Computational Support for Global Policy Making

Hannah Thoreson (squareforceone@gmail.com)
William Abbott Foster, PhD (william.foster@inta.gatech.edu)
Center for International Strategy, Technology, and Policy (CISTP)
Sam Nunn School of International Affairs
Georgia Institute of Technology
July 23, 2013

Introduction

Over the past thirty years, decision-makers have started to use a wide variety of computational tools to support the policy process both nationally and globally. To our knowledge no one has attempted to document what disciplines these systems have come out of nor have there been successful attempts to quantify the effective impact on society of these systems. If we are to design better computational tools in support of policy-making, it is valuable to review the history of such systems.

Though there is wide debate about whether it is helpful to use the term "collective intelligence" to describe computationally supported policymaking, collective intelligence is the wisdom of indityiduals acting as a group. IBM describes it as " the aggregated knowledge, insight and expertise of a diverse group" (http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-collective-intelligence.html). This is something that academia, IT, and business have studied for decades, but that has recently come to be referred to by buzzwords like "crowdsourcing". However, the merits of collective intelligence can no longer be disputed, as it has been made a part of the systems of some of the most powerful and deeply entrenched organizations in society.

One of the goals of collective intelligence is to raise the level of discourse beyond that of the lowest common denominator and to take advantage of the knowledge the group members have. MIT's Center for Collective Intelligence sums it up thusly: "*How can people and computers be connected so that—collectively—they act more intelligently than any individuals, groups, or computers have ever done before?'"* *http://cci.mit.edu/*

Many well-known and easily accessible websites now take advantage of this logic for consumers, but the path to highly-intelligent systems like Google and LinkedIn started with arcane systems used primarily by large corporations or government and the military. This began in the 1970's with projects like EIES (the Electronic Information Exchange System) and has continued through the present day for organizations with a need to focus on strategy and respond instantaneously respond to the huge influx of data that is created by modern life.

This project has made an effort to chronicle some of the most influential collective intelligence projects that have played a role in policymaking over the last 40 years. These include:

*List of Projects

EIES

EIES was one of the first collective intelligence projects. It came online at the New Jersey Institute of Technology in 1976 under the direction of Dr. Murray Turoff. The goal behind it was to raise the level output from groups to that of of the whole intelligence of the group rather than allowing discourse to sink to the level of the lowest common denominator.

In its early years, it was used by many large corporations as well as organizations belonging to the U.S. government. Companies like 3M, Exxon, and IBM used the powers of EIES to help maintain their economic status, while NASA, the U.S. Army, and the U.S. Department of Commerce used the system for strategic purposes.

EIES still exists today, though presumably looks much different than the system pictured below.

The WeLL

The WeLL is considered to be “the birthplace of the online community movement”. It has been described as “a precursor of every online business from Amazon.com to eBay”.

The WeLL was structured similarly to modern online forums or bulletin boards, but ran through a complex piece of Unix-based conferencing software called Picospan at the outset. It was a true community, spawning massive online discussions as well as offline social activity.

Salon owned The WeLL for a time as a paid online community, which was later purchased by its members (similar to a cooperative).

The WeLL's origins lie in another system studied by this project, EIES. Its founders were Larry Brilliant, who helped to eradicate smallpox, and Stuart Brand, a notable intellectual. "The Well was born in the context of the techno/hippie culture of the time. ... It was a hi-tech culture with open, counterculture values."

This is what made it a valuable source of collective intelligence for those who joined. For only $2 a month, members had access to the minds of a highly-educated and somewhat socially isolated group of people who were eager to share their knowledge on a wide variety of topics, from providing expert-level advice on arcane medical or technical subjects to simply being a source of emotional support and camaraderie.

See First-person source on The WeLL (email)

Global Business Network

Global Business Network was a human source of collective intelligence that was supported by a one of the first social computer networks that businesses could pay a consulting fee to access. It had many very well known and experienced members who would provide advice to corporate clients. These ranged from artists and musicians to some of the world's brightest science and technology minds. GBN's clients included large multinational corporations such as IBM or AT&T as well as organizations such as the Pentagon or the National Education Association. Later GBN dedicated itself to applying the principles of scenario planning to predict how future situations would impact the business models of various companies.

GBN reflects more the influence of government and policy on business than the reverse. Many of GBN's members came out of government, academia, or nonprofits. The concept of scenario planning originated in the military. Governments are constantly preparing for divergent futures that may or may not ever occur, but the idea of translating that concept into something the business world can profit off of is a relatively new concept. GBN would help businesses to prepare for very large shifts in thinking, such as what may have happened if the Y2K bug caused a catastrophe, or for fluctuations in the rate of economic growth.

NanoHUB.org

NanoHUB.org is a social networking website for scientifically-minded individuals to study nanotechnology and related concepts in a collaborative environment. Its thousands of users span numerous countries and academic institutions. It is a unique resource because previously, each university or research organization had their own website and their own network. NanoHUB.org makes a lot of course material and research available online for students and researchers in the field of nanotechnology.

It also has numerous tools that enable researchers to collaborate and avoid duplicating one another's efforts. There are virtual Linux workspaces and many software tools available on the website to help further researchers' goals. This is important as much of modern physical science involves running computer simulations in lieu of performing laboratory experiments. The resources needed to perform that kind of research are not always easily accessible by students, even on campuses with supercomputers and decent computer labs.

NanoHUB.org also has features fairly typical of a collective intelligence system, such as online group meeting rooms and the ability for researchers to post their studies and collect feedback.

NanoHUB.org has been extremely successful within the academic community. NanoHUB content has been cited over 1,000 times in scientific literature and used in over 450 classrooms at more than 150 universities. However, it has struggled to attract attention from those outside the academic world: only about 4% of the site's users work in the private sector.

IBM Watson

IBM Watson is one of the most famous computers in the world as well as a complex system of collective intelligence. In contrast to many of the projects covered by this survey, Watson's efforts are largely focused on the private sector as opposed to government or academic research.

One of the most common uses of Watson is to improve customer service. As one might expect, research has shown that traditional call center response to customer service complaints is often ineffective. Watson can help companies to leverage modern data collection methods to improve customer service.

The financial industry also enjoys the advantages of putting intelligent systems like Watson to work. Even the most talented investor cannot know all of the factors impacting their decisions, but algorithms can assist them in trying to understand the marketplace. Retail financial institutions will also be using Watson to learn more about their customers.

Most controversially, IBM Watson is also extremely involved in the healthcare field. Patient, research, and pricing data will play a growing role in the delivery of healthcare in the future, and large computer systems like Watson are up to the task of improving care through evidence and statistics. According to IBM, "the amount of medical information available is doubling every five years and much of this data is unstructured". There is simply no way for physicians and insurance companies to keep up with everything, and mistakes are often made. Intelligent systems such as Watson are the most efficient way to master the information overload in healthcare.

Watson can perform more than 80 trillion operations per second, and can read and understand 200 million pages of text in 3 seconds. Watson uses natural language and is an expert at understanding unstructured data, because it is powered by Apache UIMA (Unstructured

Information Management Architecture). It also takes advantage of Power Linux and Apache Hadoop.

Recorded Future

Recorded Future is a company based out of Massachusetts that has built a collective intelligence system with a goal of being able to predict the future. Using a technique known as "temporal analysis", its algorithms attempt to piece together a picture of who, what, when, where, and why things are (or will) happen. Analysis is performed in realtime, but "Recorded Future maintains an index with more than 100 million events, hosted on Amazon.com servers".

Recorded Future may be considered a more modern version of a system like EIES, as one of its biggest investors has been the U.S. government. In fact, the investment arm of the CIA purchased a stake in Recorded Future. Despite this, Recorded Future uses only publicly-available information when it performs searches. One of Recorded Future's biggest tasks has been to try and make predictions involving terrorism or events in the Middle East.

Google has also worked alongside Recorded Future as an investor and technology partner. Google hosts the Recorded Future API, which is where the magic happens for users of the service. Using the RF API requires a bit of tech savvy, as queries are generally performed in JSON and results are generated in JSON or CSV. Recorded Future can also be used with R or Python to simplify certain aspects of this procedure.

Recorded Future has vast application outside the world of government and national security. It is also used in business intelligence and competitive analysis as well as by hedge funds and economists.

Wikipedia

Wikipedia is one of the most popular websites on the internet, not only in English, but in 285 languages. But it is also one of the most innovative and frequently studied or imitated sources of collective intelligence. It has "harnessed the work of millions of people to produce the world's largest knowledge-based site along with software to support it, resulting in more than 19 million articles written, across more than 280 different language versions, in less than 12 years".

According to Wikipedia, it is "written collaboratively by largely anonymous Internet volunteers who write without pay". There are around 77,000 active contributors to Wikipedia, who come from "all ages, cultures, and backgrounds". The editors have all ranges of expertise, from casual readers to academic scholars. The diversity of the editor pool on Wikipedia helps to eliminate biases and produce articles that are editorially balanced.

Wikipedia is also a "live collaboration" - articles on current events may appear on the site within minutes, in contrast to traditional reference sources. Due to its digital nature, the amount of

knowledge that Wikipedia can contain is virtually unlimited, while print encyclopedias necessarily have constraints on what can be included.

The articles are also never considered complete though, and remain open for edits and improvements indefinitely. This is by design: "it is inherent in Wikipedia's editing model that misleading information can be added, but over time quality is anticipated to improve in a form of group learning as editors reach consensus... Wikipedia is a pioneer of communal knowledge building of this kind." Historically, encyclopedias have been produced through a very hierarchical process where an article is written by only one expert on a subject.

One of the drawbacks of Wikipedia's reliance on collective intelligence is that "some topics may not be covered well, while others may be covered in great depth". The minutiae of pop culture are often discussed at great length, while more scholarly subjects are neglected or left to experts with little time to contribute.

Despite this, the content shared on Wikipedia has a high level of accuracy. Articles tend to trend more accurate over time as a result of the collaborative writing process, and vandalism on the site is easily reversible as every version of each page is saved. A 2007 peer-reviewed study concluded that "42% of damage is repaired almost immediately".

Wikipedia has had several impacts on public policy. *The Wall Street Journal* found that federal courts of appeals had cited Wikipedia about 95 times between 2007 and 2012. In other instances, research from Wikipedia has been through out by the courts. Various groups with a political-oriented agenda have also seen Wikipedia as being legitimate enough to include as part of their propaganda campaigns. In the U.S., the Central Intelligence Agency is known to have altered the content of Wikipedia pages, and both sides of the Arab-Israeli conflict have attempted to present their viewpoint favorably on the website. Wikipedia attempts to control for these biases and uses tools such as WikiScanner to attempt to uncover where changes are being made by organizations that may be making edits for public relations purposes.

LinkedIn

LinkedIn can be considered a source of collective intelligence because the contacts that users are adding are generally people whom they have had some kind of contact with in the professional world. The strength or weakness of job seeker can then be evaluated based on something other than their resume. Potential employers can verify the information they have been given by candidates since a complete profile will usually contain a fairly large network and endorsements for different skills.

LinkedIn also has elaborate social graphing tools to map out the network of its users. The intelligence provided by mapping out the connections between people is used to make suggest more contacts. LinkedIn does this very accurately. However, many users have found the insights that LinkedIn's algorithms are capable of generating to be creepy, as the site can even suggest contacts which you may actually know but do not use the website.

View Your Social Graph
http://inmaps.linkedinlabs.com/network

Carinet, UCP-SARnet and sustainable community networks

UCP-SARnet is a computationally based social network dedicated to meeting the U.N. Development Goals by building a global network that supports local sustainable developemtn. Their website is used by students and community activists all over the world who are working to improve conditions in 75 countries. As might be expected, many of the projects discussed on UCP-SARnet are about ending poverty.

There are literally thousands of social networks around the world based on a wide variety of platforms. Organizers of these networks face a tough challenge as to whether to continue to develop their own platform or to migrate their community over to global social networks like Facebook and Linked-In.

Conclusion
We have reviewed a number of the key efforts to use computers to improve policymaking as a start to developing a historical awareness that can be used to design and evaluate better ways of approaching public policy. Some of these efforts like EIES and Group Systems have been cited thousands of times and researchers in these traditions have pushed cutting edge issues such as multicultural computer aided group decision-making and how to create data driven decision making processes.

We selected these ten major projects because we believe that with all of us being confronted by a wide variety of technologies for developing policy scenarios from Big Data, it is important for us to all have a shared history to base our discussions on.

The next step is to develop a schema for categorizing the major computational projects being used to support global decision making and for evaluating their effectiveness. Because we will inevitably draw from so many different disciplines in our attempt to create policy that is appropriate for our global network society, it is important that we continue to build a literature survey of milestones in computational based policy.