Before and Beyond Anticipatory Intelligence: Assessing the Potential for Crowdsourcing and Intelligence Studies

Alexander Halman
University of Pittsburgh

Follow this and additional works at: https://scholarcommons.usf.edu/jss
pp. 15-24

Recommended Citation

This Article is brought to you for free and open access by the Journals at Scholar Commons. It has been accepted for inclusion in Journal of Strategic Security by an authorized editor of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Before and Beyond Anticipatory Intelligence: Assessing the Potential for Crowdsourcing and Intelligence Studies

Abstract
Crowdsourcing is a new tool for businesses, academics, and now intelligence analysts. Enabled by recent technology, crowdsourcing allows researchers to harness the wisdom of crowds and provide recommendations and insight into complex problems. This paper examines the potential benefits and limitations of crowdsourcing for intelligence analysis and the intelligence community beyond its primary use: anticipatory intelligence. The author constructs a model and compares it to existing crowdsourcing theories in business, information science, and public policy. Finally, he offers advice for intelligence analysis and public policy.
Introduction

Crowdsourcing has become a ubiquitous term in colloquial English. ‘Crowdfunding’ websites like Kickstarter are quickly popping up on the internet, and people are beginning to harness the power of large groups thanks to a rapid proliferation in technology. These breakthroughs in technology, however, have resulted in a lag of theoretical research on the subject. James Surowiecki’s *The Wisdom of the Crowds: Why the Many Are Smarter Than the Few and How Collection Wisdom Shapes Business, Economies, Societies, and Nations* reintroduced the subject with a mass-market flavor reminiscent of Albert-Laszlo Barabasi’s *Linked*. His book begins with the familiar anecdote about Francis Galton and the first modern reference to the concept of crowdsourcing.

While attending a livestock fair in the early twentieth century, Galton overheard a contest in which the villagers were asked to guess the weight of an ox. Although each individual estimate varied greatly, the mean was within one pound of the animal’s actual weight. Thus, the wisdom of the crowds was born. Since Galton reported the results in *Nature* with his seminal paper, *Vox Populi*, researchers from mathematics to business have examined the issue. However, with little theoretical work in regards to specific subjects, scholars and practitioners can be mired into a single thought process. In other words, there are only a handful of conceptual pieces on crowdsourcing in the intelligence community, and none branch out beyond the implications for forecasting.

The Intelligence Advanced Research Projects Activity (IARPA), for instance, has focused a tremendous amount of effort and funding on crowdsourcing. Projects such as the Aggregative Contingent Estimation (ACE), Forecasting Science and Technology (ForeST), Foresight and Understanding from Science Exposition (FUSE), and Open Source Indicators (OSI) all demonstrate significant strides in crowdsourcing for anticipatory intelligence, but illustrate the stagnation on the subject as well. While the agency is supposed to epitomize high-risk high-reward research, it does little to expound upon the other aspects of crowdsourcing that have been advanced by scholars in other fields.

Ultimately, the purpose of this paper is to explore the opportunities and limitations for crowdsourcing in the field of Intelligence Studies. Intelligence studies is a relatively new field, though it borrows methodological and epistemological roots from areas such as economics, psychology, and policy sciences. With the surprise of 9/11 and bruising of the intelligence community’s reputation, Congress passed the *Intelligence Reform and Terrorism Prevention Act of 2004* to formalize, among other alterations, analytic methods. Nevertheless, despite substantial effort to improve analysis, our thinking can often times devolve into reinventing the wheel.

The paper will progress as follows: First, I conduct a literature review on the concept and evolution of crowdsourcing in academia and practice. This section will conclude with a review of the current state of research on crowdsourcing in Intelligence Studies and limitations. Next, I develop a theoretical framework to examine the potential benefits and limitations for various types of crowdsourcing beyond forecasting in the
intelligence community. This is followed by some implications for intelligence analysis scholars and practitioners. Finally, there is a brief discussion about future research and conclusions.

**Literature Review**

*Concept and Definition of Crowdsourcing*

As the usage of crowdsourcing is nearing ubiquity, so too are the definitions and concepts. Some scholars focus on factors that make crowds wise, such as diversity of opinion, independence, decentralization, and aggregation.\(^1\) Similarly, researchers find that increasing the size and diversity of sources like Wikipedia improve quality.\(^2\) These analysts primarily examine the causal mechanisms of group knowledge and elements that result in superior results. On the other hand, a vast majority of contemporary scholars seek to pinpoint the actual process and definition of crowdsourcing. For example, after reviewing a plethora of definitions from many fields, one scholar defined it as,

“...a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage that what the user has brought to the venture, whose form will depend on the type of activity undertaken”\(^3\).

There are several key aspects to this definition. First, crowdsourcing is an online activity, indicating that the concept of Web 2.0 is necessary. These websites cannot be static, and must facilitate cooperation and interaction. Technology is necessary, but not sufficient, according to this definition. The strategy and task must adequately match the requirements of an organization, individual, or institution. Finally, both sides must benefit from the transaction. Whether it is intrinsic or extrinsic motivation, the participants must receive compensation. With regards to the intrinsic factors, for example, a leaderboard or ‘game-ified’ system could encourage contributors to add content and provide a sense of accomplishment. The researchers receive useful feedback and the user can gauge his/her progress or standing among others.

---

Crowdsourcing in Business

In order to access the significance of crowdsourcing in Intelligence Studies, one must track the evolution of the concept. After the original report of wisdom of the crowds by Galton, business was one of the first modern fields to seriously consider crowdsourcing (in addition to mathematics). The aforementioned Web 2.0 allowed for a greater profit margin. Business scholars not only examined how tasks could be ‘crowdsourced’ out, but also studied what advances could be made in innovation. The concept of Crowdsourcing Inventive Activities (CIA) was defined and framed through the theory of transaction costs and the evolutionary theory of the firm. In the end, crowdsourcing was used for innovation if it was profitable; unsurprising for business literature.

Other literature looks at aptitude in business crowdsourcing. For example, an empirical test using Amazon’s Mechanical Turk (AMT) found that not all crowds are so wise. Groups with a high reputation outperformed those with a low reputation, and paying too much or too little changes work quality as well. However, simply utilizing crowdsourcing will not result in desired effects. There must be well-crafted strategies, a careful analysis of goals, and an appropriate environment. Contemporary business literature is teeming with crowdsourcing material; nevertheless, many of the articles’ focus are on the same concepts and definitions. Indeed, the modern business research on crowdsourcing seems to be mired like Intelligence Studies, but not nearly to the same effect. The next section discusses crowdsourcing within the information sciences.

Crowdsourcing in Information Sciences

Although much of the information science literature on crowdsourcing predominately discusses how technological innovation has facilitated crowdsourcing, the author would be remiss not to discuss other research in the field as some of the research is applicable to Intelligence Studies. Some scholars, for example, straddle multiple fields and focus on the relationship between networks, subject matter experts, and crowdsourcing. These researchers, still coming to terms with the concept, primarily focus on meta-analysis, literature reviews, and conceptual frameworks. Indeed, in many fields that utilize

---

8 Aitamurto et al., "The Promise of Idea Crowdsourcing."
crowdsourcing, each has conducted multiple literature reviews, but often fail to go beyond a disciplinary approach.

Additionally, some researchers in the information sciences, similar to those in the business sector, examine when to crowdsource. One empirical study looks at how reluctant individuals choose new participants for crowdsourcing; interestingly, good crowdsourcers pick good replacements. This is similar to the aforementioned research that found not all crowds are equal. Each of the studies in the information sciences also illustrates stagnation in the field and demonstrates there is a dearth of empirical research on the subject. Of all the research reviewed, only one experimented with and tested the concepts scholars had developed. The next section on the policy sciences moves closer to Intelligence Studies and reveals insight for the future of crowdsourcing in intelligence analysis.

Crowdsourcing and Policy Sciences/Public Administration

Similar to the previous sections, the literature on crowdsourcing is relatively new and facilitated by technology. However, many innovations provide the benefits of crowdsourcing to public entities, but do not explicitly mention it. Researchers in urban planning use crowdsourced GIS data to provide better services and future resources. Additionally, some government agencies use it to facilitate public-private partnerships.

The research in the public sector has started to recognize the significance of crowdsourcing and Web 2.0 technologies on the subjects of democracy and citizen participation. Many analysts in the field believe these technologies, with proper application, have the abilities to strengthen democratic institutions. Tim Van Gelder’s seminal piece examines the impetus for unstructured forums in order for citizens to discuss major issues. This article allows future researchers on the subject to hypothesize new uses for this technology. In a similar vein, studies in public policy have

---

begun to look at crowdsourcing in the context of the policy cycle.\textsuperscript{17} The figure below conceptually maps how some of the various types of crowdsourcing can be used in each part of the public policy process.

The authors visually and quantitatively demonstrate that each crowdsourcing technique is more appropriate depending on the stage in the policy cycle. This echoes the previous research on crowdsourcing and when it should be utilized. The final two sections of the literature review examine the current state of crowdsourcing in Intelligence Studies and its pitfalls.

\textit{Current State of Crowdsourcing in Intelligence Studies}

As previously mentioned, all of the available literature on crowdsourcing in Intelligence Studies is on predictions, forecasting, and participatory intelligence. IARPA’s ACE epitomizes the use of crowdsourcing in the intelligence community. Philip Tetlock’s work has shown that groups of people predict and estimate better than subject matter experts, and that individuals with certain cognitive biases perform worse than others.\textsuperscript{19} Other researchers also examine aggregation for analytic tasks\textsuperscript{20}, but there is no testing or conception for utilization beyond forecasting.

An unclassified paper has begun the discussion, but still has not reached past the ‘prediction market.’ Overall, the authors argue the Intelligence Community Prediction Market (ICPM) should support decision makers, support analysts, identify the best

\begin{figure}
\centering
\includegraphics[width=\textwidth]{crowdsourcing_in_policy_cycle.png}
\caption{Figure retrieved from the MIT Center for Collective Intelligence\textsuperscript{18}}
\end{figure}

\textsuperscript{18} John Prpić, Araz Taeihagh, and James Melton, "Crowdsourcing the Policy Cycle," \textit{2014 Collective Intelligence Conference} (June 10-12, 2014).
forecasters in the IC, and provide a test for future study. Nevertheless, there is a difference between estimating and problem definition, discovering the elements of a problem, or exploring their relationship.

Further Crowdsourcing Critiques and Limitations

Even though there is a plethora of research supporting the use of crowdsourcing in many situations, there are also researchers who dispute its merits. One researcher argues that the wisdom of the crowds could actually be the ignorance of the crowds; in other words, subject matter experts who know the material can be overruled. Wikipedia is used as a prime example of both the benefits and drawbacks of crowdsourcing. Because the content can be edited by anyone, there is the potential for incorrect information to continually surface.

Similarly, some scholars agree that crowdsourcing is best for issues involving optimization, and not very useful for creativity or innovation. This is seemingly in stark contrast to the business and information science literature, yet they can be reconciled by determining the appropriate situation to crowdsource. Other criticism is drawn from a like technique called Delphi. Delphi is comparable to crowdsourcing for subject matter experts. Sackman, nonetheless, argues that this methodology is severely flawed. He concludes that conventional Delphi is an unreliable and scientifically unvalidated technique in principle and probably in practice. Except for its possible value as an informal exercise for heuristic purposes, Delphi should be replaced by demonstrably superior, scientifically rigorous questionnaire techniques and associated experimental procedures using human subjects.

Furthermore, there is applicable research that detracts from crowdsourcing’s validity by examining culture. Indeed, if not all crowds are wise, then culture has a profound influence on how problems are solved. Finally, crowdsourcing may be committing a Type III error, or simply formulating the wrong problem. The following sections examine context validation and Type III errors for crowdsourcing in Intelligence Studies. I develop a theoretical foundation in order to access the potential for this technique beyond anticipatory intelligence and forecasting.

25 Ibid.
Conceptual Framework - Beyond and Before Anticipatory Intelligence

The Policy Cycle and the Intelligence Cycle

The aforementioned research on the policy cycle and crowdsourcing was one of the first systematic conceptual models to examine the potential of several crowdsourcing techniques in the public sector. There are many parallels that can be made between the policy cycle and the intelligence cycle. The figure below illustrates the similarities.

Currently, crowdsourcing in Intelligence Studies is stuck in the analysis phase (circled in red). The model above shows how there are other ways in which this technique can be utilized.

Problem Definition, Requirements, and Context Validation

By jumping directly to analysis, intelligence analysts and policymakers can skip and miss important questions. In the intelligence community, this could mean another terrorist or surprise attack. Ultimately, programs such as IARPA’s ACE and FoREST can answer a question or accurately predict a phenomenon, yet they may not be answering the right question. The realm of public policy and intelligence analysis is riddled with ill-structured problems; indeed, we altogether could be formulating the wrong problem.28 By utilizing the concept of context validation and Pragmatic Eliminative Induction,29 intelligence analysts can begin to answer the right questions. For example, a question on The Good Judgment Project asks: “Will Russia invade Eastern Ukraine in the next six

months?” Maybe that is not the right question. Crowdsourcing can help us get closer to the right question, and then answer it.

Policy Implications

Incentivizing Interagency Forums

The conceptual model above demonstrates one example in which intelligence agencies may be solving the wrong problems and asking the wrong questions. In the unstructured environment of uncertainty, crowdsourcing has the potential to innovate intelligence analysis through problem definition, context validation, and Pragmatic Estimative Induction. Since 9/11 and the criticism of a lack of information sharing, agencies have begun to implement interagency forums. In the end, however, there is little incentive to participant, and this is not a problem limited to the intelligence community. Academia also suffers from a lack of incentivizes to participate in non-peer reviewed journals despite the potential to provide valuable information to decision-makers.

These forums do exist, but throughout a busy day, analysts may not have time to invest in ways that do not advance their careers. There is significant evidence these programs are beneficial. For instance, the intranet website Intellipedia facilitates information sharing through a closed Wikipedia-like format. One example of its success was, in 2007, when twenty-three users from various agencies created a page that covered Iraqi insurgents using chlorine-based IEDs.30 This crowdsourcing program helped reduce stovepipes, yet if employees are evaluated on other products, they will be less likely to contribute.

Indeed, the forums epitomize both a bottom-up and top-down approach. The infrastructure required upper management to approve these technologies, but also needed employees to add content. Therefore, middle management needs to take these forums into account when evaluating employees; similarly, analysts need to highlight these contributions in their portfolios and briefings.

Tournament-based Collaboration for Processing

The research from MIT’s Center for Collective Intelligence theorized that tournament-based collaboration would be optimal for policy development. Other institutions have already begun to utilize crowdsourcing for data processing. An excellent example is the online puzzle video game called Foldit. Developed by the University of Washington, this game opens up a difficult problem of deciphering protein folding to the public. In 2011, the ‘gamified’ version of the problem was solved by the public, a feat unaccomplished by scientists in fifteen years; it was done in just three weeks!31 The lesson: establish fun,
collaborative competitions for intractable problems plaguing the community. The agencies can make a week like color war or ‘floor war’ to see who gets the best results. This can raise morale and have the potential to solve problems in-house.

IARPA and The Good Judgment Project have already found that many average people do better than the pros.32 The model in this paper lends further credence to this point and agencies should begin to open up insensitive large datasets to the public. The intelligence community has an Open Source Center; nevertheless, it is just collection from publically available sources. U.S. agencies should analyze big datasets that are insensitive and ‘gamify’ them like Foldit.

Discussion

In this section, I provide a brief discussion about the potential and limitations of crowdsourcing in the intelligence community. One issue is the difficulty of change in the bureaucracy. There are seventeen members or elements in the U.S. intelligence community, and with any large entity, there are competing interests. Deviations from the status quo are feared by management. As much as change is lamented, the post-9/11 alterations illustrate that it can be done.

Another important notion is the lack of empirical research on crowdsourcing. In each literature base in the review, it is clear there are few experimental studies on the subject. Thus, it is imperative that scholars, both in Intelligence Studies and in other fields, begin to test the models that others have put forth. Like Structured Analytic Techniques, crowdsourcing is taken to heart quickly, but is rarely tested. In reality, we may be taking these analytic techniques for granted and they may not provide the benefits touted by private contractors. Certainly, more research is warranted.

Finally, practitioners must recognize the limitations of crowdsourcing. There are various types of crowdsourcing techniques and a multitude of analytic tasks that these agencies perform. The U.S. intelligence community needs to examine how businesses and research intuitions implement them, and create lessons learned. Crowdsourcing will not solve all of the problems, but, if properly utilized, can take advantage of the wisdom of subject matter experts and dilettantes in the agencies.

Conclusions

In summation, crowdsourcing is a new technique used by many fields, organizations, and government agencies that was facilitated by technological innovation. IARPA and other intelligence agencies have started programs such as ACE to leverage aggregate estimation for anticipatory intelligence. This paper asks the question: did we skip a step? By comparing the intelligence cycle with the public policy cycle, we can see how

different crowdsourcing techniques can be used in each step of the cycle. This paper provides just one example, however, the conceptual framework can be applied to other stages of the cycle as well.

Furthermore, it is imperative that the intelligence community asks the right questions. These agencies and other public entities are operating in a ‘wilderness of ill-structured problems.” There are several actions U.S. intelligence agencies can take to fully optimize the usage of crowdsourcing. The above-mentioned recommendations are a good start, but more needs to be done. With further research and experimentation on the implementation of such programs, these agencies, which have critical missions for national security, can truly take advantage of the wisdom of the crowd.