the antithesis of net neutrality.)
The analogy might compare the operating system provider to the network operators (if we ignore the dynamic between such operating system companies and their hardware suppliers). And the computer industry also displays some of the complications we see in telecom, as in the Windows operating system being controlled by the same company that provides a key application, Office (which is also provided for another, competing operating system, that of Apple).
But let's try to think a bit abstractly and ignore this.

The main point is that application developers are key to the success of any network architecture. And at the moment they are still moving towards the current Internet (meaning IPv4 Internet). There is no incentive for them to develop applications for architectures that don't exist yet, or that are not widely used (as with the IPv4 to IPv6 transition [Perset2008]). So the IPv4 Internet is still gaining momentum, in spite of its faults. Even commercial efforts, such as IMS and NGN, are not seeing much success, although to a large extent they could be implemented invisibly to the users (and often application developers).

4. Possibility of inflection points, where the relative values of different technologies change suddenly

What could speed up a transition of IPv4 to IPv6, or the emergence of an entirely new architecture, such as GENI or NGN? Well, a surprise (similar to the rise of the Internet). Some niche application requires special features, somebody clever builds a networking technology that satisfies those needs, and it happily turns out that this technology can be used more widely, and initially can run on top of the Internet, and later also carry the Internet. And then some application developers might build things that work much better with that technology than with the Internet. And so on. But so far there are no obvious candidates for such a role. (But then there were none two decades ago, either.)

In addition to surprises, there two other factors that might accelerate spread of a new technology (such as replacement of IPv4 by IPv6).

(i) the Internet is getting "encrusted with barnacles." All that baling wire and chewing gum that hold it together are getting harder and harder to deal with, see Geoff Huston's recent column [Huston2008]. This is a gradual process, but it is proceeding. And with it, the Internet is becoming less of the "stupid network" that one just tosses bits at, and more a creaky thing one has to work very hard to make do what is necessary.

(ii) the exhaustion of IPv4 address blocks appears (and unfortunately there is no unanimity on this point) about to make the growth of the Internet much more painful. This is the inflection point referred to at the beginning, where suddenly the tradeoffs become different, with the costs of using IPv4 rising, and so a new alternative getting a much higher chance of acceptance. Some disasters are predictable in their effects, but not time, for example Hurricane Katrina, or major earthquakes in California or Tokyo. On the other hand, the IPv4 address problem seems to be pretty well predictable in its timing, but the scale of the pain it will inflict is not certain. But it is possible that it will be very painful.

[Huston2008] G. Huston, The end of end to end?, May 2008 column, available at http://www.potaroo.net/ispcol/2008-05/eoe2e.pdf

[Isenberg1997] D. Isenberg, Rise of the stupid network: Why the Intelligent Network was once a good idea, but isn't anymore. One telephone company nerd's odd perspective on the changing value proposition, Computer Telephony, August 1997, pp. 16 - 26. Available at http://www.isen.com/stupid.html.

[Kilkki2005] K. Kilkki, Sensible design principles for new networks and services, First Monday, vol. 10, no. 1, Jan. 2005. Available at http://www.firstmonday.org/issues/issue10_1/kilkki/.

[Kilkki2008] K. Kilkki, Quality of Experience in communications ecosystem, J. Universal Computer Science, vol. 14, no. 4, 2008, pp. 615-624. http://www.jucs.org/jucs_14_5/quality_of_experience_in/jucs_14_05_0615_0624_kilkki.pdf

[Odlyzko1998] A. M. Odlyzko, Smart and stupid networks: Why the Internet is like Microsoft, ACM netWorker, vol. 2, no. 5, Dec. 1998, pp. 38-46. Available at http://www.dtc.umn.edu/~odlyzko/doc/stupid.networks.pdf

[Odlyzko2001] A. M. Odlyzko, Content is not king, First Monday, vol. 6, no. 2, Feb. 2001, http://firstmonday.org/issues/issue6_2/odlyzko/

[Odlyzko2003] A. M. Odlyzko, The many paradoxes of broadband, First Monday, vol. 8, no. 9, Sept. 2003, http://firstmonday.org/issues/issue8_9/odlyzko/

[Perset2008] K. Perset, Internet address space: Economic considerations in the management of IPv4 and in the deployment of IPv6, OECD Ministerial Background Report DSTI/ICCP(2007)20/FINAL, May 2008, available at http://www.oecd.org/dataoecd/7/1/40605942.pdf

[Tanenbaum1989] A. S. Tanenbaum, "Computer Networks," 2nd ed., Prentice Hall, 1989.

GENI: Strategies for Attracting End Users

Yan Chen
School of Information
The University of Michigan

This paper identifies strategies for GENI to attract users, based on the economic theories of network externalities, and the incentives for the private provision of public goods. In what follows, I will summarize the theories and apply them to the GENI user community. Some of the ideas are elaborated in Chen, Konstan and Resnick (2008) as a chapter in a new handbook for online communities.

Theory of Network Externalities and Its Implications

The theory of network externalities was developed for products for which the utility that a person derives from consumption of the product increases with the number of other users consuming the product (Katz and Shapiro 1985, Farrell and Saloner 1985). A good example is the fax machine. Being the only person in the world who owns a fax machine does not generate much utility. A user's value for a fax machine increases with the number of other people who use the fax machines. For the community of GENI users, I assume that a user’s utility from opting in (weakly) increases with the total number of users on GENI. Suppose, among other things, users on GENI engage in the production of contents or social networking with other users, more users might translate into more contents to consume and more friends to interact with. For researchers, if there are a large number of users, they can conduct randomized experiments to test theories and improve designs. If this assumption is by and large correct, some insights from the theory of network externalities can be readily applied to the design of GENI.

The theory of network externalities, developed for competing systems, highlights at least three important issues which are relevant for GENI: expectations, critical mass, and compatibility (Katz and Shapiro 1994).

When a user makes a decision of whether to join a new online community, in this case GENI, the value she expects to derive from the community depends on the *expected* size of the community. When a community is expected to be popular, and thus have high quality content and interactions with other users, it will be popular for that very reason, a phenomenon called positive feedbacks. As a result, the number and nature of the competitive equilibria depend on users' expectations about the new community. Theory puts a rationality assumption on the expectations, i.e., they correctly incorporate all information available at any point in time. We use an example to illustrate the importance of expectations. Suppose a new online community enters the market. If no one believes that it will succeed in attracting users, no one will join it. As a result, there is a fulfilled expectations equilibrium with no one joining the community. On the other extreme, suppose each potential user believes that a large number of users will join the community. Then many would join, and this outcome is a second fulfilled expectations equilibrium.

Under certain assumptions, there exist at least two fulfilled expectations equilibria. From a designer's perspective, it is extremely important to influence potential users' expectations in order to achieve the "good" equilibrium. Strategies that influence expectations include advertisement, subsidizing early users, and providing incentives for referrals. A community's reputation could also influence which equilibrium is selected.

In communities with network externalities, because of positive feedbacks, competition tends to result in tipping or a winner-take-all effect, i.e., the tendency of one community to get ahead of its rivals once it has gained a *critical mass*. In product markets, tipping has been observed in many situations where one product eventually dominates, such as VHS vs. Beta in videocassette recorders or more recently, Blu-Ray vs. HD for DVDs. In the context of online communities, heterogeneous tastes and community differentiation might limit tipping and sustain multiple networks. A new entrant catering to users who care more about particular community attributes than network size might be able to co-exist with large incumbent communities. As GENI is a new entrant, it is important to be able to achieve critical mass for the community to be viable.

A third strategic decision for the designer is *compatibility*. An implication from theory is that new communities, such as GENI, should choose compatible technologies with large incumbents. In general, incompatibility discourages entry as it requires a new entrant to have a minimum size to be viable. For example, in the space of user-customized music sites, Last.fm was the first to offer free, on-demand, full-length music that a user might choose. When Pandora entered the market, it emulated the last.fm technology of a single box for entering artists or songs to start a new personal radio station, while offering an arguably superior recommender system. The decision to adopt a compatible technology, or even the interface (the look and feel of the site), lowers the switching cost of the users of the incumbent sites. Meanwhile, it might increase the perceived size of the network. A third advantage of compatibility for the entrant is that third party applications designed for the incumbent sites can be integrated into the new online community. For example, tianji.com, a social networking site in China, adopts the OpenSocial API standard supported by Google, MySpace, LindedIn and other major incumbents in social networking sites, which facilitates the integration of third party applications designed for the incumbent sites.

The Theories of Private Provision of Public Goods and Its Implications

GENI, as an infrastructure, can be viewed as an excludable public good. Opting into the network can also be modeled as adding to the public good, which does not preclude the consumption of private and public goods provided on GENI. Therefore, some theories of the private provision of public goods might be relevant. In what follows, I will focus on the signaling theory of Andreoni (2006).

As GENI is relatively unknown to the larger Internet community, joining GENI shares similar characteristics as contribution to a public good of unknown quality. In the domain of charitable fund-raising, when the public good is of unknown quality, a leader giving a large amount conveys a credible signal of the quality. A fundraising consultant is quoted in the New York Times as saying “When a big (leadership giver) comes in, the smaller donors pay attention. It legitimizes a fund-raising project and puts the institution on a much faster track.” Game theoretic analysis of leadership giving finds that having a leader increases the total amount of public goods provided in equilibrium.

In the GENI domain, leadership giving theory implies that the designer should proactively recruit well-known researchers and institutions to be early users. Published research conducted on GENI is likely to signal the quality and the potential of the network, and to attract new users.

References

Andreoni, James (2006). Leadership Giving in Charitable Fund-Raising. Journal of public economic theory. Vol.8, Issue 1, pp. 1 - 22.

Chen, Yan, Joseph Konstan and Paul Resnick (2008). Starting an Online Community. Manuscript.

Farrell, Joseph and Garth Saloner (1985). Standardization, Compatibility, and Innovation. *The RAND Journal of Economics*, Vol. 16, No. 1, pp. 70-83.

Katz Michael L. and Carl Shapiro (1985). Network Externalities, Competition, and Compatibility. *The American Economic Review*, Vol. 75, No. 3, pp. 424-440.

Katz Michael L. and Carl Shapiro (1994). Systems Competition and Network Effects. *The Journal of Economic Perspectives*, Vol. 8, No. 2, pp. 93-115.

Thoughts on Attracting GENI Users

Joseph A. Konstan
Department of Computer Science and Engineering
University of Minnesota
konstan@cs.umn.edu

Background

My background is in human-computer interaction, with five years of recent experience focusing on the specific problem of designing online community sites to elicit participation, and more than a decade of work on public health computing systems designed to influence individual behavior. As such, I come from a perspective that bridges computer science, social science, and design in an attempt to design computational systems that effect behavioral change in their users.

- social science as the theoretical foundation from which to draw understanding of human behavior and the models from which we can build systems to influence that behavior; psychology and economics are particularly useful foundations here

- computer science as the vehicle for employing computation to implement these models of influencing behavior; such computation may involve algorithms for identifying and responding to user state, algorithms for identifying the correct users to approach with specific work proposals, or computationally implementing incentive systems or community structures

- design as a discipline for tying together the scientific knowledge and engineering tools and crafting solutions for specific situations

My focus, therefore, is on highlighting some of the relevant work from the HCI/Social Informatics perspective to help identify ways in which such research can be of use in motivating users and groups of users to opt into the use of GENI.

Why do People Change their Behavior?

A substantial amount of research has focused on the broad question of how to elicit behavior change. This focus is not surprising, as eliciting behavior change is central to many fields, including psychology, public health, public good economics, and marketing. To oversimplify things, I group the theories in this space into three categories: economic theories, individual-based psychological theories, and social (group) psychological theories. Without being an expert in any of these, I have found them useful for problems related to eliciting behavior change.

Economic Theories. The core of microeconomics argues that we can model decision-making by understanding the *utility function* of the parties in a market. An individual will take an action (e.g., work, sell something, or buy something) when the result of that action leads to an increase in her utility. Or to use different terms, she will take an action when the benefit of the action exceeds the cost of the action.

Such core economic theories have been useful in understanding the participation in online community. In [1], we developed and calibrated an economic model of user rating in an online recommender system. This model, constructed from a combination of behavioral data and survey data, helped us identify the extent to which each individual user valued different aspects of system use (getting better predictions, self-expression, influencing others) as well as how users assessed the costs associated with adding ratings to the system (how much time and effort they spent). After calibrating the model, we found it to be well-correlated with actual user rating activity. In later work, we've looked at customizing interaction with users based on their individual utility functions.

This basic approach is clearly important when understanding how to get new users to adopt GENI as a platform for their work. Consider the case of a researcher "courted" to use GENI for a new gene expression analysis project. What does he stand to gain from the use of GENI? To what extent will this infrastructure increase his chances of success (or of speedier success that outpaces his competitors)? To what extent does it reduce his costs (and how much does he care)? And what does he lose? How much time and effort must be invested to learn to use this infrastructure? To participate in evaluations and experiments unrelated to his core work? What is the chance of failure? And what is the risk profile of the researcher?

Based on answers to these questions, it becomes possible to explore the design of services and support to make opting-in to GENI an easy decision, or perhaps to identify the right targets to pursue. When the system is still less reliable than alternatives, the key may be to identify people for whom there is sufficient motivation to outweigh the risk (either those who see this as a likely path towards getting or staying ahead, or those who are not overly risk-averse to begin with). Other design options may include tailoring the level of commitment to evaluation, workshops, etc., to the individual, or offering support tailored to individual needs. In essence, GENI can become, for some key users, a customized product offering.

Individual-Based Psychological Theories. Several different theories bear on the decision-making of individuals. Behavioral economics [2] has grown out of a set of observations that humans do not act as classical rational agents as envisioned in basic economic theories. A collection of experiments has shown a number of "anomalies" such as the endowment effect (i.e., that people value a good more once it is theirs) and inequality aversion (i.e., the phenomenon that people will forego their own benefits to prevent someone else from benefitting excessively from a transaction). Behavior change theory, on the other hand, looks more directly at the ways in which a behavior change moves from being an idea to a commitment to an action to a habit. Perhaps the most widely-referred to example is Prochaska and DiClemente's's transtheoretical model of behavior change [3] which models six stages of behavior change (from precontemplation, where there is no awareness of the possibility of life-improving change; through contemplation and preparation, where the individual considers and commits to change;

through action, where the change is implemented; through maintenance and then termination, where the change is fully incorporated into regular behavior). Other similar models exist.

The insights from these individual-based theories can be of service when designing mechanisms to elicit opt-in. Understanding behavioral economics can help in recognizing the challenges where human perceptions do not match simple objective models. Users will likely be less willing to give up GENI than to adopt it, suggesting a need to get people in the door, after which they are more likely to stay (other theoretical sources such as cognitive dissonance theory [4] suggest that once people have invested heavily in something, such as the large fee for joining a country club, they in turn convince themselves that it must be worth it; one might speculate if that is why we computer scientists are so attached to our often-arcane text editors and word processing software). Theories of behavior change can help identify the points at which the adoption of GENI is most risky, and the points at which is has been solidified, and can help tailor the approach to users based on their progress in adoption.

It is also important to note that computerized systems can be designed to identify and respond to individuals based on these theoretical models. In [5], we demonstrated how a telephone-based interview system could use gathered behavioral information and the transtheoretical model to customize exercise prompting for older women. Such a system recognizes the difference between someone who didn't exercise in a given week as part of a pattern of not exercising, someone whose failure to exercise seems to be a risk of "falling back" on old ways before new habits are solidified, and someone for whom this is an anomaly within a regular pattern of exercise.

Social Psychological Theories. Much of the decision on whether to take an action depends not only on the individual targeted, but on that individual's relationship to others. Theories such as Karau and Williams' collective effort model [6] explain individual contribution as a combination of several factors: the degree to which the individual desires the outcome, the degree to which the group desires the outcome, the affinity of the individual for the group, and the individual's estimation of the likelihood of success both with and without her individual effort. Through such a theoretical lens, it becomes clear that users can be demotivated both by the belief that the project will fail, with or without their effort, and by the belief that the project will succeed just fine without them. This latter phenomenon is also known as "uniqueness" and is a reason for assembling groups without too much duplication in ability, so that each group member feels necessary to the success of the project.

As an example of such work, [7] reports on a series of studies that show how manipulating an appeal to individuals to participate (in discussion or rating behavior) can affect the success of the appeal. Appealing to uniqueness, and setting ambitious but reachable goals both have a consistent positive effect. Interestingly, a focus on the self-benefit or the benefit to others had no effect or negative effects (this has been explained by some psychologists as a case of the articulated tangible benefit crowding out the unarticulated intangible benefits). In a separate study, we found that social comparisons (pointing out to individuals how their efforts compared to those of others in the community) could affect both the quantity and nature of work done by those receiving the comparisons [8].

Such social psychological theories have similar value in attempts to elicit users for GENI. It is important to understand whether the prospective user cares about the success of GENI (or some subproject) at all, and if so, to make a case that their effort can lead to greater success. Linking the individual to a community (whether through a common identity or through person-to-person bonds) is likely to increase affinity and therefore interest in seeing success. Yes, as obvious as it may sound, there can be value in give-aways of swag (logo items and/or apparel) as a means of inducing identity and therefore affinity.

Interesting Examples of Research in this Space

The research cited above is primarily theory and very direct applications to test theory in the domain of online communities. There are interesting examples of research that is more problem-driven, and that illustrates the type of creative thinking that can be valuable in eliciting opt-in usage.

Intelligent Task Routing. Cosley et al. [9] addressed the problem of assigning work to volunteers in Wikipedia. There are lots of volunteers. There is lots of work (even more work than available volunteers). The question was how to get people to do more work, and do it better. The insight was that people were more likely to do work that interested them, and that people who didn't find interesting work might not do any at all.

Cosley et al. developed SuggestBot, an agent that scanned all the available work and suggested work to individuals (through their talk pages). They found that intelligently chosen selections (based on a variety of algorithms including text matching and link analysis) yielded nearly four times as many edits as randomly chosen ones (the default for Wikipedia was alphabetical, which is quite close to random given the enormous number of edits). This

Network Value of Customers. Domingos and Richardson [10] asked the question of how to identify the right individuals to market to when the goal is to grow a product's market share through word of mouth. Intuitively, the idea is to identify individuals who have two properties: (1) they are likely to like the product, and (2) their preference is likely to influence many others to try (and perhaps like) the product. In the marketing world, there are companies that attempt to build panels of such "taste-makers" and distribute product samples to them (sometimes these are celebrities; often they are just well-connected in a target group).

In the online world, Domingos and Richardson found that it was possible to identify people of influence in a recommender community (based on a matrix of ratings data). This is an algorithmic influence that helps determine how strongly how many other people will receive recommendations based on the positive rating of a "high value" customer.

The same idea extends more broadly. Organizationally, there are generally human hubs who disseminate information to others, and opinion leaders who influence others to try new technologies. These form an opportunity for marketing.

Conclusion

There are many examples of effective use of social science theory in carefully designed computational systems that result in greater participation. Experiences in domains outside of GENI's will be instructive in designing a plan for recruitment of opt-in users.

References

1. An Economic Model of User Rating in an Online Recommender System, Harper, F.M.; Li, X.; Chen, Y.; Konstan, J.A. , *Proc. 10th International Conference on User Modeling*, Edinburgh, UK, (2005)
2. Kahneman, D. & Tversky, A. 'Prospect Theory: An Analysis of Decision under Risk,' *Econometrica*, XVLII (1979), 263–291.
3. Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*, 390-395.
4. Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
5. Peters, J.A., **Wyman, J.F.**, Konstan, J.A., Srinivasan, S., Bowman, A.M., & Brandt, C.L. (2000). Development and testing of a computerized system for prompting and monitoring exercise behavior in older adults. *The Gerontologist*, *40* (Special Issue), 123.
6. Karau S.J. and Williams, K.D. (1993) Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology 65*(4) 681-706.
7. Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Cosley, D., Frankowski, D., Terveen, L., Rashid, A. M., Resnick, P., and Kraut, R. (2005). Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication* 10(4) article 10.
8. Harper, F.; Li, X.; Chen, Y.; Konstan, J. (2007) Social Comparisons to Motivate Contributions to an Online Community. *Persuasive Technology Conference*, Palo Alto, CA, (2007).
9. Cosley, D.; Frankowski, D.; Terveen, L.; Riedl, J. (2007) SuggestBot: Using Intelligent Task Routing to Help People Find Work in Wikipedia, , *ACM International Conference on Intelligent User Interfaces* (IUI 2007), Honolulu, Hawaii, USA, p.32 - 41, (2007)
10. Domingos, P. and Richardson, M. (2001). Mining the network value of customers. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 57-66, San Francisco, CA, 2001. ACM Press.

This is a very raw draft/outline of a white paper for G-WEB 1 conference. At this time I am not certain if I can participate, I will submit a complete paper if I will be able to come to the conference.

Michael Schwarz

End User Experience

When an end user opts into GENI he or she agrees to participate in experiments thus effectively becoming a gini pig for testing unproven technology. GENI is not a product, rather it is a platform for experimenting with different technologies. Consequently success of GENI is not synonymous with good user experience for users who opt in. GENI will be a success if at least some technologies developed using the platform are successful; however, research is a process of trial and error, it is not unlikely that most technologies tested on GENI will be "errors" as long as the users are concerned. It is important that users who opt into GENI are not overexposed to poorly performing technology. One way to insure that users who opt into GENI receive a reasonable experience is to condition duration of experiments on the experience delivered to end users e.g. as soon as it becomes apparent that an experiment leads to poor network performance the experiment should be aborted.

There are many methods for measuring network performance, a number of parameters such as latency experienced by users, system downtime, crashes and direct user feedback[1] can be combined into a single variable that can serve as proxies for user experience.
For instance, a researcher who conducts an experiment may be allocated a quota of "bad experience" when the quota is reached the experiment is stopped.

User incentives for participation

From an individual user point of view most of the time participating end users may receive inferior performance. Thus end users may have to be compensated for participation in GENI. The traditional method of attracting experimental subject is paying the subjects either with money or with free or discounted services (for instance, patients in clinical trials often receive free care in exchange for agreement to participate). It seems that the model used to attract subjects to clinical trials may be appropriate for GENI. The cheapest method of attracting users to GENI is by offering to users a deeply discounted service that has a low marginal cost relative to the average market price. For instance, in

[1] It is worth noting that direct reports from users about their experience with the network may be equally valuable. Indeed, in the early days of networking the standard performance measure was percentage of time that the system was available. When users were informed that the system that they were using was up 99% of the time they felt that there was something wrong with that measure—in particular, the system was always down during the pick demand periods and always up when nobody was using it (which was most of the time). Of course, today's measure of system performance are far more sophisticated; however, it is worth remembering that no measure of network performance is perfect and that direct user feedback helps to make sure that technical measures of network performance are consistent with actual user experience.

many areas there is excess bandwidth. The market price of bandwidth may be fairly high, but the marginal cost is low (when the system is not at capacity), if some of the institutions participating in GENI are willing to make excess bandwidth available for the platform the users may be induced to participate in GENI because although the service may be less stable the participants may be compensated for lack of stability by faster than normal access to content that require high bandwidth. The other low marginal cost recourse is content itself. Many forms of content are either not available or expensive (e.g. most old TV and radio broadcasts are not available; many movies are available for a high fee). One can imagine a copyright holder of this sort of digital content may be willing to make it available to GENI users for free or at a low cost. Availability of free content to GENI users may be an inducement to participate (if that content unavailable or available for a fee to users who are not opted into GENI).

Availability of experimental data

GENI is somewhat akin to a supercollider or Hubble Space Telescope. Clearly, GENI is meant to be a recourse shared by many researchers both in industry and academia. It is less clear how the results of that research should be shared. For instance, if a particular researchers runs an experiment on GENI should the design of the experiment be public record? Or perhaps should the design of the experiment become public record after a period of time may be weeks or months after an experiment is conducted? Then there is a question of data. An experiment generates data, should that data become public record, should it become public record with some delay? One can make a good case in favor of requiring that non commercial researchers applications both the experimental design and the data should become available shortly after experiment (perhaps with enough delay to allow a researcher to publish his finding). I believe there are good reasons to allow commercial users of GENI to keep experimental results and designs confidential; however, the privilege of keeping experimental results confidential should be reserved for organizations that contributed resources to built GENI.

“Emergent Consumers” Can Help Develop Successful Future Ideas

Donna L. Hoffman

Thomas P. Novak

July 1, 2008

Donna L Hoffman is the Chancellor’s Chair and Professor of Marketing at the A. Gary Anderson Graduate School of Management, 900 University Avenue, University of California, Riverside, Riverside, CA 92521; Ph 951-827-4848; Fax 951-827-3970; Donna.hoffman@ucr.edu

Thomas P. Novak is the Albert O. Steffey Professor of Marketing at the A. Gary Anderson Graduate School of Management, 900 University Avenue, University of California, Riverside, Riverside, CA 92521; Ph 951-827-4999; Fax 951-827-3970; tom.novak@ucr.edu

White Paper Prepared for the GENI Opt-In Workshop, July 20-21, 2008

"Emergent Consumers" Can Help Develop Successful Future Ideas

ABSTRACT

Though marketers are well aware that typical consumers have difficulty estimating the usefulness of new products, little research has focused on *which* consumers to use in the new product development process, particularly in the consumer goods industry. We have developed a methodology to identify "emergent consumers," individuals that are able to synergistically apply intuition and judgment to improve product concepts that typical consumers will find more appealing and useful relative to ones that are developed by lead user, innovative, or "ordinary" consumers. Emergent consumers can aid in the successful development of new product ideas and concepts, help predict their acceptance by mainstream consumers and improve the chances of success in the marketplace for products eventually commercialized. As such, we believe that employing consumers high in emergent nature may prove useful in the GENI Web Opt-in Project.

“Emergent Consumers” Can Help Recognize Successful Future Ideas

Which consumers are the most appropriate to engage in the product development process is important because research has shown that typical consumers have a difficult time estimating the usefulness of new products. However, little research has focused on *which* consumers to use in the product development process, particularly in the end-user consumer arena. In the business-to-business product development literature, research has identified the lead-user, customers with an early awareness of their needs from the “leading edge” of a target segment, as an appropriate segment to improve idea generation (Lilien, et.al. 2002; von Hippel 1986), and in the consumer behavior literature, Steenkamp, ter Hofstede and Wedel (1999) identified the consumer predisposition to try new products with subsequent research linking this “dispositional innovativeness” to the likelihood to try new consumer packaged goods. However, the use of particular, as opposed to mainstream consumer target segments for product concept testing has not received nearly as much attention.

In our research, we have developed and validated a scale to identify these “right” consumers. We argue that the right consumers possess what we call an “emergent nature,” i.e., the ability to process information in a synergistically experiential and rational thinking style, and exhibit a unique set of personality traits such that they are able to recognize which product concepts mainstream consumers will find more appealing and useful, compared to mainstream, lead user, or even innovative consumers. In other words, emergent consumers are uniquely able to apply intuition *and* judgment to distinguish the best product concepts. We propose that identifying and using such “emergent consumers” for important product development tasks such as product concept testing can help predict the ultimate acceptance of products by mainstream customers and thereby improve the entire product development process.

One important implication of our research is that employing consumers high on emergent nature in the concept development process may act as an early warning system in those cases where new product concepts have the potential to be disruptive (Chandy and Tellis 2000).

In this white paper we summarize the theory on which the emergent nature construct is based and review the results of several studies designed to test our theory.

THEORY

A considerable body of research in dual-processing theory has differentiated among two types of information processing styles: experiential thinking style and rational thinking style (e.g. Epstein 1994) and substantiated the existence of individual differences in these two thinking styles (Norris and Epstein 2003a, 2003b). Rational thinking style involves goal-directed, active, logical processing, and permits consumers to make optimal judgments about the utility of adopting a particular product innovation. Experiential thinking style, on the other hand, involves holistic, emotional, associative processing. Immediate experience is critical for experiential thinking, while logic and evidence are critical for rational thinking. Recent research has also demonstrated that higher levels of experiential thinking style are associated with significantly higher levels of creativity (Norris and Epstein 2003b). Thus, a rational thinking style appears to be adaptive for good judgment in specific decision-making situations, while an experiential style is adaptive for interactions and creative pursuits.

Synergy Between Experiential and Rational Thinking Styles Creates Emergent Nature

Novak and Hoffman (2007) suggested that some tasks "might demonstrate *synergistic* effects" in which *both* experiential and rational situation-specific thinking style might correlate positively with performance. We propose that consumers with an emergent nature are high in both experiential and rational thinking style and are able to use the two thinking styles in a synergistic manner. That is, we argue that *emergent nature is defined largely by the complementary interaction between the experiential and rational thinking styles.*

How does emergent nature influence these consumers' ability to develop concepts? We propose that emergent consumers, owing to the synergies between their thinking styles, are able to engage successfully in both idea generation to enhance the original concept and logical analysis to refine and develop the concept further. In other words, consumers possessed of an emergent nature are able to synergistically apply intuition and judgment to improve product concepts. Consumers with a high emergent nature develop an intuitive, almost "instinctive" understanding of the product by, for example, visualizing the latent uses of the product concept, through a sequence of small scale, affective, and associative perceptions. They are able to generate these visualizations because they possess a high degree of experiential processing ability. The experiential processing system generates the "gut feelings" underlying the intuitive understanding of the potential appeal of a product concept. Following this automatic, associative stage, consumers high on emergent nature, owing to their high degree of rational processing ability, then employ a rational thinking style in a conscious, logical and analytic effort to further evaluate and refine the concept. In our conceptualization, the thinking styles work together in a complementary and iterative fashion, where a rational effort to analyze a product concept may activate further implicit, experiential associations about that concept, followed by another round of rational analysis, and so on. The essence of emergent nature is that consumers so possessed are able to inform their experiential impressions and associations with rational evaluation and judgment and vice versa.

Related Constructs

A related consumer-oriented approach to evaluate the potential success of products is to rely on consumer innovativeness, defined as an underlying predisposition of consumers to *buy* new and different products (Midgley and Dowling 1978), distinct from innovation adopter categories (Rogers 2003), which are determined ex-post product introduction (Midgley and Dowling 1978; Steenkamp, ter Hofstede, and Wedel 1999). Consumer innovativeness has been found to correlate positively with personality traits such as extraversion, risk taking, and impulse buying (Steenkamp et al. 1999).

The use of lead users has also received wide attention, particularly in business-to-business settings (e.g. von Hippel and Katz 2002). Lead users have a conscious awareness of their domain-specific needs, are motivated to innovate to satisfy those needs, and experience those needs earlier than most in the market (Lilien, et.al. 2002; Morrison, Roberts, and von Hippel 2000). The lack of lead user studies in consumer settings suggests that it may be hard to identify lead users in consumer markets and lead user status may not be a trait-based characteristic.

THE EMERGENT NATURE SCALE

Using our theoretical definition of an emerging consumer, we developed and validated a scale of emergent nature. The final set of items is listed in table 1. For comparison purposes, we also created a scale of domain-specific lead user status, also shown in table 1.

TABLE 1

FINAL ITEMS FOR EMERGENT NATURE AND LEAD USER SCALES

Emergent Nature

When I hear about a new product or service idea, it is easy to imagine how it might be developed into an actual product or service.
Even if I don’t see an immediate use for a new product or service, I like to think about how I might use it in the future.
When I see a new product or service idea, it is easy to visualize how it might fit into the life of an average person in the future.
If someone gave me a new product or service idea with no clear application, I could “fill in the blanks” so someone else would know what to do with it.

Even if I don't see an immediate use for a new product or service, I like to imagine how people in general might use it in the future.
I like to experiment with new ideas for how to use products and services.
I like to find patterns in complexity.
I can picture how products and services of today could be improved to make them more appealing to the average person.

Domain-Specific Lead User

Other people consider me as "leading edge" with respect to home delivery of goods.
I have pioneered some new and different ways for home delivery of goods.
I have suggested to stores and delivery services some new and different ways to deliver goods at home.
I have participated in offers by stores to deliver goods to my home in new and different ways.
I have come up with some new and different solutions to meet my needs for the home delivery of goods.

Table 2 displays the correlations and means of emergent nature, dispositional innovativeness and domain-specific lead user status with demographics based on a study of 1124 consumers. There is a slight tendency for consumers high on emergent nature, dispositional innovativeness and lead user status to be younger and better educated (although the correlation between innovativeness and education is not significant). There are very small effects for gender and country. There is a slight tendency for men to be emergent and innovators and for women to be lead users in the home delivery category. Residence in the United Kingdom is positively associated with lead user status and a positive association between being an American or from the U.K. and emergent nature and dispositional innovativeness.

TABLE 2

CORRELATIONS AND MEANS OF EMERGENT NATURE, DISPOSITIONAL INNOVATIVENESS AND DOMAIN-SPECIFIC LEAD USER STATUS WITH DEMOGRAPHICS, N=1124

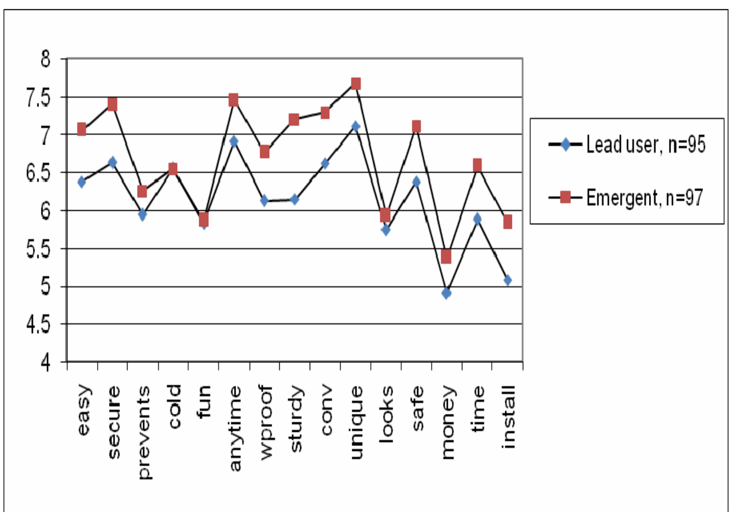
	Correlations		
	Emergent Nature	Dispositional Innovativeness	Domain-Specific Lead User Status
Age	-.076*	-.095**	-.098**
Education	.085**	.010	.073*
Men	37.81 (9.03)	15.83 (7.73)	32.06 (7.5)
Women	36.18 (10.26)	13.70 (7.59)	34.71 (8.68)
Effect size η_p^2 (p-value)	.007 (.006)	.019 (.000)	.025 (.000)
Australia	34.73 (9.30)	12.96 (6.77)	33.17 (7.35)
Canada	35.08 (9.54)	13.33 (7.06)	32.18 (8.04)
United Kingdom	36.77 (9.83)	14.51 (8.59)	36.57 (8.51)
United States	37.33 (9.99)	14.82 (7.80)	33.88 (8.54)
All other countries	39.87(7.45)	18.01 (8.08)	33.01 (6.57)
Effect size η_p^2 (p-value)	.017 (.001)	.021 (.000)	.012 (.010)

**Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level. All tests are two-tailed.

We conducted a concept development study in which five mutually exclusive concept development groups, classified according to their scores on scales for emergent nature, dispositional innovativeness, and lead user status, each further developed a technology concept called the SmartBox for the home delivery of goods (see Appendix for the concept description) in an online bulletin board setting. Participants were instructed to "further develop the SmartBox concept so that it will be successful in the marketplace as a home delivery solution for average consumers" and to "develop the concept so that the SmartBox will be as appealing as possible to the average consumer for the home delivery of goods and that they will want to buy it." Participants developed and refined multiple concepts, and iterated until they had a single SmartBox concept. These five concepts were submitted to a thorough evaluation by a new sample of 631 consumers. The results showed that the concept produced by consumers high on emergent nature was rated the highest and the concept produced by consumers high on lead user status was rated the next highest.

A new independent sample of consumers rated both of these concepts on fifteen attributes (easy to use, secure, prevents breakage, keeps cold foods cold, fun to use, can use anytime, is waterproof, sturdy, convenient, has a unique design, looks good, safe to use, saves money, saves time, and easy to install). The means for the two groups on each attribute are displayed in Figure 1. The concept developed by the high emergent group was rated significantly higher than the one developed by lead users on 11 attributes; there were no significant differences on the remaining four attributes.

FIGURE 1
Attribute ratings of Emergent and Lead User Group Concepts



SUMMARY

We have developed a highly reliable and valid scale to measure emergent nature in consumers and showed that the emergent nature construct is empirically distinct from other product development constructs such as lead user status or dispositional innovativeness. We have also demonstrated that consumers high on emergent nature can develop product concepts that are perceived by typical consumers as significantly better than concepts developed by groups high on domain-specific lead user status or dispositional innovativeness. For these reasons, our conceptualization of emergent nature and the corresponding measurement scale may be useful tools in the GENI Web Opt-in project.

REFERENCES

Chandy, R., G. J. Tellis. 1998. Organizing for Radical Innovation: The Overlooked Role of Willingness to Cannibalize. *Journal of Marketing Research* 35 (November) 474–487.

Epstein, S. (1994), "Integration of the Cognitive and the Psychodynamic Unconscious," *American Psychologist*, 49, 709-724.

Lilien, Gary L., Pamela D. Morrison, Kathleen Searls, Mary Sonnack, and Eric von Hippel (2002), "Performance Assessment of the Lead User Idea-Generation Process for New Product Development," *Management Science*, 48 (8), 1042-1059.

Morrison, Pamela D., John H. Roberts, Eric von Hippel (2000), "Determinants of User Innovation and Innovation Sharing in a Local Market," *Management Science*, 46 (12), 113-1527.

Novak, Thomas P. and Donna L. Hoffman (2007), "The Fit of Thinking Style and Situation: New Measures of Situation-Specific Experiential and Rational Cognition," UCR Sloan Center for Internet Retailing Working Paper, August 27.

Norris, P. and S. Epstein (2003a), "The Investigation of Some Fundamental Issues Concerning Rational-Analytical and Intuitive-Experiential Thinking Styles with a Short Form of the Rational-Experiential Inventory," working paper.

Norris, P. and S. Epstein (2003b), "Objective Correlates of Experiential Processing." Working paper.

Steenkamp, Jan-Benedict, Franekl ter Hofstede and Michel Wedel (1999), "A Cross-National Investigation into the Individual and National Cultural Antecedents of Consumer Innovativness," *Journal of Marketing*, 63 (April), 55-69.

Von Hippel, E. (1986), "Lead Users: A Source of Novel Product Concepts," *Management Science* 32(7), 791–805.

Von Hippel, E., and R. Katz (2002), "Shifting Innovation to Users via Toolkits," *Management Science*, 48 (7), 821-833.

APPENDIX: The SmartBox Product Concept

Now we'd like you to evaluate a new product concept called the "SmartBox." The SmartBox is depicted in the drawing below.

Regardless of whether it's laundry, dry cleaning, groceries or most anything else, the SmartBox should make home pickup and delivery secure and convenient even if no one is home.

Presuming you had a choice of many styles, sizes and installation locations, please imagine that a device similar to this is on, by, or close to your home – or, if you live in an apartment, that a cluster of them is by your building. Suppose that FedEx, UPS and the Postal Service as well as grocers, drycleaners and anyone else you want to authorize could use it to make secure pickups and deliveries. Built-in intelligence enables authorized deliveries only and sends notification to both consumer and merchant whenever a delivery is made.

Thinking about the SmartBox concept, please answer the questions below.

TO: G-WEB 1: GENI Workshop for End-user opt-in Broadening (Cycle 1)

FROM: William Lehr (wlehr@mit.edu, 617-258-0630)
CSAIL (32-G814)
Massachusetts Institute of Technology
32 Vassar Street
Cambridge, MA 02139

RE: Discussion/Position paper for workshop

DATE: June 2, 2008

GENI will provide a national experimental platform for research in next-generation networking and computer science. It is intended to provide a test bed for "at-scale" networking research and industry/academic collaboration that will play a central role in promoting innovation and a healthy networking ecosystem. The resources devoted to GENI will be substantial and so it is worthwhile investing advance effort to help ensure GENI attracts appropriate participation from the academic and commercial communities.

In addition to providing a test bed for pure computer science/networking research, I believe it will be critical for GENI to provide a test bed for multidisciplinary research, engaging collaborative research with social scientists (economists, legal scholars, political scientists, sociologists, and psychologists).[1]

In the balance of this note, I articulate my vision of the computing/networking future and why this implies that multidisciplinary research will be increasingly important; the role that GENI may play and some of the associated challenges; and conclude by identifying several areas that are most likely to offer fruitful opportunities for further development.

Vision of the Networking Future

From an end-user perspective, the future of computing and telecommunication services will be one of ***pervasive computing***:

- ***Ubiquitous, always on, 24/7***: users will expect services to be available every where (at home, at work, on the street; in urban and rural locales; and with roaming capabilities), on all devices (interoperable), at any time (24/7 access), and immediately (near-zero latency).

[1] The focus of GENI is on promoting innovation in networking and computer science research, not in providing a stable, high-performance "tool" or test bed for other types of research. While hard science research may seek to use GENI as an experimental computing platform (e.g., physics, genetics, meteorology, biology), these computationally intensive hard science disciplines already have well-defined requirements for computing/communication test beds that may be inconsistent with the goal of GENI to experiment with networking approaches. Their general focus on performance (speed) and lack of emphasis on manageability, scalability, or security/trust in complex networking environments make the hard sciences less than ideal candidates for participation in GENI. This is not true for hard science research that seeks to exploit new computer science/networking technology such as sensor nets or parallel processing.

- ***Mobile/portable***: users will want their services to be mobile with context-relevant modalities (high-speed and nomadic mobility, flexible moves and installs, portable on person or with car).
- ***Multimedia***: the traffic will be multimedia and bursty, mixing text, data, real-time and streaming video and audio.
- ***Interactive***: services should be responsive to real-time feedback and user interaction. Users will originate services and content as well as consume.
- ***Unaware***: a lot of the computing and communication will take place in the background, without requiring direct end-user intervention or attention.
- ***Personalized***: users will want choice (of suppliers and how provided) and customization of services tailorable to their usage context (location, device, tastes). Communication/computing services become truly mass market when they are personalizable.

Delivering the above experience will require deployment (investment and design) of substantial new infrastructure to enable the next-generation Internet which will be:

- ***Broadband***: capable of supporting multimedia traffic at data rates in excess of 1Mbps per user,[2] with long-tail of higher-speed services that may depend on location (new or legacy, urban or rural) and service context (mobile or fixed). Lower speed services will be available from wireless devices (appliances) while higher speed services will be available from fiber optic fixed line services.
- ***Wireless (and wired, integrated)***: to realize pervasive computing, most of the last-link connections will be wireless (to sensors, to mobile devices). At the same time, higher-speed wired connections will continue to proliferate and many of the new wireless connections will be low-power to base stations with wired trunking. To meet end-user demand for seemless interoperability the many types of wired and wireless access technologies will need to be integrated. This means that provision must be made for physical interconnection as will as application portability across networking platforms.
- ***Context aware and dynamic services***: to accommodate the resulting heterogeneity in demand (with respect to timing, location, and context/application), networks will need to be more dynamically responsive to share and reconfigure resources (share common resources, fault-recovery, enable context-aware customization). In wireless services, this will mean more scope for Dynamic Spectrum Access (DSA), enabled by such technologies as cognitive radios. In wired and wireless services, it will mean more support for QoS-differentiated services.[3]

---

[2] There is a debate as to what constitutes broadband service. The first generation of mass market broadband services are on order 1Mbps, and today's best cable modem services support on order 10Mbps. Between 50-100Mbps is likely to be needed to support a "triple play" offering (voice, data, and video entertainment) and future FTTH systems may deliver upwards of 1Gbps per home. How much is enough is an active topic of public debate.

[3] To date, ISPs have not implemented significant QoS, although such capabilities are widely in use in VPNs. Historically, it has been deemed less expensive to over-provision than to share capacity more

- ***Intelligent networks and devices***: in addition to having more distributed intelligence in networks, edge-devices (under end-user control) also will have more embedded intelligence. AI-enabled agents will assist in supporting unaware computing/communication and will be needed to manage the greater complexity of network management (adaptive interoperability, trust protection, and context-aware customization).

While the end-user experience articulated above is yet to be fully realized, it is worth noting that computing and electronic communication technologies are already pervasive in our social and economic lives. Telephone and Internet access are essential basic infrastructure for business and most (all?) consumers. Computing chips are embedded in our appliances, our cars, and throughout building and factory control systems. Increasingly these computers are being networked and supplemented with sensors to further enhance their ability to be contextually aware of their local environments. While the realization of seemless interoperability, ubiquitous connectivity, or context aware services is nascent, much work is on-going toward realization of this future.

Competition,[4] globalization,[5] deregulation,[6] and convergence[7] have combined to reshape the Information and Computing Technology (ICT) industry value chain, and the capabilities brought by ICT have facilitated the reconfiguration of value chains in many other industries. One key example of this is the modern capability to ***virtualize*** business functions – it is now possible to outsource almost everything. This enables modular, scalable, and flexible business organization. Complementary developments such as the growth of venture funding, increased labor mobility, deeper markets for outsourced business functions (customer support, intermodal transport, contract management), and on-line markets – each of which are themselves dependent on the availability of advanced ICT services -- also contribute toward enabling the virtualization of business enterprises.

These trends have important implications for the design of next generation Internet infrastructure and computing/communication services. In the future ecosystem for ICT, ownership and control will be distributed and decentralized, with end-users (or their devices on

---

dynamically. In the future, the need to accommodate diverse e2e networking requirements are likely to drive wider adoption of QoS.

[4] Competition has accelerated across most ICT segments because of technical convergence and Moore's Law-driven cost reductions. The ability to implement services in software or hardware facilitates cross-industry/cross-technology platform competition and lower costs facilitates scalable entry and exit.

[5] Service and equipment markets are increasingly global in scope, enabled in part by the ability of ICT to lower distance/time-sensitive "transportation" costs that previously separated geographic markets.

[6] Communications regulation has transitioned away from traditional public utility regulation of telecommunication providers and separate regulation of broadcasters toward liberalization and increased reliance on market forces.

[7] Convergence has blurred industry boundaries between broadcasting and telecommunications, between computers and communications, and between network services and end-user equipment-based platforms. Convergence helped provide the driver to deconstruct traditional silo-based regulation of communication services.

their behalf) playing a bigger role in service definition and management.[8] This means that end-to-end management will be more challenging, and potentially strategic.[9] It also has important implications for innovation because the benefits of basic research will be less readily appropriable to private entities and the social costs of strong intellectual property rights protection likely will be higher.[10] Finally, positive and negative externalities will be a common feature of the networking environment.[11]

Need for multidisciplinary research

Assuming the above vision is correct, the future Internet will need to embody in its design greater attention to economic and policy issues. The push for this is two-fold. First, as the Internet becomes *the* platform for our global communications networks and even more inextricably bound into our economic and social fabric (unaware, pervasive, critical infrastructure), the range of stakeholder interests impacted by technical design choices will expand and become even more strategic. Industry convergence, globalization, and liberalization are all impacted by Internet design and rational stakeholders will seek to influence market mechanisms and institutions to advance their particular interests.

Second -- but closely related -- with the increase in intelligence at the edges, the potential for more distributed/local strategic decision-making will make the behavior of the Internet more like an economic/social system in its own right. The adaptive, autonomous systems that will partially comprise the future Internet (agents, gateways, networks) will be capable of evincing self-interested and learning individual and group behaviors, rendering them intelligent economic agents.

Evidence of the above forces has been around for a number of years and includes telecommunications and broadcasting regulatory reform;[12] the increased strategic importance and contentiousness of technical standardization;[13] call for Internet regulation and governance

---

[8] An alternative scenario is that industry consolidation will occur to address the challenges posed by coordinating complex end-to-end networks and increased competition. In arguendo, I will assume public policy will continue to successfully promote competition and will preclude any attempt to monopolize the ICT value chain.

[9] This should be contrasted to provisioning of an end-to-end service by a vertically integrated entity.

[10] The decline or refocusing of basic research toward more narrow, strategically-focused research by the big industrial research labs offers one example of this. With lower costs for personal mobility and enterprise start-up, it is harder for companies to capture the benefits of innovation which, ceteris paribus, may lower private incentives to innovate.

[11] Positive externalities will be associated with network effects and innovation spillover benefits; negative externalities will be associated with congestion/interference and transaction/coordination costs.

[12] We are in the midst of a global transition from silo-based public utility regulation toward increased reliance on markets, driven in part by convergence of computing and telecommunications technologies. Markets require different institutional frameworks for effective self-regulation which are in the process of emerging.

[13] With convergence and globalization, and increased reliance on open interfaces, technical standardization has become increasingly contentious and strategic. Legacy institutions such as the ITU, ANSI, ETSI, and IEEE face significant challenges in remaining responsive while protecting consensus decision-making.

reform;[14] growth of pervasive computing;[15] demise of industry basic research;[16] and, restructuring of the ICT value chain with greater reliance on multi-vendor solutions and open interface architectures.[17] Each of these changes increases the need for a national, publicly-funded, experimental test bed for networking research.

These changes have important implications for computer science and networking research directed toward the design of the next generation Internet:

- Design research needs to become more consciously multidisciplinary, engaging social science research perspectives such as economics, public policy, and sociology/psychology. Future computing/communication systems need to be designed with an eye toward embedding them in social and economic market-based systems in ways that are consistent with human/social norms and public policies. Network design needs to be policy-aware and public policy toward networks needs to be technically informed. This work is needed at the basic research level, before commercialization when stakeholders' strategic positions regarding differing technical trajectories have hardened.
- Internet architecture design will become increasingly contentious and collective coordination and dispute/conflict resolution mechanisms will become more important. This will include anticipation of institutional reforms in the environment in which the Internet will operate, including regulatory reform and market evolution.
- Policy-management capabilities and market mechanisms will be increasingly built directly into the infrastructure. Information sharing mechanisms (e.g., Knowledge Plane),

---

[14] The growth, globalization, and commercialization of the Internet raise important questions about whether traditional decentralized management/governance structures embodied in such entitites as ICANN and the IETF are robust to address future challenges. The debates at WPIS and the failure to coordinate timely migration from IPv4 to IPv6, or to adopt best-practices for end-to-end QoS highlight the stresses confronting the existing processes.

[15] From appliances to cars, computer chips are in everything. Mobile phones (increasingly data-enabled) are everywhere. We are on the cusp of pervasive computing. Always on/everywhere connected services, in many cases operating in the background ("unaware"), will have profound implications for our collective and individual notions of privacy and identity.

[16] The growth of ICT has facilitated a world in which traditional legacy industry structures are increasingly yielding to more distributed, decentralized, and fluid forms of industrial organization. It is now possible to outsource almost anything. The "virtualization" of firms has increased the dynamism of industries and markets, allowing more scalable and faster entry/exit across a wide-range of industrial sectors. One downside of this is that basic research by industry is much less easily sustainable. The decline of leading industrial research labs or their refocusing from basic toward more narrowly strategically valuable research is a by-product of this.

[17] The industry value chain for ICT (from semiconductors to equipment to applications to services) is more decentralized/distributed and mixed-integrated than ever before. Industry convergence which has blurred industry boundaries between service and equipment providers, communications and computers, and networks and edges coupled to the rise of open interfaces have increased competition and the need for cross-industry, multi-vendor collaboration. This enhances the likelihood that innovation will result in significant spillover benefits.

resource markets, virtualization, and distributed QoS are all technical manifestations of this.

- Internet architecture and design needs to be considered within the larger ecosystem of regulatory, social, and market mechanisms (industry structures) in which it evolves. The larger ecosystem and the Internet design need to co-evolve and design needs to consciously reflect that co-evolution. The capabilities and architecture of the Internet will influence the ease with which Internet users may use (design new services) or abuse (DDoS) the Internet and the ease with which policy-makers may regulate it (e.g., implement open access policies). For example, it is a design choice whether the Internet ought to embed capabilities for intellectual property protection or whether such protection ought to be addressed at another level (e.g., legal and regulatory enforcement).

Social science basic research also will benefit from closer multidisciplinary collaboration with computer science/network researchers. For example, the new computational tools and greater access to a faster-clockspeed experimental platform afforded by the GENI network can advance economic theory regarding the dynamics of complex, interactive economic systems. Econometrics, experimental economics, game theory, biological/evolutionary economics, and information economics are several of the economic sub-disciplines which could benefit from collaboration with computer science/network researchers. In addition to having access to better tools, economists can benefit from exposure to computer science/networking theory in the form of such tools as graph theory, information theory, and AI/bounded rationality.

GENI as a viral incubator

GENI can play an important role in helping to foster an environment for collaborative research in next generation Internet architectures and services. It will take time to build a collaborative community of researchers with sufficient shared expertise to facilitate communication. To date, most economists have viewed network technology as a black box, while most computer scientists/networking researchers poorly understand economics. A shared experimental test bed which recognized the common challenge and need for multidisciplinary work would provide a framework for building the needed collaborative capability.

Building the community of multidisciplinary researchers will require a long-term commitment of resources. The ecosystem for multidisciplinary work is poorly developed. Academia does not typically reward multidisciplinary work, which makes it especially risky for young faculty. The rewards for multidisciplinary collaboration are greater in industry which better understands the business imperatives, but commercial enterprises are less able to undertake basic research. Many of the most successful multidisciplinary research projects in this space have been undertaken by industry/academic collaborations, with industry providing the funding and contributing to the strategic direction of the research (if not also contributing directly to the research).

Publishing multidisciplinary research is also a challenge. There are a lack of suitable journals and in-depth multidisciplinary work imposes difficult challenges for readers. Good multidisciplinary research needs to avoid over-simplification, and ideally, should sustain critical evaluation from specialists in each of the disciplines involved.

GENI can play an important role in enabling experimentation and collective learning about network innovations that are too radical to test in the existing Internet. For example, coordinating adoption of new Internet standards or technologies such as the conversion to IPv6 or replacement for the BGP routing protocol involves significant switching costs. GENI could offer a mechanism for experimenting with alternatives at scale and with a faster clockspeed to examine stability and long-term equilibria effects. GENI experimentation could provide a vehicle for viral adoption, serving as a nursery for nascent ideas that are pre-commercial.

GENI's success in promoting multidisciplinary research will be enhanced if its selection of research topics are carefully chosen to emphasize problem domains where the need for collaborative research of the sort discussed herein is most obvious and immediate.

Topic Areas for Multidisciplinary GENI Research

Several likely candidates for GENI-based multidisciplinary research include:

- Innovation in on-line markets: eCommerce continues to grow in importance and with the emergence of social networking, virtual worlds, and non-traditional (i.e., non-neoclassical market) resource allocation models (e.g., commons) there are a wealth of market mechanisms that exist principally in the Internet but have obvious economic implications for the off-line world as well. Understanding the interaction effects between such emerging on-line phenomena and "bricks and mortar" markets and user behavior, valuing such activity, and identifying appropriate institutions/mechanisms for managing such emerging markets are all interesting questions for economists, lawyers, and psychologists/sociologists. Some key challenges include enabling interoperability across markets (e.g., virtual worlds) and appropriate public policies for consumer protection (including privacy). Some interesting social science questions include understanding the formation of public opinion and value in such markets.
- Internet as a measurement platform: as traffic continues to grow in volume and heterogeneity, and as industry control fragments and direct regulatory oversight is relaxed, there is a growing need to generate, collect, and share traffic metrics. Such data is needed for market regulation, for enforcing QoS contracts, and for network planning and management. The shear volume of possible information and the overhead costs associated with processing it make it necessary to identify appropriate summary statistics. Determining what to collect, how to share the information (securely, privacy preserving), and how to recover the costs of data measurement pose difficult incentive compatibility/mechanism design challenges and require new types of data collection technology.
- Digital rights management (DRM): Increasingly, the Internet is a platform for sharing content that is subject to differing intellectual property rights regimes. This ranges from copyrighted material to creative commons. Most of the DRM proposals are incompatible and many of the mechanisms have over-emphasized the need for protection by content owners over the fair-use rights of end-users. GENI could provide a platform for experimentation in novel DRM management regimes. The adoption of an appropriate DRM mechanism for the Internet may benefit from GENI as a platform for viral adoption analogous to the role of GENI in helping coordinate the adoption of other network-wide technologies. The DRM challenge represents a compelling public policy and business challenge to the entire content-distribution value chain in both the on-line and off-line world and so this is an area of strong multidisciplinary research interest.

- Dynamic spectrum access (DSA): as noted earlier, the future of networks will be increasingly wireless and the future of wireless will involve greater unbundling of applications and infrastructure, and infrastructure and RF spectrum. Historically, wireless infrastructure has been dedicated to specific frequencies and spectrum management was based largely on command and control. In the future, access to the RF will be managed more by markets and frequencies will be shared more intensively across infrastructure, applications, and uses. In short, spectrum access will be more dynamic. The rise of spread spectrum technologies provides one obvious example of this trend. The emergence of cognitive/software radios is another example. Building the ecosystem for DSA will require a number of complementary developments that GENI could contribute to. For example, DSA will require better local information about sharing opportunities (or interference) which will require better information about wireless usage and ownership. GENI could provide a testbed for RF sensing integration with improved spectrum rights database management. Alternatively, GENI could provide a test bed for simulating collective behavior of cognitive radio systems to test the efficacy of protocols and coordination strategies. Some specific problem areas which would benefit from better testing would include management of complex radio (including cognitive radio) infrastructure, public safety radio sharing, and white space access.

Position Paper

For: GWEB1 Workshop

David Waterman, Indiana University

May 27, 2008

Prospects for a Virtual Video Store on GENI

I. Introduction

Copyrighted broadband entertainment, especially recent movies, is surely attractive to Internet end-users. A way to take full advantage of the Internet's architecture to offer these products, by sweeping away high download times, limited variety, and user-unfriendly features is enticing—especially with the diffusion of high definition (HD) software and hardware.

Commercial IPTV distribution of movies and other copyrighted products essentially began with *cinemanow.com* in about 2000, followed by *movielink.com* in 2002 and by iTunes and Amazon in 2006. Though rapidly growing in percentage terms, little commercial value has so far been produced. In 2007, Internet download-to-own (DTO) or video-on-demand "rental" (VOD) services generated a fraction of 1% of studio revenues from distributor of theatrical films in the U.S., in notorious contrast to the 52% from physical distribution of DVDs.[1] (Appendix Table 1 shows a more aggregated historical summary of U.S. distributor revenue by media for theatrical films.) Foreign market contrasts are even greater, mostly reflecting more nascent IPTV movie distribution. TV series account for around 10% of DVD sales and rental volume in the U.S. Internet distribution of these programs, mostly with advertiser support, has gotten a lot of publicity, but the revenues (apart from any promotional value) are surely negligible. By far the most commercially successful use of the Internet for broadband entertainment to date (apart from IPTV pornography) is for ordering DVDs to be shipped by Netflix and similar services via the postal service.

[1] *Screen Digest* (Sept. 2007, pp. 275) estimated U.S. spending for all online movies (DTO + VOD) to be $32 million in 2006, and projected $110.4 million for 2007. Roughly half of retail spending accrues to the distributors.

Do the gaping contrasts between commercial success of "actual" video stores and what many observers see as an ongoing commercial disappointment of VOD via IPTV reveal the shortcomings of the Internet as a commercial movie or TV distribution medium?—or do they just highlight the vast potential for a virtual video store when technology barriers finally fall?

In my view, the powerful intuitive appeal of the widely held “anything, anytime” vision for a seamless online video delivery system has a real basis, and progress toward it is inevitable. But, such a system may be less of a Holy Grail for the future of entertainment delivery than some have imagined. History, law, and program distribution economics suggest obstacles likely to endure beyond the collapse of technological constraints.

II. Seven desirable characteristics of a commercially successful virtual video store

In general, useful experimentation on GENI with a virtual online video store should mimic our best vision of what a real commercial system would look like. Here is an incomplete list--some obvious, some perhaps less so.

(1) High speed

Current constraints are evident, especially with high definition (HD), in spite of innovations to shorten delays to begin viewing. How much download times affect demand is not clear, but it is surely substantial.

(2) High variety

Consumers prefer a system that allows one stop shopping and offers the highest possible proportion of all available programming. Though a disproportionate amount of viewing is accounted for by a small number of entertainment products, market experience suggests that “option demand”—the perceived freedom to indulgence any wish—is significant. The several thousand titles of Blockbuster’s brick-and–mortar stores and the 85,000 + inventory of Netflix are examples of apparently successful marketing advantage. So far, Internet video services have offered poor variety of both recent and catalog products (more on this below).

(3) Security/Copy protection

DRM systems for legitimate Internet distribution of movies and other copyrighted products are sophisticated and apparently secure, but the threat of user copying and redistribution is always a concern for copyright holders, especially outside the United States.

(4) A seamless TV set interface

Regular people want to watch movies or TV programs on their TV sets, and this interface, along with the access and ordering system, needs to be seamless and intuitive. At this writing, critics agree there is a long way to go.

(5) Early release programming

The same movies or programs available sooner or at the same time on other media may be the biggest current constraint on commercial IPTV distribution. Studios have steadily moved toward simultaneous VOD and DVD movie release in 2007-08, but most films have appeared on Internet VOD systems (and also on multi-channel PPV or VOD systems) about 30 days after DVD release—by which time DVD demand is largely satisfied.

Though often attributed to myopic attitudes, the Hollywood studios have some compelling economic incentives to move cautiously or to maintain the status quo for IPTV release dates. Major feature films are released in a timed sequence, usually beginning with theaters, that is designed to segment high from low value audiences and thus to effectively price discriminate--or maximize the total revenue per viewer according to their willingness to pay. For example, distributor net revenue per transaction for physical DVD rental is far lower than for electronic VOD rentals, giving studios apparently strong incentives to favor the latter market. DVD sales, however, may appeal to higher value consumers than

DTO service because of the packaging and the DVD extras--thus favoring earlier DVD release. A complication is that copyright law essentially dictates that DVD sales and rentals must be simultaneous; marketing, piracy, and other factors probably enter the equation.

As technology develops, current pecking order of the movie release system might be overturned. When calculated on a revenue-per-number-of-viewings basis, theater attendance compares favorably with a DVD sale, which market research studies show to generate several individual viewings on average. However, diffusion of high quality HD home theater systems could attract enough very high value viewers to induce VOD (and DVD) release simultaneously with that of theaters—undoubtedly enhancing commercial online movie demand

In summary, it is a good guess that studios will continue their move to simultaneous DVD and VOD release—and otherwise to do what they can to replace the large but very inefficient physical DVD rental system with electronic DVD.   It would be unwise, however, to over-parse the studio's complex release decisions to anticipate much further.

For TV series, the market seems headed toward (virtually) simultaneous Internet and broadcast release, mostly on a within-program advertising model. Compared to movies, TV series seem to be low value programming with relatively low direct payment potential.

(6) **Consumer friendly pricing**

A la carte pricing, such as theaters or VOD now use, is a very likely component of a virtual video store; it has compelling market skimming advantages for high quality products released near the front of the sequence.

A subscription component, however, is also possible. Consumers have resisted a la carte pricing in various contexts, and subscription plans like those of the Netflix

DVD postal service, have done well. As the Appendix Table implies, consumers currently spend much more money on monthly subscription movie networks than they do on all a la carte priced movies combined.

One can imagine lots of pricing models, including online sales tied with DVD sales, various other forms of bundling, or advertising—especially for older IPTV TV programs or movies. Also entering the equation is likely to be tiered pricing of broadband access itself based on download quotas.

(7) Competitive advantages over other media

To be successful, an IPTV virtual video store must create real value for consumers relative to alternative media. In particular, cable television is a highly efficient technology for one way broadband distribution. Operators are earnestly competing to offer VOD services and are increasingly mimicking the Internet's advantages with interactivity, lack of capacity constraints, target marketing, bundling, and other attributes.

It is also rash to assume that brick-and-mortar rental stores or Netflix-type DVD retailers will fade away. DVDs, for example, offer packages for home or gifts, and stores offer visual observation and other differentiated features.

III. Effects on other media: a non-zero sum game?

If distribution of copyrighted video products via the Internet is commercially successful, it will obviously impact other media negatively--especially it would seem, DVD and its retailers.

The rapid decline of music CDs is a foreboding model, but the movie industry's own historical experience (as illustrated by the Appendix Table), suggests that improvements in market segmentation brought forth by new movie media, such as IPTV, will result in a net expansion of total content industry revenues.

IV. Other considerations for experimental design

It is hard to predict the extent to which the desirable characteristics of a viable commercial virtual video store listed above may already be achieved by the time GENI is available for experimentation. It seems likely, however, that there will remain a frontier for serious improvement.

To the extent that faster speed and other technical improvements over the public Internet can be accomplished by GENI, it offers a natural environment by which studios can simply observe the workings of their existing business model, but with fewer constraints. Periodically since the 1960s, studios have also shown willingness to experiment with different models, such as earlier release dates or different price points. Recently, several studios reportedly cooperated with Comcast, a major cable operator, to offer a "virtual video store" in Pittsburg and Denver, that involved simultaneous VOD and DVD release, while the usual 30 day + PPV window remained in the rest of the U.S.

Potentially, GENI can offer a variety of experimental opportunities with earlier release dates and different pricing models. Any experiment, however, must garner a high level of studio cooperation or it will be ineffective. Major limitations on studio flexibility with release dates or pricing schemes that vary from the national pattern, however, must be recognized. Movies are expensive production investments that are launched with expensive national advertisements, so release date experiments, for example, have an unnatural aspect. The major studios envision their future world differently as well, and behind every individual movie is also an individual producer who is justifiably averse to the risk of any experiment. Early experimental release also tends to carry a negative quality signal about the product.

On the margin, however, better understanding the effect of window shifts and alternative pricing points is a subject of intense interest to studios that will encourage their cooperation.

Certainly other entertainment models are also viable for experimentation. For example, more realistic environments and faster response times in virtual worlds are possible both for commercial purposes and for academic experimentation.

Appendix
Sources of Revenue from Distribution of Theatrical Feature Films by Media: U.S. Distributors (%)*

	1981	1985	1990	1995	2000	2005	2006
Domestic Market							
Theaters (%)	50	37	30	26	26	22	22
Total home video (%)	14	36	38	51	49	54	52
Subscription pay TV (%)	16	15	14	8	8	7	7
All PPV/VOD (%)**	0	< 1	< 1	2	2	2	3
Broadcast + basic cable (%)	20	12	17	12	15	14	15
Total (%)	100	100	100	100	100	100	100
Total domestic revenues ($ billions)	$2.4	$4.3	$6.9	$10.2	$15.3	$21.9	$22.1
Foreign Market							
Theaters (%)	84	43	35	28	26	20	20
Total home video (%)	8	42	40	45	38	55	56
Subscription pay TV (%)	0	< 1	5	9	14	10	10
All PPV/VOD (%)**	0	0	< 1	< 1	< 1	< 1	1
Broadcast + basic cable (%)	8	14	20	17	21	13	13
Total (%)	100	100	100	100	100	100	100
Total foreign revenues ($ billions)	$1.1	$2.1	$6.1	$9.3	$13.3	$22.4	$23.1
Total revenues ($ billions)	$3.5	$6.4	$13.0	$19.5	$28.7	$44.3	$45.2

***Domestic merchandise licensing revenues and international merchandise licensing revenues are not included in the table.**

****includes hotel PPV, airlines and "other." Separate data for "home PPV/VOD" (Internet + cable + DBS +other multichannel providers) became available for the domestic market after 1988, and since at least 1995, home PPV/VOD has accounted for more than 75% of the "All PPV/VOD" category total.**

Sources: Waterman (2005), Hollywood's Road to Riches (Harvard Univ. Press): Appendix C, pp. 288-292, based primarily on Motion Picture Investor, Kagan Research, various issues; also Motion Picture Investor, August, 2006, p. 4; Aug, 2007, pp.6-7.

(*Sung Wook Ji/ David Waterman, 5/28/08*)

GENI User Opt-In
White Paper

Barbara van Schewick,
Stanford University[1]

July 15, 2008

Getting a significant number of end users to use GENI applications is critical for GENI's success. Without real traffic, GENI will not be able to reach one of its main goals – validating networking research at scale and under realistic traffic loads.[2]

In this White Paper, I would like to focus on two potential risks. Ironically, they only become relevant if GENI manages to meet its goal and one or more applications attract a large number of users.

First, in the absence of a network neutrality regime banning this type of behavior, the traffic created by highly successful GENI applications may motivate Internet Service Providers to block or discriminate against GENI-related traffic.

The opt-in concept assumes that GENI (or the architectures running on top of the GENI facility) will offer mechanisms that will enable users whose end hosts are not directly connected to the GENI infrastructure ("off-GENI users") to use the networking protocols or applications offered by GENI.[3] This implies that a very successful application that is used by a lot of off-GENI users will drive large amounts of traffic to or from the GENI network. Depending on the interconnection agreements of the home networks of these users, the increased traffic flows may significantly increase network providers' interconnection fees. This is a natural consequence of the opt-in concept. In fact, papers such as Ratnasamy, Shenker & McCanne (2005) that rely on end-user opt-in as an

---

[1] Assistant Professor of Law and (by Courtesy) Electrical Engineering, Co-Director, Center for Internet and Society, Stanford Law School. E-Mail: schewick@stanford.edu.
[2] Computing Research Association GENI Community Advisory Board (2007), p. 3.
[3] See, e.g., Clark, Shenker & Falk (2007), pp. 87-88.

important mechanisms to change network providers' incentives to deploy new network protocols view the flow of money from ISPs who have not deployed a new networking protocol to the ISP that has deployed the new networking protocol as a crucial component of the mechanism that provides incentives to deploy the new protocol: "What we assume here is that if IPvN [the new Internet Protocol deployed by an innovative network provider, the equivalent to GENI] attracts users, then revenue will flow towards those ISPs offering IPvN. […] We also posit that an ISP that attracts new traffic, by offering IPvN, will also gain revenue possibly due to increased settlement payments (traffic from non-offering ISPs to offering ISPs would increase)."[4]

While the increase in GENI-related traffic leaving or entering the users' home networks will increase network providers' costs, it will not increase their revenue, as long as flat-rate pricing remains dominant. This may motivate network providers to throttle or block GENI-related traffic in order to limit interconnection fees.[5] Instead of outright blocking or throttling, network providers may send GENI-bound traffic over cheap, congested wholesale links to constrain their costs. Such behavior would significantly limit the performance GENI applications could realize.

A look at industry white papers suggests that this is not just a hypothetical concern. For example, according to a White Paper produced by Sandvine,[6] the vendor that Comcast seems to have been using to manage BitTorrent and other peer-to-peer applications, blocking external lechers (i.e. peers on other networks that want to download a certain file) from connecting to internal seeders (i.e. peers on the network provider's network that make a file available for upload) is the default strategy employed by Sandvine's session management policy. The White Paper justifies this approach by the significant savings in interconnection fees which can be realized using this strategy.

4 Ratnasamy, Shenker & McCanne (2005), pp. 314-315.

5 For a general description of this problem, see MIT Communications Futures Program and Cambridge University Communications Research Network Broadband Working Group (2005) who call this „the broadband incentive problem". See also van Schewick (forthcoming 2009), chapter 5.

6 Sandvine (2004).

Home network providers may also be motivated to interfere with highly successful GENI applications if these send a significant amount of traffic into the network, putting pressure on the upload portion of the last-mile access networks. Comcast's behavior towards BitTorrent and other peer-to-peer applications is an example of such behavior.[7] All this suggests that highly successful GENI applications that create a significant amount of non-local traffic for the home networks of off-GENI users or send significant traffic over the upload portion of the last-mile links may be blocked or throttled by the users' home network providers.

A network neutrality regime that prohibits this type of behavior and limits congestion management to non-discriminatory methods that do not single out specific applications or classes of applications would mitigate this risk.[8] To make sure GENI's interests are considered by the FCC, the GENI project office or the National Science Foundation could submit its concerns to the ongoing FCC proceeding regarding broadband industry practices.[9] At a minimum, careful thought should be given to the technical and financial impact of GENI-related traffic on the home networks of off-GENI users and to measures that may mitigate these problems. Implications of the opt-in concept for GENI's interconnection strategy should be considered.

Second, building successful applications that attract a significant number of users may become a problem under laws that limit competition by public entities with private enterprises, if these applications compete with existing applications produced by private entities operating for profit.

There are a number of strategies for attracting end users: one is to offer applications that enable users to do things they could not do with applications that are available in the current Internet. This could be completely new applications, applications with new

[7] Comcast has shown to interfere with peer-to-peer applications such as BitTorrent and Gnutella by sending TCP reset packets, see, e.g.,, Schoen (2007); Eckersley (2007). Comcast admits that it sends TCP reset packets, but claims that it only to terminate unidirectional sessions that are dedicated to uploading content from a Comcast user, see Cohen (2008).
[8] See generally, van Schewick (2008a); van Schewick (2008b).
[9] Broadband Industry Practices, WC Docket No. 07-52.

features, or applications that simply offer better performance than existing applications. Another strategy may be to offer an application or service for free that on the existing Internet is only available for a fee. While the second strategy may be quite effective in attracting users, it may raise legal concerns. If the existing applications are produced by private entities operating for profit, they may be able to claim that the GENI applications constitute "unfair competition by the government". I'm not an expert in this area of law, but several states have laws that prohibit state entities from competing with private entities. For example, Iowa Code §23 A prohibits public entities, including public institutions of higher education, from competing with private enterprises unless this is explicitly authorized by statute, ordinance, rule or regulation. Iowa Code §23 A.2(2) allows the Board of Regents to adopt exemptions by rule.[10] Other states such as Maine have special committees such as the Advisory Committee on Fair Competition with Private Enterprise that are responsible for "approving services and goods sought to be provided by state agencies that are not otherwise allowed by law, and may compete with private enterprise."[11] A company whose business was destroyed by GENI may rely on these laws (or at least appeal to the rationale underlying them) to claim that using public funding to make an application available for free constitutes unfair competition by the state. While it is unclear whether these rules would apply to GENI-related efforts, it may be worth exploring the issue to make sure that GENI does not run into problems here.

References

Clark, David, Scott Shenker & Aaron Falk. 2007. *GENI Research Plan (Version 4.5).* (No. 06-28). GENI Research Coordination Working Group and GENI Planning Group.

Cohen, David. 2008. *Written Testimony before the Federal Communications Commission at its First En Banc Hearing on Broadband Management Practices.* Federal Communications Commission, http://www.fcc.gov/broadband_network_management/022508/cohen.pdf.

[10] For a brief overview of the exemptions, see State of Iowa Board of Regents (2007).
[11] http://www.maine.gov/dafs/advcom/index.htm.

Computing Research Association GENI Community Advisory Board. 2007. *Report on the Review of the GENI Research Plan, January 15-16, 2007.*

Eckersley, Peter. 2007. Comcast is also Jamming Gnutella (and Lotus Notes?). *EFF Deeplinks Blog*. www.eff.org/deeplinks/2007/10/comcast-also-jamming-gnutella-and-lotus-notes

MIT Communications Futures Program and Cambridge University Communications Research Network Broadband Working Group. 2005. *The Broadband Incentive Problem (White Paper).* Cambridge, MA.

Ratnasamy, Sylvia, Scott Shenker & Steven McCanne. 2005. "Towards an Evolvable Internet Architecture". Paper presented at the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '05). Philadelphia, PA.

Sandvine. 2004. *Session Management: Bit Torrent Protocol: Managing the Impact on Subscriber Experience.* Waterloo, Ontario, Canada: Sandvine.

Schoen, Seth. 2007. EFF Tests Agree with AP: Comcast is Forging Packets to Interfere with User Traffic. *EFF Deeplinks Blog.* http://www.eff.org/deeplinks/2007/10/eff-tests-agree-ap-comcast-forging-packets-tointerfere.

State of Iowa Board of Regents. 2007. *Annual Report on Competition with Private Enterprises.* http://www2.state.ia.us/regents/Meetings/DocketMemos/07Memos/may07/0507_ITEM42.pdf.

van Schewick, Barbara. 2008a. *Oral Testimony before the Federal Communications Commission at its Second En Banc Hearing on Broadband Management Practices.* Federal Communications Commission. www.fcc.gov/broadband_network_management/041708/vanschewick-oral.pdf.

van Schewick, Barbara. 2008b. *Written Testimony before the Federal Communications Commission at its Second En Banc Hearing on Broadband Management Practices.* Federal Communications Commission. www.fcc.gov/broadband_network_management/041708/vanschewick-written.pdf.

van Schewick, Barbara. forthcoming 2009. *Architecture and Innovation: The Role of the End-to-End Arguments in the Original Internet.* Cambridge, MA: MIT Press.

GENI End-User Opt-In: Incentive-Centered Design For Virtual Organizations

Jeffrey K. MacKie-Mason*

1 GENI end-user opt-in: virtual organizations

In this white paper I discuss ideas — and needed research — that address challenges in designing GENI applications to increase end-user opt-in.

To focus ideas, I concentrate on a class of applications that seem plausible candidates for early implementation on GENI: virtual organizations (VOs) for research and education, including collaboratories, community data systems and distributed research centers. Examples include large-scale scientific equipment sharing, data analysis and instrumentation projects, such as the National Earthquake Engineering System, genomics databases, and AIDS research collaboratories.. NSF and many scholars believe that virtual organizations will be a major application built on top of modern cyberinfrastructure. As such, they are receiving attention and research support from the new Office of Cyberinfrastructure (OCI) at NSF.[1] VOs certainly are only one type of promising end-user application for GENI, but they are one clear target. In any case, many of the ideas I discuss below also apply to designing for opt-in to other GENI applications as well.

One of the pivotal features of most VOs is support for, and dependence upon knowledge-sharing. For the most part the data, analysis, images, stored discussions, research papers and so forth, are content created, contributed and then shared by the participants in the VO. Thus, VOs rely on *user-contributed content* (UCC), much as Wikipedia and Flickr rely on the voluntary contribution of content as their perhaps most important input to production. UCC is a particularly apt consideration for GENI, because its existence and quality depend inextricably on end-user opt in. Getting users to opt-in — that is, to voluntarily expend time and effort to create, annotate and organize, deliver, maintain and support content — is the core problem I discuss below.

2 Knowledge-sharing in virtual organizations

> A virtual organization (VO) is a group of individuals whose members and resources may be dispersed geographically, yet who function as a coherent unit through the use of cyberinfrastructure (CI). A VO is typically supported by, and *provides shared and often real-time access to, centralized or distributed resources, such as community-specific tools, applications, data, and sensors, and experimental operations* [NSF Workshop on Virtual Orgs, 2008, 23, p. 4, emph added].

Virtual organizations (VOs) to support science, engineering and education are made possible by exponential declines in the cost of computing and communication *technology*. However, VOs depend critically on the activities and inter-activities of the *humans* who comprise them. Performance follows from the contributions and efforts of autonomous, motivated people, and the quality of those efforts. In particular, I focus on *knowledge-sharing* in VOs, which is a crucial activity for almost all VOs, and is the main purpose for many.

One role of a traditional *centralized* organization is to command, or to provide strong motivations, to induce participants to share their knowledge in ways that benefit the organization. Unless the motivations of autonomous participants are considered explicitly in the design and operation of a *virtual* organization, this pivotal function will not be provided. The quantity and quality of knowledge sharing will be haphazard at

*Associate Dean, School of Information, University of Mcihgan; Arthur W. Burks Collegiate Professor of Information and Computer Science; Professor of Economics and Public Policy. Contact information: http://www.jeff-mason.com/.

[1] E.g., the currently running grants competition for Virtual Organizations as Sociotechnical Systems (VOSS), http://www.nsf.gov/pubs/2008/nsf08550/nsf08550.htm.

best and under quite general circumstances, formalized in the public goods and mechanism design literatures, effort and quality will be below the socially-desirable level.

In particular, the success of a VO depends on reaching a critical mass (quantity) of active participation, and on the quality of the participation. In VOs, participation is voluntary, and fluid; this is especially the case for *loosely-federated* VOs, which are plausibly of especial interest for GENI. Humans are autonomous, and their participation cannot be compelled, nor can their effort towards quality. I focus here on *designing incentives congruent with behavioral motivations* for knowledge sharing. Indeed, one of the key research challenges singled out by the NSF Workshop quoted above is "Motivation and Rewards":

> Given that collaboration is not mandatory in many sciences, what motivates people to participate in VOs? We need to better understand how to design socio-technical systems that leverage or enhance existing motivations, given the nature of the science. . . . In some cases, external incentives and rewards will be the draw whereas for others the costs of or the barriers to participation are stronger determinants. [23, p. 33]

In a major study based on analysis of dozens of existing VOs, [13, p. 53] ask "Why are scientific collaborations so difficult to sustain? . . . only a few of these efforts have succeeded in sustaining long-distance participation, solving larger-scale problems, and initiating breakthrough science." For several important categories of VOs, they identify individual and interpersonal motivation issues as critical. For example, they conclude that "Community Data Systems [group created, shared information resources] can be viewed as public goods projects that may find themselves in a social dilemma related to motivating contributions" [13, p. 61]. "Virtual Communities of Practice" (networks of knowledge-sharing individuals who share a research area) "Must work hard to maintain energy and participation rates with a shifting set of participants" [13, p. 63]. "Distributed Research Centers" (functioning like university centers but at a distance, sharing knowledge and resources in a unified area of interest) "as the most organizationally ambitious project type, these collaboratories....must gain and maintain participation among diverse contributors" [13, p. 65].

Once a corpus of knowledge exists, is organized, and is stored in digital format, sharing is not a difficult problem (except perhaps for bandwidth and other scale issues, but for virtual "organizations" these are usually not binding constraints, as they may be for certain Internet-scale knowledge sharing problems). The information can be replicated and distributed at near-zero incremental cost, and in most cases information is *non-rivalrous*: its use by one person does not degrade its value to another (in contrast to, for example, a chair, in which, for the most part, only a single person can sit at a time). Of course there are interesting research questions concerning information organization, storage and retrieval. But for the creation and maintenance of successful knowledge-sharing VOs the specific problem is this: From where does the knowledge-to-be-shared come, and of what quality is it? I focus on issues of *content contribution*.[2] In particular,

1. Why will potential contributors participate, and make the effort to *contribute* knowledge to the shared resource? How can a VO be designed to encourage greater contributions?

2. Why will participants be motivated to exert effort to provide not just quantity, but *quality* (e.g., organization, documentation, metadata, tools for using and accessing the knowledge)?

These pivotal problems are well-known to those who develop and participate in VOs, as well as in other loosely-federated knowledge-sharing communities. For example, the first problem is often referred to as obtaining "critical mass". For a knowledge resource to attract participants to a VO requires that there be enough active participants in the first place to create the knowledge resource.[3] The problem is also familiar from the explosion of loosely-federated online social, technical and educational communities for whom *user-contributed content* determines success or failure (Wikipedia, health information sharing services, CiteULike, Amazon book reviews, Yahoo! Answers, and the Stanford Encyclopedia of Philosophy are among the successful; the equally interesting are the thousands that have *not* achieved critical mass).

The second problem (effort to provide quality) is very familiar, for example, to organizations that seek to collect and manage scientific datasets for sharing and reuse by researchers [9, 21, 5, 14]. Although these

[2] In other contexts known as problems of "user-contributed content".

[3] This is known as a network externalities problem: due to the feedback loop there is a problem of achieving critical mass; see, e.g.,[38, 26, 25].

motivational problems are well-recognized, they have received less attention in research on VOs.[4] The quality problem, of course, feeds back into the opt-in participation problem: if the knowledge resources available in a VO are not of high-quality, then potential users will choose not to participate.

3 Incentive-centered design

I argue that we should approach these problems through incentive-centered design (ICD): an emerging, multi-disciplinary research approach to designing systems whose performance depends unavoidably on *autonomous*, *motivated* human behavior. Human participants are smart, distributed and — crucially — autonomous components, with their own information sets, beliefs and motivations. We can draw directly on the sciences of motivated behavior — in particular, microeconomics, social psychology and game theory — to model incentives, individual responses to them, and inter-individual strategic awareness and behavior.[5] Because humans are non-programmable, unlike the information *technologies* they employ, I advocate an eclectic approach that employs mathematical modeling, numerical methods, human subject experiments [6] and field studies[7]

The design of a virtual organization — and the information technology to support it — can greatly influence the incentives people have for participation and quality effort. For example, in a site that relies on user-contributed ratings to make personalized movie recommendations (`http://www.movielens.org`), under-contribution is common. More than 22% of the movies listed on the site have fewer than 40 ratings, insufficient for the software to make accurate personalized predictions about which users would like these movies [20][8]. One design characteristic is that the site is not social, i.e., social norms and goals were not visible to users. Based on insights from social comparison theory, my colleagues [17] designed a field experiment to test the use of social information on MovieLens contributions. After receiving information about the median user's total number of movie ratings, users below the median exhibit a 530% increase in the number of monthly movie ratings, while those above the median do not necessarily decrease their ratings.

The theoretical foundations for ICD research are deep. The sciences of motivated behavior have been important since at least the early part of the 1900s. Just in economics several Nobel prizes have been awarded for modern advances in these areas[9]; fundamental contributions are numerous and important in other fields as well, especially in social psychology. ICD has recently been applied to research on a number of social and organizational information systems. For example, my colleagues and I have addressed incentive mechanisms for managing Internet congestion [e.g., 46, 51, 50], least-cost interdomain routing [28], community knowledge sharing (CommunityLab) [e.g., 33, 17, 15], sharing large-scale experimental apparatus across a nationwide scientific collaboratory (NEESgrid) [71]), social firewalling for home computer security [76, 77], manipulation-resistant recommendation systems[64, 30, 65], evaluation (reviewer) systems [6], online contributions to a shared knowledge resource (IPL) [16], incentives to contribute to a question-answer database (Google Answers) [18], incentives to share content in a peer-to-peer network [35], online auction design [e.g., 53, 52, 59, 63, 79], exit from a knowledge-sharing community (Wikipedia) [36], incentive mechanisms to reduce inappropriate content contributions (e.g., spam) [19], reputation feedback in electronic markets [54], and persistent pseudonyms for online identity [31].

4 Designing to motivate participation

The basic problem: why would a busy person volunteer her time to prepare, package and share knowledge in a VO, when she herself already has the benefits of access to the information? Knowledge and information

[4] "Many early projects ran into motivation and incentive issues as unanticipated and poorly understood roadblocks....A few high-profile collaboratories have documented these issues...including the Upper Atmospheric Research Collaboratory, the Environmental Molecular Sciences Laboratory, SEQUOIA, and WormBase" [11, p. 251].

[5] My research colleagues and I also draw on computer science and engineering when we apply ICD to information systems.

[6] For example, my colleagues and I recently conducted an online experiment to test multiple theories of motivating financial contributions to support an online information community [16], and and my colleagues conducted human subject laboratory experiments to test the effectiveness of allocation mechanisms for equipment time in large scientific virtual organizations [71].

[7] For example, I implemented and studied the large-scale PEAK system for online distribution of scholarly publications [e.g., 10, 47, 49]. In a more recent example, I have begun a field data study of the dynamics of Wikipedia editor participation, and in particular, their motivations when they choose to depart the project, [36].

[8] There are currently more than 125,000 members on MovieLens.

[9] E.g., Simon (1978), Harsanyi, Nash, Selten (1994), Mirrlees and Vickrey (1996), Akerlof, Spence, Stiglitz (2001), Kahneman and Smith (2002), Aumann and Schelling (2005), Hurwicz, Maskin, Myerson (2007); see `http://nobelprize.org/nobel_prizes/economics/laureates/`.

are *nonrivalrous*: their use by one does not materially reduce the value for another [73, p. 414]. As a familiar example, once National Public Radio broadcasts a program, consumption by one listener does not crowd out consumption by another. Nonrivalry is a defining characteristic of a *public good* [67].

The classic solution is government provision (e.g., installing street lights). But sharing VO participant knowledge requires that autonomous, motivated *individuals* voluntarily contribute knowledge. This is a problem in the *private provision of public goods*, which generally results in underprovision [67, 8]. (The social psychology literature refers to the problem as "social loafing".) Theorists have proposed several possible approaches [32, 75, 7, 1, 55]. However, the solution for knowledge-sharing in VOs is not straightforward for two reasons. First, barring incentives such as monetary payments, those who already have the knowledge may already enjoy most or all of its benefits without incurring the cost of sharing. Second, nearly all of the proposed mechanisms are for contributions made as money (as in charitable giving) which is homogeneous, substitutable and additive, none of which are characteristics of knowledge. Furthermore, in many VOs, monetary mechanisms are not an option. Therefore, I believe that it is important to pursue research on mechanisms with *non-monetary* incentives to motivate knowledge-sharing in VOs.

The (natural) limitations of the economic literature for the knowledge-sharing problems of VOs justifies a multidisciplinary ICD approach combining the well-developed economic theories on the private provision of public goods with the *behavioral and psychological* literatures on intrinsic, non-monetary motivation. For example, researchers have measured responses to motivators such as a sense of fairness [39], and affect associated with altruism or "warm glow" [27, 2, 3, 4], or simple fun [22]. Building on the large empirical social loafing literature in psychology, [37] propose a Collective Effort Model; empirical research on virtual communities supports their prediction that contributions increase when subjects are informed of the impact of their contribution on the aggregate outcome [45, 43, 61]. [72] and [12] find, as predicted by social identity theory [70], that people work harder for their "in-group" than for groups with which they do not identify. Social comparison theory [29] concerns different ways in which people are motivated by comparing their performance to others'. For example, comparison with superior others can motivate those seeking self-improvement [80], but in other circumstances may discourage effort [69]. Leader boards (e.g., "Top 10 contributors") are one example of comparison to superior others in virtual communities.

Stylized design ideas include:

- Providw direct *private* benefits, with the *sharing* resulting as a side-effect, or *spillover*. Apparently, the private benefits of mobile personal bookmark management are one of the main motivations responsible for `del.icio.us`'s success as a provider of spillover shared public goods [78, 60]. [33] found that users in the MovieLens virtual community rate movies to improve the recommendations they receive themselves.

- Make knowledge-sharing fun. [57] report that both fun and ideological reasons attract contributors to Wikipedia, but it is fun that causes a high level of contributions. [33] find MovieLens users rate movies for fun as well. [74] created a game in which pairs of players try to guess each other's choice of tags to the same images, which has been licensed by Google (see also `http://www.gwap.com/gwap/`).

- Structure knowledge-sharing to provide contributors with a valuable learning experience. Learning is an important reason for open source programmers to participate [62, 34].

- Design knowledge-sharing to provide social comparison feedback (e.g., [4, 66]). In the case of open source software, social comparison feedback can generate career opportunities [42].

My colleagues and I recently commenced a new set of ICD studies on the question of increasing participation in VOs. In one paper we are applying the theories of private provision of public goods and hidden information to develop a model using *exclusion* to motivate knowledge contribution and quality in a VO. We have initiated a series of lab and field experiments on question answering in VOs. We have also started to analyze field data to test hypotheses on motivations for VO members to continue or terminate their participation over time [36].

5 Designing to motivate quality

I now briefly discuss the also critical problem of inducing desirable *quality* in VO knowledge-sharing. Willingness to opt-in to a knowledge-sharing VO typically depends critically on the quality of the information available to participants. Knowledge is not a homogeneous commodity. There is no generally agreed upon way to measure its quantity, and thus no standard way to characterize "how much" knowledge a participant is contributing to a VO.[10] Though informal, it is useful to say that knowledge is distinguished by (multi-dimensional) qualities. For example, in economics, scholars distinguish between "vertical" and "horizontal" qualities. Vertical are characteristics on which everyone agrees in their preference ordering: a monotonic gradient from worst to best [56]. An example for information might be its degree of objective accuracy. Horizontal refers to characteristics on which preference orderings vary: some people prefer red, some blue, and a few chartreuse [e.g., 41].[11]

The quality problem activates at least two behavioral issues: *hidden action* (or moral hazard) and *hidden information* (or adverse selection). The first: how do we induce participants to make desired *effort* to provide quality when we cannot observe their effort? One approach studied by theorists [40, e.g.,] is to tie extrinsic rewards to measures that *are* observable, and are caused in part by the underlying hidden action. In knowledge systems, this might be done with a rating system on contributions, with the feedback tied to the contributor's reputation or another motivator.

Hidden information matters when contributors know something about the quality of their contribution that others do not, and may use this to their advantage if their motivations are not congruent. One well-known example is information *pollution*: information contributed to advance the interests of the contributor but that is outside the objectives of the VO. Spam is an example (whether in email, blogs, discussion boards, etc.) [19, 44], as is not-arm's-length recommendations (such as pseudonymous self or friend book reviews on Amazon [24]). In our new work we have a preliminary model of an ICD voting screen to limit contributions to those above a minimum quality threshold.[12]

Several social psychology findings can be applied to the design of quality motivations. For example, a VO might post a statistic about the feedback scores (e.g., the median) as a social comparison motivator (as do universities when they return teaching evaluation scores). In one of our current studies we are comparing social comparison and social norm motivators in several question-and-answer field experiments.

[10]Shannon's entropy [68] to quantify the amount of information is not a pragmatic construct for characterizing a set of specific utterances, and in particular is specific to a narrow definition of information that does not encompass what is generally meant by knowledge.

[11]Still others care deeply about the difference between chartreuse green and chartreuse yellow, see `en.wikipedia.org/wiki/Chartreuse_(color)`.

[12]I have shown that a related mechanism might work to discourage information pollution in a virtual community [48].

References Cited

[1] Anat R. Admati and Motty Perry. Joint projects without commitment. *Review of Economic Studies*, Jan 1991.

[2] James Andreoni. Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Public Economics*, 97:1447–1458, December 1989.

[3] James Andreoni. Impure altruism and donations to public goods: a theory of warm-glow giving. *Economic Journal*, 100:464–477, 1990.

[4] James Andreoni. Cooperation in public goods experiments: Kindness or confusion? *American Economic Review*, 85:891–904, 1995.

[5] P. Arzberger, P. Schroeder, A. Beaulieu, G. Bowker, K. Casey, L. Laaksonen, D. Moorman, P. Uhlir, and P. Wouters. An international framework to promote access to data. *Science*, 303(5665):1777–1778, 2004.

[6] Christopher Avery, Paul Resnick, and Richard Zeckhauser. The market for evaluations. *American Economic Review*, 89(3):564–584, June 1999. URL `http://ideas.repec.org/a/aea/aecrev/v89y1999i3p564-584.html`.

[7] Mark Bagnoli and Barton L. Lipman. Provision of public goods: Fully implementing the core through private contributions. *Review of Economic Studies*, Jan 1989.

[8] Theodore Bergstrom, Lawrence Edward Blume, and Hal Varian. On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49, January 1986. URL `http://repositories.cdlib.org/cgi/viewcontent.cgi%3Farticle%3D1024%26context%3Dpostprints`.

[9] National Science Board. Long-lived data collections: Enabling research and education in the 21st century. Technical report, National Science Foundation, Arlington, VA, 2005.

[10] Maria S. Bonn, Wendy Pradt Lougee, Jeffrey K. MacKie-Mason, and Juan F. Riveros. The peak project: A field experiment in pricing and usage of a digital collection. University of Michigan Scholarly Publishing Office, Ann Arbor, MI, 2008. URL `http://hdl.handle.net/2027/spo.5621225.0001.001`. Available online at `http://hdl.handle.net/2027/spo.5621225.0001.001`.

[11] Nathan Bos. Scientific collaboration on the internet. In Olson et al. [58], chapter Motivation to Contribute to Collaboratories: A Public Goods Approach, pages 251–274.

[12] Nathan Bos, N. Sadat Shami, Judith S. Olson, Arik Cheshin, and Ning Nan. In-group/out-group effects in distributed teams: an experimental simulation. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 429–436, New York, NY, USA, 2004. ACM Press. ISBN 1-58113-810-5. doi: http://doi.acm.org/10.1145/1031607.1031679.

[13] Nathan Bos, Ann S. Zimmerman, Judith S. Olson, Jude Yew, Jason Yerkie, Erik Dahl, Daniel Cooney, and Gary M. Olson. Scientific collaboration on the internet. In Olson et al. [58], chapter From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories, pages 53–72.

[14] C. J. Bowser. Historic data sets: Lessons from the past, lessons for the future. In W. K. Michener, editor, *Research data management in the ecological sciences*, pages 155–179. University of South Carolina Press, Columbia, 1986.

[15] Yan Chen and Xin Li. Group identity and social preference. Manuscript, 2007.

[16] Yan Chen, Xin Li, and Jeffrey K. Mackie-Mason. Online fund-raising mechanisms: A field experiment. *Contributions to Economic Analysis and Policy*, 5(2), 2006. URL `http://www.bepress.com/bejeap/contributions/vol5/iss2/art4/`.

[17] Yan Chen, Max Harper, Joseph Konstan, and Sherry Xin Li. Social comparisons and contributions to online communities: A field experiment on movielens, 2007. working paper.

[18] Yan Chen, Teck-Hua Ho, and Yong-Mi Kim. Knowledge market design: A field experiment at google answers. Manuscript, 2007.

[19] Ben Chiao and Jeffrey K. MacKie-Mason. Using uncensored communication channels to divert spam traffic. In *34th Research Conference on Communication, Information and Internet Policy (TPRC)*, 2006. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=941207.

[20] Dan Cosley, Pamela Ludford, and Loren Terveen. Studying the effect of similarity in online task-focused interactions. *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*, pages 321–329, 2003.

[21] National Research Council. Bits of power: Issues in global access to scientific data. Technical report, National Academies Press, Washington, D.C., 1997.

[22] M. Csikszentmihalyi. *Beyond Boredom and Anxiety: The Experience of Play in Work and Games*. Jossey-Bass, San Francisco, 1975.

[23] Jonathon Cummings, Thomas A. Finholt, Ian Foster, Carl Kesselman, Diana Rhoten, and Katherine A. Lawrence. Beyond virtual: A multidisciplinary agenda for designing, building, and evaluating virtual organizations. A Report from the NSF Workshop to Identify Research, Infrastructure and Organizational Needs to Accelerate Collaboration in Science and Engineering (Draft 2.0), February 2008.

[24] Chrysanthos Dellarocas. Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, 52(10):1577–1593, 2006.

[25] Nicholas Economides. The economics of networks. *International Journal of Industrial Organization*, Oct 1996.

[26] Nicholas Economides and Charles Himmelberg. Critical mass and network size with application to the us fax market. Technical report, New York University Stern School of Business, New York, 1994. URL http://www.stern.nyu.edu/networks/95-11.pdf. Discussion Paper no. EC-95-11.

[27] Ernst Fehr and Urs Fischbacher. The nature of human altruism. *Nature*, 425(6960):785–791, October 2003. doi: 10.1038/nature02043. URL http://dx.doi.org/10.1038/nature02043.

[28] Joan Feigenbaum, Christos Papadimitriou, Rahul Sami, and Scott Shenker. A BGP-based mechanism for lowest-cost routing. *Distributed Computing*, 18(1):61–72, 2005. http://theory.csail.mit.edu/~sami/papers/FPSS.pdf.

[29] Leon Festinger. A theory of social comparison. *Human Relations*, 7:117–40, 1954.

[30] Eric Friedman, Paul Resnick, and Rahul Sami. Manipulation-resistant reputation systems. In Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani, editors, *Algorithmic Game Theory*, chapter 27. Cambridge University Press, 2007.

[31] Resnick Paul Friedman Eric J. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199, June 2001. URL http://ideas.repec.org/a/bla/jemstr/v10y2001i2p173-199.html.

[32] Theodore Groves and John Ledyard. Optimal allocation of public goods: A solution to the “free rider” problem. *Econometrica*, 45(4):783–809, May 1977. URL http://links.jstor.org/sici?sici=0012-9682%28197705%2945%3A4%3C783%3AOAOPGA%3E2.0.CO%3B2-8.

[33] Maxwell Harper, Xin Li, Yan Chen, and Joseph Konstan. An economic model of user rating in an online recommender system. In *User Modeling: Proceedings of the Tenth International Conference*, 2005. URL http://www.si.umich.edu/~yanchen/papers/um-paper-vfin.pdf.

[34] Alexander Hars. Working for free? motivations for participating in open-source projects. *International Journal of Electronic Commerce*, 6(3):25–39, Spring 2002.

[35] Lian Jian and Jeffrey MacKie-Mason. Why share in peer-to-peer networks? In *Proceedings of Workshop on the Economics of Networked Systems (NetEcon06)*, Ann Arbor, MI, June 2006. URL `http://www.cs.duke.edu/nicl/netecon06/`. Available from `http://www.cs.duke.edu/nicl/netecon06/`.

[36] Lian Jian and Jeffrey K. MacKie-Mason. Why leave wikipedia? In *iConference*, UCLA, February 2008. Extended abstract and poster.

[37] S. J. Karau and K. D. Williams. Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, pages 681–706, 1993. URL `http://web.ebscohost.com.proxy.lib.umich.edu/ehost/pdf?vid=3&hid=102&sid=413cab98-a869-4809-ace5-09bd1b320c90%40sessionmgr106`.

[38] Michael L. Katz and Carl Shapiro. Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–440, 1985. URL `http://ideas.repec.org/a/aea/aecrev/v75y1985i3p424-40.html`.

[39] Daria Knoch, Alvaro Pascual-Leone, Kaspar Meyer, Valerie Treyer, and Ernst Fehr. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, Jan 2006. URL `http://www.sciencemag.org/cgi/content/abstract/314/5800/829/`.

[40] Jean-Jacques Laffont and David Martimort. *The Theory of Incentives*. Princeton University Press, 2001.

[41] Kelvin Lancaster. A new approach to consumer theory. *The Journal of Political Economy*, 74(2):132–157, 1966.

[42] Joshua Lerner and Jean Tirole. Some simple economics of open source. *Journal of Industrial Economics*, 50(2), June 2002. URL `http://ideas.repec.org/a/bla/jindec/v50y2002i2p197-234.html`.

[43] K. Ling, G. Beenen, P. Ludford, X. Wang, K. Chang, D. Cosley, D. Frankowski, L. Terveen, A. M. Rashid, P. Resnick, and R Kraut. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, 10(4), 2005. URL `http://jcmc.indiana.edu/vol10/issue4/ling.html`.

[44] Thede Loder, Marshall van Alstyne, and Rick Wash. An economic response to unsolicited communication. *Advances in Economic Analysis and Policy*, 6(1), 2006. URL `http://www.bepress.com/bejeap/advances/vol6/iss1/art2/`.

[45] Pamela J. Ludford, Dan Cosley, Dan Frankowski, and Loren Terveen. Think different: increasing online community participation using uniqueness and group dissimilarity. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 631–638, New York, NY, USA, 2004. ACM Press. URL `http://doi.acm.org/10.1145/985692.985772`.

[46] Jeff Mackie-Mason, Scott Shenker, and Hal Varian. Service architecture and content provision: The network provider as editor. *Telecommunications Policy*, 20(3), April 1996. URL `http://ideas.repec.org/p/wpa/wuwpmi/9601002.html`.

[47] Jeffery K. MacKie-Mason, Juan F. Riveros, and Robert S. Gazzale. Pricing and bundling electronic information goods: Experimental evidence. In B. Compaine and I. Vogelsang, editors, *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy*. MIT Press, 2000.

[48] Jeffrey K. MacKie-Mason. Tom sawyer production on the internet: Getting the good stuff in, keeping the bad stuff out. Invited Plenary Lecture at the Southern Economic Association Meetings. Current version available from author., November 2007.

[49] Jeffrey K. MacKie-Mason and Juan Riveros. Economics and electronic access to scholarly information. In B. Kahin and H. Varian, editors, *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*. MIT Press, Cambridge, Massachusetts, 2000.

[50] Jeffrey K. MacKie-Mason and Hal R. Varian. Pricing congestible network resources. *IEEE Journal on Selected Areas of Communications*, 13(7):1141–1149, 1995. Available from http://www-personal.umich.edu/~jmm/papers/Pricing{_}Congestible{_}Resources.pdf.

[51] Jeffrey K. MacKie-Mason and Hal R. Varian. Pricing the internet. In Brian Kahin and James Keller, editors, *Public Access to the Internet*, pages 268–314. MIT Press, Cambridge, MA, 1995. URL http://ideas.repec.org/p/wpa/wuwpco/9401002.html.

[52] Jeffrey K. MacKie-Mason and Michael P. Wellman. Computational markets. In Leigh Tesfatsion and Kenneth L. Judd, editors, *Handbook of Computational Economics*, volume 2, chapter 28, pages 1382–1431. Elsevier, Amsterdam, 2006. Available from http://www-personal.umich.edu/~jmm/papers/handbook4apr05.pdf.

[53] Jeffrey K. MacKie-Mason, Anna Osepayshvili, Daniel M. Reeves, and Michael P. Wellman. Price prediction strategies for market-based scheduling. In *Fourteenth International Conference on Automated Planning and Scheduling*, pages 244–252, Whistler, BC, 2004. Available from http://www-personal.umich.edu/~jmm/papers/dexter-icaps04.pdf.

[54] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting honest feedback: The peer prediction method. *Management Science*, 51(9):1359–1373, September 2005.

[55] John Morgan. Financing public goods by means of lotteries. *Review of Economic Studies*, 67(4), 2000. doi: 10.1111/1467-937X.00153.

[56] Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18(2): 301–317, August 1978. URL http://ideas.repec.org/a/eee/jetheo/v18y1978i2p301-317.html.

[57] Oded Nov. What motivates wikipedians, 2007. URL http://opensource.mit.edu/papers/Nov_Wikipedia_motivations_opensource.mit.edu.pdf. Accepted for publication in the Communications of the ACM.

[58] Gary M. Olson, Ann S. Zimmerman, and Nathan Bos, editors. *Scientific Collaboration on the Internet*. MIT Press, Cambridge, MA, Forthcoming.

[59] Anna Osepayshvili, Michael P. Wellman, Daniel Reeves, and Jeffrey K. MacKie-Mason. Self-confirming price prediction for bidding in simultaneous ascending auctions. Working paper, University of Michigan, 2005.

[60] Emilee Rader and Rick Wash. Tagging with del.icio.us: Social or selfish? In *Computer Support Cooperative Work*, 2006. Poster Session.

[61] Al M. Rashid, Kimberly Ling, Regina D. Tassone, Paul Resnick, Robert Kraut, and John Riedl. Motivating participation by displaying the value of contribution. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 955–958, New York, NY, USA, 2006. ACM Press. URL http://doi.acm.org/10.1145/1124772.1124915.

[62] Eric S. Raymond. *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary (O'Reilly Linux)*. O'Reilly, October 1999.

[63] Daniel M. Reeves, Michael P. Wellman, Jeffrey K. MacKie-Mason, and Anna Osepayshvili. Exploring bidding strategies for market-based scheduling. *Decision Support Systems*, 39:67–85, 2005.

[64] Paul Resnick and Rahul Sami. The influence-limiter: Provably manipulation-resistant recommender systems. In *Recommender Systems*. ACM, 2007. URL http://www.umich.edu/~rsami/papers/recsys.pdf.

[65] Paul Resnick and Rahul Sami. The information cost of manipulation resistance in recommender systems. In *ACM Electronic Commerce (EC '08)*, Chicago, IL, July 2008. ACM.

[66] Richard Ryan and Edward L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78, 2000.

[67] Paul A Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4): 387–389, November 1954.

[68] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27: 379–423, 1948.

[69] J. Suls, R. Martin, and L. Wheeler. Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, 11(5):159–163, 2002. URL `http://dx.doi.org/10.1111/1467-8721.00191`.

[70] Henri Tajfel and John Turner. An integrative theory of intergroup conflict. In Stephen Worchel and William Austin, editors, *The Social Psychology of Intergroup Relations*. Brooks/Cole, Monterey, CA, 1979.

[71] Kan Takeuchi, John C. Lin, Yan Chen, and Thomas A. Finholt. Scheduling with package auctions. University of Michigan Manuscript, 2008.

[72] Mark Van Vugt and Claire M. Hart. Social identity as social glue: the origins of group loyalty. *Journal of Personality and Social Psychology*, 86:585–98, 2004.

[73] Hal R. Varian. *Microeconomic Analysis*. W. W. Norton & Company, 1992.

[74] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the ACM CHI, 2004*, 2004. URL `http://citeseer.ist.psu.edu/vonahn04labeling.html`.

[75] Mark Walker. A simple incentive compatible scheme for attaining lindahl allocations. *Econometrica*, Jan 1981. URL `http://ideas.repec.org/a/ecm/emetrp/v49y1981i1p65-71.html`.

[76] Rick Wash and Jeffrey K. MacKie-Mason. Incentive-centered design for information security. In *USENIX Hot Topics in Security (HotSec 06)*, Vancouver, BC, 2006. USENIX. URL `http://www.usenix.org/events/hotsec06/tech/wash.html`.

[77] Rick Wash and Jeffrey K. MacKie-Mason. Security when people matter: Structuring incentives for user behavior. In *International Conference on Electronic Commerce (ICEC-07)*, Minneapolis, MN, August 2007. ACM.

[78] Rick Wash and Emilee Rader. Public bookmarks and private benefits: An analysis of incentives in social computing. In *American Society for Information Science & Technology*, 2007.

[79] Michael P. Wellman, William E. Walsh, Peter R. Wurman, and Jeffrey K. MacKie-Mason. Auction protocols for decentralized scheduling. *Games and Economic Behavior*, 35:271–303, 2001.

[80] Joanne V. Wood. Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, 106:231–248, 1989.

Penn Law
UNIVERSITY of PENNSYLVANIA LAW SCHOOL

Christopher S. Yoo
Professor of Law and Communication
Founding Director, Center on Technology, Innovation, & Competition Policy

3400 Chestnut Street
Philadelphia, PA 19104-6204
Tel 215.746.8772 Fax 215.573.2025
csyoo@law.upenn.edu

GENI WHITE PAPER

Christopher S. Yoo

University of Pennsylvania

May 31, 2008

GENI holds the promise of being a landmark in promoting the migration to the next generation network technology. The idea of providing a laboratory for experimenting with different architectures with real-world end users has tremendous potential. I am grateful to have the opportunity to be a part of it.

Attracting end users to GENI represents a critical element of the project. Without the ends users, the experiment will fail. It is also the aspect of the project over which the project participants have the least control.

Marketers have long recognized the key role that early adopters play in facilitating the adoption of a new technology. These individuals are often among the most interested in new technologies. They are also typically quite technologically savvy and generally more capable of managing new technologies that often contain unresolved technological issues and receive less than normal technical support. Most importantly, these early adopters are leading opinion makers who exert a strong influence over the decisions of subsequent consumers.

The GENI project would thus be well served if it developed an organized strategy to identify and reach out to these early adopters. One approach might be to target key populations, such as college students, who have the time, interest, access to computing and network resources, and the technical support to manage interfaces with new systems. Another model for such an outreach program is the marketing program for the recently launched Wii gaming platform. As part of its entry strategy, Sony identified key individuals in communities across the country likely to serve as early adopters and focal points for information.

That said, for GENI, the key opinion makers are more likely to be different. To the extent that a particular application on GENI is a direct replacement for an older technology, powers users of the previous technology are logical candidates. The GENI staff should consider entering into discussions with Sony and other corporations pursuing similar strategies to see if the strategies they employed can be used to recruit end users. In addition, the GENI project should consider sponsoring research into the technology adoption process. Other key opinion makers may include bloggers or other information intermediaries. The GENI staff should consider devising a strategy and expending resources to foster relationships with key members of the blogging community. The public service oriented aspects of the project are likely to make them receptive to the project's overall goals.

In addition, the GENI project should develop an infrastructure for communicating with the end user community on an ongoing basis, to inform them of solutions to technological problems and to notify them of new product offerings. As everyone knows, information that is not well targeted at the person receiving the information will not read it and will cause considerable frustration. It would thus be worth sponsoring research into the most effective means of communication

In addition, change management scholarship has often emphasized the importance of achieving a few early successes. GENI project participants should thus choose the initial projects with care. Strong candidates would include applications that simply replace other applications already in use on a one-for-one basis. Applications that require educating consumers about new functionalities about which the consumer does not already know or require significant changes in habits are attractive candidates for early success.

The GENI project must also carefully monitor a number of legal issues that may arise with respect to end user opt in. In terms of privacy, it is inevitable in such a program that end user activities will generate end user information that GENI participants will use to improve their products and services. It is important that GENI participants provide end users with full disclosure about the information that is being capture and how it will be used.

Every end user should be required to enter into an agreement on certain matters. Although it is not completely clear whether the legal issues discussed have implications primarily for firms offering experimental services through GENI or for the GENI project itself, it would seem prudent for the GENI project to protect its interests through some for of end user agreement.

For example, as is typically the case with experimental technologies, the agreement should include a specific disclaimer of warranties, specifying that the products and services are provided on an "as is" basis without any warranty of merchantability or fitness for any particular purpose.

The agreement should contain a limitation of liability, disclaiming any responsibility for incidental and consequential damages, including lost revenues and profits.

The agreement should specify that all GENI participants agree to keep all trade secrets confidential unless given written permission. It should specify language to ensure that GENI usage does not affect the patentability of any inventions by causing the on-sale bar to run.

The agreement should include explicit terms about the extent to which GENI participants can publicize activities and the extent to which they have obligations to maintain confidentiality.

The agreement should probably contain an arbitration clause, agreeing that any disputes under it should be resolved through arbitration. This provision may require a higher level of assent than other provisions.

Judicial decisions suggest that this agreement can be undertaken through a clickwrap license, although contrary decisions exist as well. There may be some provisions, such as the arbitration clause noted above, that may require a higher level of assent.

Comments on Getting Real World Users into GENI

David C. Parkes
Harvard University
parkes@eecs.harvard.edu

June 24, 2008

To meet its call of facilitating transformative networks research with an eye on a "revolutionary future Internet" [1], GENI will need to embrace and support rich real world applications.

Echoing in part the useful discussion in this GENI "research plan" document, it is apparent that this will require at least:

- GENI facilities on which third parties can develop and deploy applications
- Enough stability in candidate network architectures that are tested on GENI to make development of applications worthwhile
- Connections to the today's Internet: end users are not on GENI, they are on the Internet
- Strategic resource allocation: a clear policy and mandate for allocating GENI resources to competing architectures, and from architectures to applications.
- Is it technologically plausible to conceive of application developers being able to "purchase" the competing services on underlying network architectures? Could a market be developed to facilitate innovation and learning, with special rates for spending on new services?
- Can we in turn allow these network architectures to compete for GENI resources, but without "locking-out" other upcoming and in development architectures?
- Can for-profit/not-for-profit partners be developed in strategic areas including
    1. Distributed Virtual Gaming Environments
    2. Peer-to-Peer Television
    3. Distributed Back-up
- A long-term plan that supports the commercialization of real-world applications [e.g., will there be opportunities to invest in a "fenced in" part of GENI for use by a dedicated application, or shared amongst a group of investors?]
- Can market-forces (even w/ virtual currencies) be used to encourage innovation amongst application developers in advance of mass commercialization, perhaps with votes etc. from end users, in part as an effort to prevent a GENI office running a "beauty-contest" in choosing winners and losers
- Basic challenge: explore and exploit... a basic research question seems to exist in terms of the best way to manage the strategic allocation of GENI resources. Related to area of "two-sided markets."

References

[1] David D. Clark and Scott Shenker (Chairs). GENI research plan. Technical Report GENI Design Document 06-28, Research Coordination Working Group, 2007.

GENI: Privacy and User Opt-in

White Paper

Alessandro Acquisti
July 2008

Preliminary draft. Not for general distribution.

This white paper delineates privacy concerns and opportunities that may arise in the development and deployment of a Global Environment for Network Innovations (GENI), with particular attention to issues that may affect GENI users' opt-in. This document is based on the vision for the GENI project which can be inferred from version 4.5 of the GENI Research Plan (dated April 23, 2007).

The GENI project aspires to provide scientists and industry players with a facility where network experiments can be ran and evaluated, and research on an enhanced future Internet can be supported. The scope and breadth of the project (as well as those of the experiments it would make possible) could rise significant privacy issues but also create new opportunities in the field of privacy research. This document discusses some of those issues and opportunities, focusing exclusively on the *informational privacy of GENI end-users* (such as the individuals who would engage in GENI services and participate in its experiments).

In what follows, I discuss: 1) potential privacy issues associated with GENI services and experiments; 2) strategies for fair and convincing communication of GENI privacy matters to potential end-users; and 3) opportunities GENI may offer for novel privacy solutions.

1) Privacy issues

Experiments in which end users' behavior (and its impact on GENI services and applications) is monitored at the individual or aggregate level are at the core of the GENI initiative. Consequently, a trade-off arises between researchers' interests in granular information flows and end users' privacy rights.

Very little of what one does on the Internet is expected to remain private – emails sent or received, pages visited or files downloaded, and so forth. However, given the distributed nature of the Internet and – in smaller measure - the legal or contractual protections afforded to individual data by particular entities (such as an ISP), it is difficult for a third party to monitor end-to-end Internet transactions of diverse types of a particular set of users. On the other hand, while the future topology of the GENI network and the mechanisms of the experiments therein designed are still malleable, scenarios where researchers will gain direct access to vast amounts of usage and personal data by individual users are reasonable to forecast. Depending on the extent of end users participation, GENI systems development, and experiments design, the GENI researcher (individually, or as a community) may therefore assume a position similar to the *global passive* and *active attacker* that privacy researchers have long studied. This raises considerable issues. For instance, global sensing architectures and ubiquitous health care services are explicitly listed in the 2007 Research Plan as GENI's possible areas of applications. Such applications are powerful examples of the breadth and sensitivity of potentially identifiable information flows that may be generated by GENI end users participating in the network. They also highlight how end users of the network may not be necessarily aware of the extent of the monitoring made possible by the GENI initiative.

As a prerequisite to properly communicating to potential opt in end users the extent to which their personal data may be used during their participation in the network, GENI may first embark in a privacy impact assessment (PIA) and establish policies and guidelines answering the following questions (in this regard, the GENI initiative may consider the lessons learnt in the field of ubitquous ubicomp privacy and location privacy[1]):

- What amount of personally identifying information will be collected about end users and shared with GENI researchers at the moment they opt-in the GENI portal? What usage and behavioral data (such as the reactions to an experimental treatment) may be monitored and gathered? What policy protocols (such as default de-identification) and specific technological solutions (such as k-anonymity) may be enforced to protect individual data (as well as any aggregate data which may carry a risk of individual re-identification)?
- At what level of granularity will end users' registration and usage data be made available to researchers? Will experimental data be exchanged across researchers or

[1] See, e.g., http://www.cs.berkeley.edu/~jfc/ubicomp-privacy2004/.

shared across experiments? If so, will it be at an *ad hoc* basis or from a universally accessible database? What forms of access controls to personal data will be established for the GENI research community? Will granular or aggregate data be made available outside of GENI?

– What common processes will be put in place for the collection, analysis, storage, and further communication of individual data? Specifically, will all GENI end users (as well as participating researchers and industry players) adhere to one single, overarching privacy policy upon joining the GENI network - or will special policies have to be drafted depending on the service, the experiment, the application, and its participants? If so, what coordination will take place across services/applications/experiments and their diverse policies so not to obstruct the benefits of information sharing across researchers? But also: how will GENI policies coordinate with national legislations and with participant Colleges' IRB boards? Will end users' consent be required for each specific experiment?
– …

2) Communicating privacy

As noted above, several policy and design decisions may need to be taken before a coherent picture of the network's privacy implications can be offered to GENI end users. Given such complex privacy challenges, two problems arise in regard to the strategy for communicating with potential end users about the initiative: a) how to communicate fairly to potential end-users the privacy implications of participating in GENI; b) how to do so without jeopardizing – in fact, incentivizing – users' opt-in.

a) Attention should be paid to explaining to end users how participating in GENI is similar and how it is different – technically and legally – from using "traditional" Internet applications. Since some users may not realize the extent of personal data they may reveal through their usage of GENI systems (especially with embedded sensors or other ubiquitous devices, the user may not even know when some information is being communicated), particular care should be afforded to the consideration of cognitive and behavioral fallacies human exhibits in matters involving personal information security and privacy (from problems associated with incomplete information to bounded rationality[2]).

b) On the other hand, depending on how GENI will be marketed to end users, some potential participants may end up fearing the pervasive access to user data that may make possible. If proper safeguards have been put in place during the development of GENI architecture and its privacy polic(ies), however, openness and straightforwardness with end users may be the optimal strategies in both the short and long run. The history of the Internet is full of privacy debacles that show how end users dislike *less* a company that is straightforward about its intentions of monitoring and using consumer data, than a company that does so sneakily or clumsily. The latter is often punished in the marketplace, while the former can thrive. In some cases, the incentives necessary to attract users and convince them to provide large amounts of personal information are almost worringly small. In fact, the experience of companies such as comScore (and, in particular, its Media Metrix products) may offer valuable lessons for the opt in of GENI end users. One challenge, however, will be represented by the inherent tension (a sort of chicken and egg problem) between the positive network externalities that end users will derive in case of mass adoption of GENI applications (and which will counter and soothe privacy concerns), versus the need to address potential privacy concerns in first instance, when strong externalities may not yet be apparent to end users, in order to win early adopters and start growing the network.

3) Opportunities

GENI also offers unprecedented opportunities for privacy researchers. Some of those opportunities may even be leveraged as incentives and act as drivers of end users' participation.

From a *normative* standpoint, privacy is notoriously hard to study in laboratory experiments: the need to avoid deception in experimental settings may conflict with the need to avoid priming subjects about the scope of the experiment (once subjects know that the focus of a study is privacy, their privacy relevant behavior is affected); in addition, realistic privacy incentives and trade-offs are hard to replicate in the lab. However, precisely because of the scope of individual

[2] See, e.g., Alessandro Acquisti, Privacy in electronic commerce and the economics of immediate gratification, ACM EC, (2004).

monitoring made possible by future GENI systems and applications, researchers may be able to run realistic experiments in the wild which may cast a better light on privacy-sensitive individual behavior than lab experiments.

From a *positive* standpoint, GENI may offer a sandbox for the deployment and test of novel privacy technologies. In particular, in recent years significant attention has been directed towards the development of protocols which protect the privacy (or even the anonymity) of their participants, while ensuring accountability[3] (unaccountability through anonymity can sometimes increase incentives for sub-optimal social behavior[4]). Rather than succumbing to the temptation of making a new-generation Internet secure by requiring the equivalent of "passports" and ID cards from its users for each Internet transaction, such research could balance security and confidentiality goals. For instance, participation in the GENI network and its applications may by default imply the adoption – in addition to strong authentication and access controls - of encrypted mail communications, of TOR-like protected transactions, and of applications that are designed to take into consideration end users' natural cognitive and behavioral difficulties in the handling of security technologies.[5] Such offering of a secure, private, *and* usable environment could counterbalance possible end users' concerns about and act as incentives for end users' opt in.[6]

---

[3] See, e.g., Roger Dingledine, Michael J. Freedman, David Hopwood, David Molnar, A Reputation System to Increase MIX-net Reliability, *Proceedings of the 4th International Workshop on Information Hiding*, (2001).

[4] See, e.g., Eric Friedman and Paul Resnick, The Social Cost of Cheap Pseudonyms, *Journal of Economics and Management Strategy*, 10(2): 173-199, (2001).

[5] See, e.g., Alma Whitten and J. D. Tygar, Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0, *USENIX*, (1999).

[6] See, e.g., Janice Tsai, Serge Egelman, Lorrie Cranor, Alessandro Acquisti , The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, WEIS, (2007).

GENI and User Opt-In

Rahul Telang
The Heinz School, Carnegie Mellon University

GENI (Global Environment for Network Innovations) is being envisioned as “a national facility to explore radical designs for a future global networking infrastructure”. The goal of this facility is to encourage and promote researchers to use the facility to create prototypes for the path breaking ideas. In particular, it is expected to attract those projects which cannot be easily implemented using current Internet and network architecture and have significant potential. Thus the facility is a bridge between an idea and its eventual commercial roll-out (if successful).

Beyond the technical feasibility of GENI, some key challenges would need to be resolved to make GENI successful. These challenges include attracting users to use these facilities, delineating policies for using the facility, monitoring end user behavior, outlining privacy policies etc. The first issue (user participation) is of considerable importance. Without significant user acceptance, the facility is unlike to generate the social benefits it is envisioned to. In this white paper, I focus on this issue and discuss various strategies to broaden user opt-in.

Users: It is important to distinguish between GENI users. There are two distinct constituents who are expected to use this environment. The first sets of users are the researchers who will explore their ideas, build the prototypes, and run the experiments which would sometimes involve real end user participation. The second sets of users are the “end” users who will participate in the experiments run on these facilities. The experiments will allow end users to perform certain tasks. The success of the experiment will depend upon demonstrating how well the application is likely to perform (scalability, reliability, security etc) in a real world setting.

While there are obvious differences between these segments, and the focus of this workshop is presumably on the latter segment, acceptance by the former segment is equally imperative. As I will outline below, GENI would need to attract enough researchers to offer interesting applications and experiments. Without researchers’ participation, end users are unlikely to opt-in as well. I will continue to distinguish between the two in the following.

There is a significant body of work in economics, Information Systems and in behavioral economics on what leads to acceptance, use and diffusion of new technology (Hall and Khan 2003). Tangible and intangible costs and benefit, user ability and their attitudes, user perception, network effects, all play an important role in users’ adopting a platform or technology. At a conceptual level, the user opt-in in GENI depends on whether users expect to receive a positive utility from adoption. User utility depends on the (perceived) expected benefits and (perceived) expected costs. Unless, the benefits outweigh the costs significantly, we are unlikely to see significant adoption. Designers of GENI can affect both these elements. Let me start with the cost.

A significant cost for the researchers (and end users) using this facility is the “learning” cost. While a researcher is likely to be technically savvy, it will undoubtedly be a challenging proposition for a beginner to learn to use the systems. Before an experiment can be set up a researcher will have to get his/her proposal cleared, request appropriate systems element (bandwidth, devices, software/hardware), configure the systems to his/her specifications, set up the experiments by uploading software code, invite and attract end users (if needed), monitor, modify and scale the experiment. All these steps would require incur a significant fixed learning cost to get off the ground, and continual learning cost as the experiment progresses.

A key challenge for GENI designers is to lower these costs and/or increase the perceived benefits as much as possible. Clear and precise documentation, hands-on tutorials, transparency of operations, hiding unnecessary details from users is critically important. The following seem to be relevant strategies for lowering these costs -

Standardization: Adoption and rate of adoption will be affected by user interface and processes. It is expected that GENI will allow diverse sets of application to be run on its facilities. However, requesting resources on the systems, inviting real users to participate in experiments, interaction between end users, end users and experiments, and end users and researchers need to be standardized. Thus a researcher (and an end user) should be able to interact with the facility in a standardize fashion independent of application, users, and systems. This is useful for two reasons. First, once a researcher and end user learns to use the system, additional use will require little or no learning. Second, standardization would allow for end users to participate in multiple experiments or a researcher to explore multiple projects, or end users and researchers to share their knowledge with each other. Standardization will presumably also reduce the cost for GENI by streamlining the processes.

“How-to” help: Despite documentation and tutorials, implementing an experiment in GENI would be a complex task. Each implementation would bring its own set of unique problems. Some of these questions cannot be anticipated and explained in documentation. These real-time questions would need quick resolution. GENI would have to plan for and provide this help to researchers and to end users. One obvious possibility is to have extensive FAQs on GENI pages, provide an email address which is monitored by technically savvy staff at GENI which can resolve these issues as quickly as possible.

Another attractive option is allow other researchers (and end users) to contribute and respond to questions and clarifications. Community participation is an efficient way to manage user questions and many firms have harnesses this very effectively (for example Dell computers). Managed effectively, such participation can lead to quicker, real time, and better response which is disseminated rapidly. It also leads to a sense of community participation and higher level of loyalty. Presumably, it would reduce the costs for GENI too.

Plug-and-Play: Of particular importance to end users will be the ease of use. Most of the end users may or may not be technically savvy users. Moreover, most users are accustomed to plug-and-play nature of Internet and software. Whether it is peer-to-peer network, or making phone calls on Internet, or Internet chat rooms, most users simply expect to either download a software client or use the Internet browsers to participate in such activities. If researchers (and GENI) want to attract real end users to participate in experiments, the users interface has to be easy to use and have plug-and-play properties.

The costs that I note are of significant relevance once a researcher and users are ready for participation. However, a big challenge for any platform is to attract enough participants in the first place. In short, some stimulus to "kick start the network" is needed before sufficient participation level is reached that the platform becomes self-sufficient.

Identify the leaders: In the initial phase, GENI needs to identify a key group of researchers in universities or industry who would have interest in taking their projects to GENI environment. GENI would have to work with them closely, gather their inputs, and provide some handholding to launch these projects. While GENI offers tremendous potential, the environment is still uncertain. It is not easy to foresee which applications or experiments are likely to be more successful, or what sort of technical or non-technical difficulties are likely to arise. Due to the uncertain and dynamic nature of the platform, working closely with particular experiments in the initial stages will provide the necessary feedback to manage the operations, foresee the challenges, and ensure that initial experiments are carried out successfully. It will also provide necessary template for future projects. Thus, these projects can act as prototypes for future researchers and end users about what can be accomplished in GENI. Without clarity in understanding the feasibility and potential benefits of GENI, a researcher will be reluctant to invest the efforts to get his/her experiments on the platform.

Learn from other Initiatives: While GENI offers unique potential, similar initiatives have been carried out in the past. Internet2 is one such initiative (though with significantly fewer features).Internet2 also offers opportunities for researchers to use those facilities to carry out their experiments. Already, researchers are carrying out experiments on Internet2. Their inputs and feedback along with the analysis of problems in running these projects can provide valuable insights into making GENI successful.

A lot can also be learned by start-up firms who develop new applications. The success of these products critically depends on attracting end users. In that regard, an application running on GENI is like a beta version of the applications like Skype, or BitTorrent. These applications are successful because they offer clear value to its users, they are able attract a loyal user base, and they have encouraged its users to build and provide third party applications which make these applications even more valuable. In the same vein, on GENI, researchers will have to bear some responsibility in conveying the value of its application to end users. GENI can provide the tools to make it easy for the end users to use them but unless end users see a significant value, some applications are unlikely to be successful. In fact, one can easily imagine the heterogeneity in the ability of different experiments running on GENI to attract end users. Some of them would lead to success while other may not.

This is where the standardization and network externalities play important role. If the end user interface is standardized across applications, it will be relatively easier to attract end users from one experiment into the other.

Involve the educators: Another attractive option to broaden GENI opt-in is to involve educators across the universities. A large number of universities and educators offer courses in networks, security and telecommunication that require hand-on student participation. If GENI can provide the opportunities for the educators to set up lab experiments on its facilities, it can increase GENI's reach to a coveted user group.

Many courses already use software simulators like OpNET or computer labs within the institutions to teach these concepts. GENI offers a significantly better environment, with a lot of flexibility and real time nature to these projects. The end users (typically students) are the appropriate target. As they get more familiar with GENI during their college years, they are more likely to be end user participants (or even future researchers as graduate students and faculty) later. While offering education opportunities may not provide direct research benefits, it is likely to have far large participation impact. The students can also be encouraged to be participants in other research experiments. Many social science experiments indeed are conducted with student audience.

To this end, GENI would need to provide appropriate templates and tutorials, and provide some examples of how GENI may be used. It would need to attract and incentivize educators and text book writers to write their own lab tutorials, and encourage widespread dissemination.

Key Challenges

Much of the discussion above focuses on operational issues when the experiments are up and running. However, there are many challenging questions regarding who can use these facilities (end users as well as researchers), how their activities be monitored, and what the privacy implications are.

First is the process to determine who can use GENI facilities. What is the threshold for allowing an experiment to be run on these facilities? A transparent methodology to determine the eligibility is important. Once the experiments are up and running and end users are participating, how will the activities be monitored? What would constitute acceptable use? A researcher experimenting with peer-to-peer network may attract end users who are sharing copy-righted material, or socially unacceptable material (like pornography). Clear guidelines would have to be imposed on acceptable use. What would be opt-out policies? How will privacy of end users and even researchers be maintained?

Research Opportunities:

This is beyond the scope of this workshop, but GENI environment offer interesting opportunities for social science research. User adoption, use, and diffusion of technology have been an active area of research for social scientists. Over time, as GENI use increases, different data sets can be created which include the usage data, experiment success and failure data, user attitudes data etc. These datasets will provide researchers a fertile avenue to examine various issues highlighted above.

References:

Bronwyn H. Hall and Beethika Khan (2003), “Adoption of New Technology”, *Institute of Business and Economic Research,* Paper E03-330.