CNS

OCCASIONAL PAPER

#30 · AUGUST 2017

Crowdsourcing Systems and Potential Applications in Nonproliferation

Bryan Lee

Middlebury Institute *of* International Studies at Monterey

James Martin Center for Nonproliferation Studies

Acknowledgements:

The author would like to thank Julia Diamond and Ryan Genzoli for their research and preparation of this study.

James Martin Center For Nonproliferation Studies
www.nonproliferation.org

The James Martin Center for Nonproliferation Studies (CNS) strives to combat the spread of weapons of mass destruction by training the next generation of nonproliferation specialists and disseminating timely information and analysis. CNS at the Middlebury Institute of International Studies at Monterey is the largest nongovernmental organization in the United States devoted exclusively to research and training on nonproliferation issues.

Middlebury Institute of International Studies at Monterey
www.miis.edu

The Middlebury Institute of International Studies at Monterey, a graduate school of Middlebury College, provides international professional education in areas of critical importance to a rapidly changing global community, including international policy and management, translation and interpretation, language teaching, sustainable development, and nonproliferation. We prepare students from all over the world to make a meaningful impact in their chosen fields through degree programs characterized by immersive and collaborative learning and opportunities to acquire and apply practical professional skills. Our students are emerging leaders capable of bridging cultural, organizational, and language divides to produce sustainable, equitable solutions to a variety of global challenges.

James Martin Center for Nonproliferation Studies
Middlebury Institute of International Studies at Monterey
460 Pierce St., Monterey, CA 93940, U.S.A.
Tel: +1 (831) 647-4154
Fax: +1 (831) 647-3519

CONTENTS

EXECUTIVE SUMMARY

This report provides an overview of crowdsourcing systems and identifies the key elements in their successful operation. It also offers an examination of best practices for US government implementation of crowdsourcing projects. Finally, it describes four instances of crowdsourcing projects with applicability to arms control and nonproliferation.

Crowdsourcing systems have been used to address a variety of problems, but government applications are typically restricted to one of four types:

» Knowledge discovery and management

» Distributed human intelligence tasking

» Broadcast search

» Peer-vetted creative production

Successful crowdsourcing outcomes are the result of both proper problem identification and an appropriately structured crowdsourcing management system. This study identifies the key management tasks as user management, task management, contribution management, and workflow management. In addition, two design tasks were identified, incentive design and crowd curation.

Using publicly available information, the report examines four case studies where public and private sector actors implemented crowdsourcing systems. The cases cover various proliferation-relevant topics, such as border management, scientific problem solving, arms trafficking, and crime reporting and tracking. In each case, success or failure was largely attributed to the degree of management involvement and the implementation of the key management and design tasks mentioned above.

The growing body of government-sponsored crowdsourcing projects as well as the establishment of a government crowdsourcing center of excellence should give additional insights into the development and application of crowdsourcing systems. Until the results are widely shared and adopted, however, an examination of the key tasks provided here together with a review of the selected case studies can provide a shortcut for implementation. With the appropriate degree of management focus, there is every reason to believe nonproliferation and arms-control efforts can also benefit from crowdsourcing methods and technologies.

INTRODUCTION

The idea that two heads are better than one is a truism no one should ignore. The famous statistician Sir Francis Galton realized this at a town fair in 1907 during a competition to guess the weight of a prize-winning ox. To make things more difficult, contestants had to guess the weight of the fully dressed animal—that is, what it would it weigh after it had been slaughtered and prepared for sale.[1] Dairy farmers, butchers, and other experts sized up the animal, and their best estimates were entered into the prize book. Of course, the townspeople also made their guesses—the prize was too good to pass up—even if they knew nothing about livestock and had to pay an entry fee of one sixpence. In the end, Galton asked for the entry records and witnessed something amazing. The actual weight of the ox was 1,198 pounds. The mean estimate of the 787 entries Galton reviewed was 1,197 pounds. With enviable understatement Galton concluded, "The result is more creditable to the trust-worthiness of a democratic judgment than might have been expected."[2]

The Galton story has since been repeated many times and has become the foundation of the "wisdom of crowds" idea. Many point to it as proof that the concept of crowdsourcing, or using a large external group to generate ideas or solutions, is nothing new. Others, however, reject that history and believe that crowdsourcing in the modern sense can only exist thanks to the Internet and its ability to facilitate information sharing and aggregate the results.

One argument for the newness of crowdsourcing is the difficulty of defining what it is and what problems it can solve. One recent study found more than forty different definitions of crowdsourcing in the academic and business literature. [3] Federal and state governments also seem to be in experimental mode, with over 400 examples of crowdsourcing efforts showing decidedly mixed results.[4] Despite well-publicized examples such as the Defense Advance Research Project Agency's (DARPA) "Red Balloon" challenge or the State Department's Tag Challenge, progress in developing best practices for crowdsourcing has been slow both inside and outside of government.[5] This has been particularly true with respect to the potential to apply crowdsourcing in the field of nonproliferation and arms control.

This paper attempts to address that shortfall by looking at government best practices for crowdsourcing in general and in the nonproliferation context in particular. The remainder of the paper is organized as follows. First, it will provide a detailed overview of crowdsourcing and identify the common features of successful crowdsourcing efforts. It will then describe a set of best practices in a governmental context

1 Francis Galton, "Vox Populi," Nature, No. 75 (1907): 450–51.
2 Ibid, p. 451.
3 Enrique Estellés-Arolas and Fernando González-Ladrón-de-Guevara, "Towards an Integrated Crowdsourcing Definition," Journal of Information Science, Vol. 38, No. 2 (April 2012), p. 189.
4 Kwasi Mitchell et al., "The Craft of Incentive Prize Design: Lessons from the Public Sector," Deloitte University Press, June 18, 2014, <https://dupress.deloitte.com/dup-us-en/topics/social-impact/the-craft-of-incentive-prize-design.html>.
5 DARPA, "MIT Red Balloon Team Wins DARPA Network Challenge," December 5, 2009, <https://web.archive.org/web/20101111082411/https://networkchallenge.darpa.mil/darpanetworkchallengewinner2009.pdf>; Alex Rutherford et al., "Targeted Social Mobilization in a Global Manhunt," PloS One, Vol. 8, No. 9 (2013), p. e74628.

taken from one of the leading researchers in the field. Finally, it offers four case studies demonstrating the use of crowdsourcing in settings related to nonproliferation and arms control.

CROWDSOURCING SYSTEMS

In the context of modern web 2.0 systems, crowdsourcing is still in its infancy. Nevertheless, many scholars have accepted the challenge to try to better understand the phenomenon by creating models and decomposing systems. Fortunately, because so many of the crowdsourcing examples come from the business world, much of the research is focused on real-world application and is refreshingly clear of academic jargon. In fact, many academic studies use Jeff Howe's 2008 popular work of nonfiction, *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business,* as a springboard for their systematic studies. The end effect has been to approach crowdsourcing systems more from the bottom up, rather than the top down and typically theory-driven academic approach.

This approach begins with the simple question: what are the common features of any crowdsourcing system? There is, of course, some debate, but four fundamental challenges are common: how to recruit and retain users; what contributions can those users make; how to aggregate the user contributions; and how to evaluate those contributions.[6] Many authors also point to the nature of collaboration (explicit or implicit) or to the type of problem (creating or deciding) as further distinguishing factors.[7]

For practitioners, a problem-centric approach is likely the most useful. Such an approach was developed by Darren Brabham in 2009 specifically for the public sector. This approach begins by asking whether a manager has an information-management problem or an ideation problem.[8] Based on the answer to that question, the manager then decides whether the need for information is to locate and assemble it or to analyze information that is already available. For ideation problems, the manager asks whether the answer must be empirically verifiable or is a question of taste or fashion. Depending on which problem-solution pathway one takes, one can end up at one of four different types of crowdsourcing systems.[9]

Brabham calls the first type of system "knowledge discovery and management" and says it is ideal for information gathering, organization, and reporting problems.[10] The second type of system is "distrib-

6 Anhai Doan, Raghu Ramakrishnan, and Alon Y. Halevy, "Crowdsourcing Systems on the World-Wide Web," Association for Computing Machinery. Communications of the ACM, Vol. 54, No. 4 (April 2011), p. 86.
7 Ibid, p. 87; Thomas W. Malone, Robert Laubacher, and Chrysanthos Dellarocas, "Harnessing Crowds: Mapping the Genome of Collective Intelligence," MIT Center for Collective Intelligence, February, 2009, <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1381502>.
8 Daren C. Brabham, "Using Crowdsourcing in Government," 2013, pp. 10-11, Figure 2, <http://www.cbs.dk/files/cbs.dk/using_crowdsourcing_in_government.pdf>.
9 Ibid.
10 Ibid, p. 11, Table 1.

uted human intelligence tasking." This system is best for large-scale data analysis that cannot be done effectively by machines, such as handwritten transcription or object labeling in photos.[11] The third type is "broadcast search," which is used for complex problem-solving with a verifiable solution, such as identifying a better recommendation algorithm for Netflix choices.[12] Last comes "peer-vetted creative production," which is aimed at finding solutions that are matters of subjective judgment, such as logo designs or architecture competitions.[13]

Once the basic type of crowdsourcing problem has been identified, the next step is to design a crowdsourcing system with the best chance to find the solution to the problem. In other words, what features and functions must be included to resolve the four challenges of crowdsourcing mentioned above?

An additional factor is that these features and functions must be congruent with government use. For example, it is common in the business world to run design challenges for items such as logos or mascots and not offer any financial compensation for the participants. Such practices frequently come in for criticism by supporters of labor rights and would likely be highly controversial in the public sector.[14] However, research has also shown that the crowds attracted to public-sector challenges tend to be more intrinsically motivated and may actually be discouraged by monetary incentives.[15]

What this means is that crowdsourcing in the public sector faces greater management challenges than the private sector. For this reason, three components of crowdsourcing systems stand out above the rest: the management mechanism, the incentives, and the crowd selection.

A 2015 study of government crowdsourcing open-innovation efforts found that only 44 percent made any effort to implement the crowd-submitted ideas.[16] In fact, the vast majority were ill-conceived, poorly executed, and produced few tangible results.[17] This is not surprising, as a large number of crowdsourcing attempts in the business world fail as well.[18] What is also unsurprising is that these failures are principally management failures. That is, the problem is not with crowdsourcing *per se* but with a poor attempt to implement a crowdsourcing solution. Because of the criticality of the management function, it is worth a deeper examination. The chief components of any crowdsourcing system should be aligned with key management tasks. A comprehensive literature review by Lars

11 Ibid, pp. 11, 16, Table 1.
12 Ibid, p. 11, Table 1.
13 Ibid, p. 11, Table 1.
14 Daren C. Brabham, "The Myth of Amateur Crowds," Information, Communication & Society, Vol. 15, No. 3 (April 1, 2012), p. 394.
15 Kevin Cupido and Jacques Ophoff, "A Model of Fundamental Components for an E-Government Crowdsourcing Platform," Electronic Journal of E-Government, Vol. 12, No. 2 (December 2014), p. 154.
16 Ines Mergel, "Opening Government Designing Open Innovation Processes to Collaborate With External Problem Solvers," Social Science Computer Review, Vol. 33, No. 5 (October 1, 2015), pp. 605, 608.
17 Ibid, 605–608.
18 Henri Simula, "The Rise and Fall of Crowdsourcing?," in 2013 46th Hawaii International Conference on System Sciences (HICSS) (IEEE, 2013), pp. 2783–2791, <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6480180>; Johann Füller, "Die Gefahren Des Crowdsourcing," Harvard Business Manager, June 27, 2012, <http://www.harvardbusinessmanager.de/blogs/a-840963.html>.

Hetmank identified these tasks as: user management, task management, contribution management, and workflow management.[19]

USER MANAGEMENT

User management components include functions to register users, evaluate them, and enable coordination.[20] User registration is simply a means to record participation and does not imply any evaluation. Usually, this is accomplished with a simple website form. There is some evidence to suggest, however, that decreasing anonymity by using social media profiles can increase trust and improve results.[21]

The evaluation function is related to the challenge of crowd selection and includes preselection mechanisms, such as specific skills or affiliations, and post-selection mechanisms, such as reputation management functions. These types of mechanisms are very common in human intelligence tasking systems and also play a very important role in broadcast search systems.

One function that is crucial but often overlooked is enabling coordination. This is especially the case in broadcast search crowdsourcing, because solutions to complex problems are frequently the result of integrating knowledge. One broad study of innovation challenges found that winning solutions were often buried in comment discussion threads and not presented as standalone ideas.[22] The most successful systems, therefore, are those where the group is capable of fostering collaboration and coordinating the selection of the best ideas.

Equally important for coordination is a means of providing two-way communications between the crowd and the task provider. In this way, task guidance can be clarified, and promising contributions can be encouraged. Integrating social media accounts into the project has been shown to be an effective means of facilitating coordination and two-way communications.[23] Provider feedback has also been shown to be an important success factor in effective incentives design.

---

19 Lars Hetmank, "Components and Functions of Crowdsourcing Systems - A Systematic Literature Review," 2013, pp. 61, 63–65, <http://wi2013.de/dateien/WI2013_Proceedings_Volume_1.pdf>.
20 Ibid, p. 63.
21 Klinger, Jörn, and Lease, Matthew. "Enabling Trust in Crowd Labor Relations Through Identity Sharing." Proceedings of the Association for Information Science and Technology, Vol. 48, No. 1 (2011), pp. 1–4.
22 Arvin Malhotra and Ann Majchrzak, "Managing Crowds in Innovation Challenges," California Management Review, Vol. 56, No. 4 (Summer 2014), p. 112, <http://search.proquest.com/docview/1559118855/D13C50D353414B9DPQ/1>.
23 Huiji Gao et al., "Harnessing the Crowdsourcing Power of Social Media for Disaster Relief" (DTIC Document, 2011), <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA581803; Yefeng Liu et al., "Drawing on Mobile Crowds via Social Media," Multimedia Systems, Vol. 18, No. 1 (February 1, 2012), pp. 53–67>.

TASK MANAGEMENT

Task management is the process of designing a task for the crowd and assigning it to the right person.[24] We will discuss incentives and crowd selection later in this section, but for now our focus will turn to task specification.

In reporting on the more than fifty highly successful crowdsourcing efforts it has conducted, NASA administrators said, "Challenge design should be informed by desired outcomes that align with NASA's mission."[25] NASA's experience also emphasizes the importance of management alignment. One key lesson was something as simple as ensuring that all members of the management team were using the same terms of reference and had the same understanding of what success was for the project.[26]

Common sense alone would conclude that poorly articulated challenges would not lead to successful outcomes. On the other hand, an overly restrictive task description hinders creative thinking and will prevent problem solvers from reaching an optimal solution. Research has shown there is something of a Goldilocks effect in specifying tasks. If the task appears too difficult, crowds are unlikely to participate. If the task appears too easy, crowds anticipate becoming bored and also will not participate.[27]

The solution to this conundrum is to carefully think through the problem that needs to be solved, and then offer explicit guidance on the requirements for a successful solution. Once this is done, the task provider must be prepared to engage with problem solvers by answering questions, providing encouragement, and making refinements to the requirements. There is also good evidence that including explicit instructions to integrate knowledge produced by other participants leads to better solutions.[28] So, the ideal task specification is in alignment with the organization's mission, is agreed upon by the project implementers, and is spelled out in enough detail to encourage participation without stifling creativity.

Leaving aside for now the question of crowd selection, task management must first decide how tasks will be assigned. There are three choices. A task can be sent to a single person, which is commonly the case for human intelligence tasking systems such as Amazon's Mechanical Turk. Alternatively, the task can be sent to a subset of the larger crowd. This is often seen in both broadcast searching and peer-reviewed types of projects. Finally, the task can be open to the crowd at large, which is usually the case for large information seeking projects, such as NASA's stellar nursery Disk Detective project.[29] Again, the type of crowd most appropriate to the task should be identified in advance. It is only once that is complete that the task manager can link the choice to the broader issue of crowd selection.

---

24 Hetmank, "Components and Functions of Crowdsourcing Systems - A Systematic Literature Review."
25 Jennifer L. Gustetic et al., "Outcome-Driven Open Innovation at NASA," Space Policy 34 (November 2015): 12-13.
26 Ibid.
27 David Geiger and Martin Schader, "Personalized Task Recommendation in Crowdsourcing Information Systems — Current State of the Art," Decision Support Systems, No. 65 (September 2014): 3–16, doi:10.1016/j.dss.2014.05.007.
28 Malhotra and Majchrzak, "Managing Crowds in Innovation Challenges."
29 Hetmank, "Components and Functions of Crowdsourcing Systems – A Systematic Literature Review"; "Disk Detective," Disk Detective, n.d., <https://www.diskdetective.org/#/about>.

One other aspect of task management that is frequently overlooked is the question of time. This includes the overall project time, such as any deadlines that must be met in order for the solution to be usable, as well as the time allotted for users to accomplish the particular task. For human intelligence type tasks, individual assignment time takes precedence and is frequently automated. Amazon's Mechanical Turk, an online marketplace for work that demands human intelligence, operates on such a basis.[30] For the other types of crowdsourcing systems, time is usually specified in terms of deadlines and overall project duration.[31]

CONTRIBUTION MANAGEMENT

Once task management has been resolved, the manager must next consider how to manage crowd contributions. There are two issues at play here. The first is how to separate legitimate entries from illegitimate ones, and the second is how to evaluate the entries that make it through this legitimacy screen. Let us first consider the problem of legitimacy.

The issue of malicious users plagues all online systems, regardless of whether users remain anonymous or have public profiles.[32] Indeed, one of the most celebrated Department of Defense crowdsourcing contests, the DARPA Shredder Challenge, was engulfed in a cheating scandal when it was revealed one of the teams had systematically deleted entries from another competitor.[33] Submitting false or misleading information is also a concern, as seen in another famous Department of Defense crowdsourcing contest, the DARPA "Red Balloon" challenge.[34] This concern becomes even more serious in military or disaster-relief operations where lives may be at stake. Clearly, determining the legitimacy of entries must be a major part of the contribution-management system.

Legitimacy is typically insured by voting mechanisms, i.e. the same question is submitted to multiple users and the consensus answer is the one accepted. This approach is frequently used in human-intelligence testing systems. An alternative method is expert evaluation, where a team of preselected experts review entries at a series of project stage "gates" to determine legitimacy. If this approach is used, research shows the best results are obtained using scale rating systems instead of absolute measures.[35]

The final method often used is that of reputation management. This method only accepts entries from users whose contributions have been deemed acceptable in the past. This acceptability can

30 "Amazon Mechanical Turk," <https://www.mturk.com/mturk/help?helpPage=overview>.
31 "Seeker FAQs," InnoCentive, 2016, <https://www.innocentive.com/offering-overview/seeker-faqs/>; "How Threadless Works," Threadless, accessed November 23, 2016, <https://www.threadless.com/how-it-works/>.
32 Victor Naroditskiy et al., "Crowdsourcing Contest Dilemma," Journal of the Royal Society Interfact, Vol. 11, No. 99 (August 20, 2014), pp. 1, 6–7.
33 Nicolas Stefanovitch et al., "Error and Attack Tolerance of Collective Problem Solving: The DARPA Shredder Challenge," EPJ Data Science, Vol. 3, No. 1 (September 30, 2014), pp. 3, 15.
34 John C. Tang et al., "Reflecting on the DARPA Red Ballon Challenge," Communications of the ACM, April 2011, 82–84, <http://cacm.acm.org/magazines/2011/4/106587-reflecting-on-the-darpa-red-balloon-challenge/fulltext>.
35 Ivo Blohm et al., "Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing," Information Systems Research, Vol. 27, No. 1 (2016), 27–48.

be based on user ratings, previous or management preselection. There is some evidence to suggest, however, that reputation-management systems may be prone to manipulation by Machiavellian personality types.[36] For this reason, managers must maintain clear ratings guidelines and be prepared for intervention if necessary.

IDEA EVALUATION

Once the legitimacy of the contribution has been determined, the manager must assess it from both technical and management perspectives. These will not necessarily differ depending on the organization and the problem, but aspects of quality are usually considered, such as novelty, feasibility, relevance, and elaboration.[37] On the technical side, assessment may be done with tools such as rating scales and prediction markets, comment evaluation, or voting systems.[38] On the management side, assessment includes defining the metrics or dimensions for evaluation and ensuring the solution aligns with the organization's mission.

Returning to quality, the aspect of novelty is the consideration of whether the idea is new or particularly innovative. Feasibility explores whether the proposed solution can be applied in the context of the organization or its operating environment, as well as the ease with which the idea can be "brought to market."[39] This latter consideration is often ignored in breathless reports of crowdsourcing success. Perhaps the most famous example is the $1 million Netflix algorithm prize, a crowdsourcing contest to devise a better film recommendation system for Netflix customers. Although the prize was awarded with much fanfare and the project is frequently cited as a crowdsourcing success, it was later announced that the algorithm would not be implemented because "the additional accuracy gains that we measured did not seem to justify the engineering effort needed to bring them into a production environment."[40]

Idea relevance is sometimes called usefulness and captures the degree to which the solution solves a tangible problem. Relevance may also incorporate the financial advantages of the idea.

---

36 Katja Hutter et al., "Machiavellianism or Morality: Which Behavior Pays Off In Online Innovation Contests?," Journal of Management Information Systems, Vol. 32, No. 3 (July 3, 2015), p. 198.
37 Ivo Blohm et al., "Does Collaboration Among Participants Lead to Better Ideas in IT-Based Idea Competitions? An Empirical Investigation," International Journal of Networking and Virtual Organisations Vol. 9, No. 2 (2011), pp. 106–122.
38 Blohm et al., "Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing"; Malhotra and Majchrzak, "Managing Crowds in Innovation Challenges," 105–106; M. C. Yuen, I. King, and K. S. Leung, "A Survey of Crowdsourcing Systems," in 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011, pp. 766–767.
39 Hua Ye et al., "Collaboration and the Quality of User Generated Ideas in Online Innovation Communities," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, August 7, 2012), <https://papers.ssrn.com/abstract=2546421>.
40 Xavier Amatriain and Justin Basilico, "Netflix Recommendations: Beyond the 5 Stars (Part 1)," The Netflix Tech Blog, April 6, 2012, < https://medium.com/netflix-techblog>.

In a government context, this might primarily mean cost savings. And relevance can also include the degree to which a solution creates a durable and innovative competitive advantage.[41]

The final dimension of idea quality is elaboration. Elaboration means "the extent of being complete, detailed and well understandable."[42] This is a particularly important concept for government projects because of the turnover of government personnel and the requirement in most cases for any solution to be implemented across multiple agencies.

WORKFLOW MANAGEMENT

Workflow management is the process of soliciting and coordinating inputs to optimize task output. The process has two components, namely definition and supervision. The specific workflow is usually determined by the task requester and is designed with the requirements of the particular organization in mind. It is also possible, however, for the crowd itself to define the workflow. This is the case in such well-known crowdsourcing initiatives as Wikipedia. It should be emphasized, however, that workflow definition is frequently an iterative process requiring frequent feedback and refinement.

A systems mindset is beneficial in managing the workflow. At all times in the process, the balance must be maintained between multiple parameters such as input quality, task-response latency, cost, participation, and complexity. For this reason, a graphical representation of the workflow may be useful both to understand the dynamics of the process as well as a management tool to communicate to the team the information flow and intended outcomes.

In summary, the comprehensive survey conducted by Hetmank identified four major components common to all crowdsourcing systems: user management; task management; contribution management; and workflow management. Each of these components in turn was made up of several functions such as registering users, assigning tasks, evaluating contributions, or managing workflows.

The critical insight when considering crowdsourcing systems and their components is how much the success of the system ultimately relies on management decisions. Each step in a crowdsourcing process needs to be thought through and defined in advance. Given the complexities involved, it is increasingly common for crowdsourcing projects to be hosted on various third party crowdsourcing platforms. In fact, the US government has made extensive use of InnoCentive's platform to support numerous scientific and technical projects.

While it is true that these technical platforms can automate and simplify many management functions, such as workflow definition and user coordination, they do not relieve management of its fundamental responsibilities. These are to determine whether crowdsourcing is appropriate for the

41 Blohm et al., "Does Collaboration Among Participants Lead to Better Ideas in IT-Based Idea Competitions? An Empirical Investigation."
42 Ibid, 5–6.

organizational problem; to align the crowdsourcing process with the organization's mission; and to ensure that any outcomes can be ultimately incorporated into the organization's processes.

INCENTIVES

Perhaps no aspect of the crowdsourcing literature has received more interest than the issue of crowd incentives. This is likely due to the fact that it crosses several disciplines, including psychology, business, economics, and mathematics. For the most part, the practical findings are very consistent. The incentive structure in a crowdsourcing system is very important, and in most cases some type of monetary reward leads to improved performance.[43] Because these basics are so non-controversial, this section will focus on the peculiarities of incentives with respect to government crowdsourcing.

To begin with, users attracted to projects with a perceived public benefit typically show a high degree of intrinsic motivation. This means that users are attracted to these types of projects because they believe the projects will provide opportunity to do something of benefit for their fellow citizens. It is counterintuitive, but offering a monetary incentive for these types of public benefit projects can lead to lower participation and poor outcomes.[44] This is not universally the case, however, and seems to hold mostly for peer-vetted creative production types of projects.[45]

As the complexity of the crowdsourcing task increases, so does the complexity of crowd motivations. Computer programmers, for example, are notorious for their desire to demonstrate competency and seem to be very motivated by public recognition such as leader boards or other types of bragging rights.[46] "Creatives" and other artistic types frequently cite the sense of community as a major motivation in their participation.[47] And scientists and engineers seem to respond best to a combination of all of these things together with reasonable amounts of prize money.[48]

Bringing all of these aspects together is difficult and may be why a recent study from Deloitte University Press was entitled "The Craft of Incentive Prize Design."[49] So many factors overlap that effective incentive design truly is more art than science.

---

43 Mokter Hossain and Ilkka Kauranen, "Crowdsourcing: A Comprehensive Literature Review," Strategic Outsourcing: An International Journal, Vol. 8, No. 1 (2015), pp. 2–22.
44 Gee-Woo Bock and Young-Gul Kim, "Breaking the Myths of Rewards: An Exploratory Study of Attitudes about Knowledge Sharing," Pacis 2001 Proceedings, 2001, p. 78.
45 Daren C. Brabham, "Motivations for Participation in a Crowdsourcing Application to Improve Public Engagement in Transit Planning," Journal of Applied Communication Research, Vol. 40, No. 3 (June 12, 2012), pp. 307–328.
46 Ned Gulley and Karim Lakhani, "The Determinants of Individual Performance and Collective Value in Private-Collective Software Innovation," Harvard Business School Technology & Operations Mgt. Unit Working Paper, No. 10–065 (2010), <http://papers.ssrn.com/sol3/Delivery.cfm?abstractid=1550352>.
47 Daren C. Brabham, "Moving the Crowd at Threadless," Information, Communication & Society, Vol. 13, No. 8 (December 1, 2010), p. 1124.
48 Karim R. Lakhani et al., "Prize-Based Contests Can Provide Solutions to Computational Biology Problems," Nature Biotechnology, Vol. 31, No. 2 (2013), pp. 108–111.
49 Kwasi Mitchell et al., "The Craft of Incentive Prize Design: Lessons from the Public Sector," Deloitte University Press, June 18, 2014, <https://dupress.deloitte.com/dup-us-en/topics/social-impact/the-craft-of-incentive-prize-design.html>.

To be effective, incentives must align with user motivations. As mentioned earlier, intrinsic motivation is often prevalent in crowdsourcing projects with a clear public benefit.[50] Extrinsic motivations, such as financial motivations, are common for human intelligence tasks. However, as task complexity rises, motivations tend to become mixed.[51]

Motivations also shift over time. In a study of user participation in a photography crowdsourcing platform, users reported different reasons for starting to participate than they did for continuing to participate.[52] Initially, users were drawn to the financial reward offered by the platform. Once established, however, they continued to participate out of enjoyment of the process and a sense of social connectivity. The implication is that extrinsic rewards are enough to incite curiosity, but once the curiosity is satisfied, the social aspect is useful and perhaps even necessary in maintaining commitment.

One helpful way to think about structuring incentives is to take a step back and think about user participation in general. Any crowdsourcing process makes demands on users. Heavy demands such as editing comments or managing forums require high amounts of commitment. Light demands, such as simple voting or "Liking" a webpage, require little. It should come as no surprise, then, that these two types of users tend to form groups, with heavy users placing an emphasis on community and light users functioning as part of the crowd.[53]

To review, an effective incentive structure has to accomplish several things. First, it has to balance between intrinsic and extrinsic motivation, while taking into account which type of motivation will have the most appeal for the specific type of project. Next, the incentive structure must be able to encourage initial participation as well as continue to support sustained participation. Ideally, the incentive structure should be able to adjust over time. Last, the incentive design must recognize that different groups of users will have different requirements, such as users who see themselves as just one of the crowd versus users who see themselves as part of a community.

What would such an incentive structure look like? Unfortunately, there are no hard and fast rules. At best, current research can only offer a series of guidelines.

---

50 Mokter Hossain, "Users' Motivation to Participate in Online Crowdsourcing Platforms," paper delivered at the International Conference on Innovation Management and Technology Research (ICIMTR), Malacca, Malaysia, 2012, p. 311.
51 M. Six Silberman, Lilly Irani, and Joel Ross, "Ethics and Tactics of Professional Crowdwork," XRDS, Vol. 17, No. 2 (December 2010), pp. 39–43; Dana Chandler and Adam Kapelner, "Breaking Monotony with Meaning: Motivation in Crowdsourcing Markets," Journal of Economic Behavior & Organization, No. 90 (June 2013), pp. 123–33.
52 Wael Soliman and Virpi Kristiina Tuunainen, "Understanding Continued Use of Crowdsourcing Systems: An Interpretive Study," Journal of Theoretical and Applied Electronic Commerce Research, Vol. 10, No. 1 (January 2015), pp. 9–10.
53 Nama R. Budhathoki and Caroline Haythornthwaite, "Motivation for Open Collaboration: Crowd and Community Models and the Case of OpenStreetMap," The American Behavioral Scientist, Vol. 57, No. 5 (May 2013), pp. 551–556.

With this in mind, we offer six considerations for incentive design for government crowdsourcing projects:

» Curiosity

» Financial compensation

» Recognition

» Community

» Feedback

» Usability

Again, none of these are absolutes. Taken together they can be seen as necessary but insufficient considerations for the incentive part of a crowdsourcing system.

CURIOSITY

An underappreciated aspect of incentive design is the requirement to create tasks that are stimulating and generate curiosity.[54] Interestingly, the link between a stimulating task and monetary reward falls along a U-shaped curve. If the task is simple and dull, users will demand financial compensation. Likewise, if the task is difficult, and especially if it requires a high degree of tacit knowledge, users will again require financial compensation.[55] As most government crowdsourcing projects will fall between these two extremes, the need to create stimulating tasks becomes more important.

One easy way to ensure tasks are stimulating is to focus on the conditions under which the task must be performed, not on the task itself. Several studies have shown the importance of subjective experience in user satisfaction.[56] Increasing user autonomy, interaction, and sense of competence are all ways to increase the amount of task stimulation without changing the task itself.[57]

---

54 Miia Kosonen and Kaisa Henttonen, "Cheer the Crowd? Exploiting Crowdsourcing as a Problem-Solving Strategy," in ISPIM Conference Proceedings (Manchester, United Kingdom: The International Society for Professional Innovation Management (ISPIM), 2013), p. 5, <http://search.proquest.com/docview/1690237887/abstract/42C132170A8A4703PQ/1>.

55 Haichao Zheng, Dahui Li, and Wenhua Hou, "Task Design, Motivation, and Participation in Crowdsourcing Contests," International Journal of Electronic Commerce, Vol. 15, No. 4 (Summer 2011), p. 57.

56 Miia Kosonen et al., "Users' Motivations and Knowledge Sharing in an Online Innovation Community," in ISPIM Conference Proceedings (Manchester, United Kingdom: The International Society for Professional Innovation Management (ISPIM), 2012), pp. 3–5, 12, <http://search.proquest.com/docview/1368545676/abstract/A9CC495784254960PQ/1>; Haichao Zheng, Dahui Li, and Wenhua Hou, "Task Design, Motivation, and Participation in Crowdsourcing Contests," International Journal of Electronic Commerce, Vol. 15, No. 4 (Summer 2011), pp. 61–64, 76–77.

57 Kosonen et al., "Users' Motivations and Knowledge Sharing in an Online Innovation Community," pp. 3–5, 12; Zheng, Li, and Hou, "Task Design, Motivation, and Participation in Crowdsourcing Contests," pp. 60, 76–77, 80.

FINANCIAL COMPENSATION

The financial compensation aspect of incentive design seems straightforward. Common sense would seem to dictate that higher rewards would lead to improved performance, but this is not always the case. Even for human intelligence tasks, researchers have found a complicated connection between money and task performance.[58]

RECOGNITION

Recognition is different than feedback in that it is often nonspecific and can be as simple as accruing points for submissions. Recognition schemes vary, but many researchers have found that common elements from video games translate well into the crowdsourcing environment.[59] Incentives such as leaderboards, badges, and scoring schemes all seem to increase both the quantity and quality of submissions.[60] This conforms to the general expectation that many participants and crowdsourcing systems have a desire to demonstrate competence and find "bragging rights" an important aspect of the overall reward system.[61]

COMMUNITY

The requirement for community incentives differ according to the crowdsourcing objective. For example, there is little need for a community if the type of system is designed for human intelligence tasks. Likewise, a knowledge-discovery management system also exhibits little need for community. Broadcast searches and peer-vetted creative production, however, should explicitly encourage communities in order to optimize performance due to the "heavy" demands on users. This can be done in the task-management phase by structuring the challenge so it is only open to teams. Another method that has been shown to be effective is to tie community participation to the recognition scheme. In other words, participants are rewarded not only for submission of successful solutions, but also for commenting on solutions, total number of posts, and even for providing overall encouragement to the group.[62]

---

58 Winter Mason and Duncan J. Watts, "Financial Incentives and the 'Performance of Crowds,'" SIGKDD Explorations Newsletter, Vol. 11, No. 2 (May 2010), pp. 100, 106–107.

59 Benedikt Morschheuser, Juho Hamari, and Jonna Koivisto, "Gamification in Crowdsourcing: A Review," in 2016 49th Hawaii International Conference on System Sciences (HICSS) (IEEE, 2016), pp. 4375–4384, <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7427729>.

60 Tak Yeon Lee et al., "Experiments on Motivational Feedback for Crowdsourced Workers," in Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, 2013, <http://www.jamierasmussen.com/publications/ibmerswhotweet.pdf>.

61 Maria J. Antikainen and Heli K. Vaataja, "Rewarding in Open Innovation Communities – How to Motivate Members," International Journal of Entrepreneurship and Innovation Management, Vol. 11, No. 4 (January 1, 2010), pp. 440–56; Nikolay Archak, "Money, Glory and Cheap Talk: Analyzing Strategic Behavior of Contestants in Simultaneous Crowdsourcing Contests on TopCoder.Com," in Proceedings of the 19th International Conference on World Wide Web, WWW '10 (New York, NY, USA: ACM, 2010), pp. 21–30.

62 Malhotra and Majchrzak, "Managing Crowds in Innovation Challenges," 120–121.

FEEDBACK

Just as it is in the offline world, feedback is an important component of encouraging participation and high quality results. Feedback is thought to increase intrinsic motivation for tasks and has been shown to have a strong positive relationship with user intent to participate in a crowdsourcing system.[63] The feedback does not always have to be after a user action. Providing clear instructions and guidelines is also thought to increase a user's "internalization" and commitment to the project.[64] Furthermore, feedback helps task managers to emphasize knowledge sharing and integrate knowledge, which leads to better results.[65]

USABILITY

The final consideration in incentive design is usability. Just as advertisers struggle to present attractive and enticing messages that appeal to limited attention spans, crowdsourcing system designers must also create appealing and intuitive interfaces to encourage and maintain participation by users who may have other demands on their time. In his study of motivations for participation in a public design project, Darren Brabham found that both usability and a perception of a low barrier to entry were cited time and again by participants as very important. This is a common insight in the world of website designers, but perhaps underappreciated in government crowdsourcing projects. Indeed, Brabham's conclusion is that "usability and design can make or break a crowdsourcing venture."[66] This insight was confirmed in a separate study of public crowdsourcing projects in South Africa, which found "Easy to use" to be a statistically significant predictor of public intention to participate in a project.[67]

CROWD CURATION

Crowd selection or "curation" is the final aspect of crowdsourcing systems we will review. Returning to our initial overview of how crowdsourcing works, it is important to remember that much of the power in crowdsourcing comes from the intellectual diversity of the crowd. In their famous study of crowdsourced scientific problems involving more than 12,000 scientists, Jeppesen and Lakhaini found that "technical and social marginality, being a source of different perspectives and heuristics, plays an important role in explaining individual success in problem solving. The provision of a winning solution was positively related to increasing distance between the solver's field of technical

63 Richard M. Ryan and Edward L. Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," American Psychologist, Vol. 55, No. 1 (January 2000), pp. 68–70, 76; Cupido and Ophoff, "A Model of Fundamental Components for an E-Government Crowdsourcing Platform," p. 15.
64 Ryan and Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," p. 68.
65 Malhotra and Majchrzak, "Managing Crowds in Innovation Challenges," pp. 104–106, 112, 120.
66 Daren C. Brabham, "Motivations for Participation in a Crowdsourcing Application to Improve Public Engagement in Transit Planning," p. 323.
67 Cupido and Ophoff, "A Model of Fundamental Components for an E-Government Crowdsourcing Platform," pp. 151–152, 154–155, Table 3.

expertise and the focal field of the problem."[68] This does not mean, however, that a randomly selected group will outperform experts at complex problem-solving. Rather, it implies that a well-designed crowdsourcing system allows and encourages people who have achieved expertise in a wide variety of fields to apply their expertise to a new arena. Or, perhaps another way to think about it is that crowdsourcing encourages creativity.

Obviously, the amount of intellectual diversity required to be successful is related to the task at hand. Practically speaking, the task provider must try to decide where the task falls along the spectrum of "open to everybody" versus "open to a few."[69] This is usually done with some type of preselection process. Preselection may be qualification based, requiring the demonstration of certain skills, or context based, which means limiting participation to certain groups such as employees, people with security clearances, or medical doctors, etc.[70]

With the exception of the employee-based crowd, where the employer presumably has some say in mandating participation, crowdsourcing participants must self-identify. That is, users must review the incentives offered and decide whether to participate or not. This is problematic for two reasons. First is the issue of tacit knowledge, especially with complex problems. Tacit knowledge is the accumulation of experience, background knowledge, learned procedures, and other information that is rarely codified and difficult to pass on.[71] Tasks requiring a high degree of tacit knowledge are difficult to crowdsource because users are thought to perceive such tasks as "risky," or exhibiting a low probability of a successful solution.[72]

The second issue related to self-identification is that for complex problems, it is precisely those users least qualified to solve the problem who will assess themselves as having the highest chance of success. This well-known cognitive bias is called the Dunning Kruger effect, named after the researchers who first documented it.[73] Their research found that incompetent individuals routinely overestimate their ability to successfully complete a task, while competent individuals underestimate their ability to do so.

This self-identification paradox, that complex tasks will simultaneously attract users who are either incompetent, or if competent, underestimate their own ability to solve the problem, has no easy solution. There are several ways to help mitigate the problem, however. A good preselection process is the

68 Lars Bo Jeppesen and Karim R. Lakhani, "Marginality and Problem-Solving Effectiveness in Broadcast Search," Organization Science, Vol. 21, No. 5 (October 2010), p. 1016.
69 Feller, Joseph, Patrick Finnegan, Jeremy Hayes, and Philip O'Reilly, "Leveraging 'The Crowd': An Exploration of how Solver Brokerages enhance Knowledge Mobility," ECIS, 2010, pp. 7, 10, <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1152&context=ecis2010>.
70 Geiger, David, Stefan Seedorf, Thimo Schulze, Robert C. Nickerson, and Martin Schader, "Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes," AMCIS, 2011, p. 6.
71 Michael Polanyi, Personal Knowledge: Towards a Post-Critical Philosophy (Chicago: University of Chicago Press, 2015).
72 Zheng, Li, and Hou, "Task Design, Motivation, and Participation in Crowdsourcing Contests," p. 64.
73 Justin Kruger and David Dunning, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," Journal of Personality and Social Psychology, Vol. 77, No. 6 (December 1999), pp. 1121.

first one. Limiting the pool of applicants to those expected to be able to perform is a good first step. Providing clear problem formation and explicit guidance on acceptable solutions will decrease tacit knowledge requirements and is another way to reduce the negative impact of self-identification.

The final way to resolve the paradox is to use a preexisting crowdsourcing system. Amazon's Mechanical Turk, for example, has built in crowd-vetting and preselection procedures that have been tested under real-life conditions against thousands of human intelligence task crowdsourcing projects. Likewise, InnoCentive's crowdsourcing system for broadcast search has a long and successful track record of complex problem-solving using its pool of preregistered scientists and its ability to publicize new challenges through its strong ties to industry and academia. An additional advantage of using a preexisting platform is that it typically has a tested and successful user interface, which ensures this critical component of a successful project is properly incorporated into the crowdsourcing system.

CROWDSOURCING CASES

In choosing examples to demonstrate the uses of crowdsourcing, we have focused on cases which have broad applicability to nonproliferation. Returning to Brabham's typology, this has resulted in a focus on two specific types: cases dealing with knowledge discovery and management and cases dealing with broadcast searching.

According to Brabham, knowledge discovery and management is ideal for "information gathering, organization, and reporting problems, such as the creation of collective resources."[74] This is exactly the type of problem seen in the most well-known example of government crowdsourcing, namely the DARPA Network or "Red Balloon" challenge. It is different from distributed human intelligence tasking in that it is not relying on aggregating data or analyzing it. Problems of this type are similar to the "If You See Something Say Something™" campaign sponsored by the Department of Homeland Security.

With respect to nonproliferation, this type of crowdsourcing problem-solving is also reminiscent of "societal verification" efforts popularized by Seymour Melman in the 1950s and reinvigorated by the surge of interest and capability of social media.[75] The idea here is to use the crowd as a type of sensor, providing distributed information that can then be aggregated for analysis and insight.

The second type of crowdsourcing we will focus on, broadcast search, is best suited for problems with verifiable solutions. This type of crowdsourcing has been used extensively in the scientific community, particularly in the health field. The famous InnoCentive platform has used exactly such an

74 Daren C. Brabham, "Using Crowdsourcing in Government," 2013, p. 11, <http://www.cbs.dk/files/cbs.dk/using_crowdsourcing_in_government.pdf>.
75 Seymour Melman, "How Can Inspection Be Made to Work?," Bulletin of the Atomic Scientists Vol. 14, No. 7 (1958): 270–272, doi:10.1080/00963402.1958.11453870.

approach for over a decade to solve major research problems, including identifying biomarkers or developing new techniques to clean oil spills.[76]

This type of crowdsourcing is relevant for nonproliferation in that it helps to develop new technologies and find novel applications for existing technologies. Again, it differs from distributed human intelligence tasking in the complexity of the problems and also differs from creative production in that the solutions are concrete and not a matter of taste or opinion.

What follows is a brief collection of use-cases that are broadly illustrative of crowdsourcing capabilities. We have been guided by several considerations in selecting the cases. With one exception, we have focused on cases of the government crowdsourcing, that is, crowdsourcing projects that were initiated by government agencies at the federal, state, or local levels.

Our next requirement was that the selected case must be applicable to problems faced in the nonproliferation domain. For this reason, we have emphasized knowledge discovery and management and broadcast search over distributed human intelligence tasking and creative production. That said, the boundaries between varieties of crowdsourcing tasks are not always crystal clear.

Last, we have balanced our cases along examples of success and failure. Although the details of the cases will show a variety of reasons for each, the common thread running through all of them is the strength or weakness of the management process. This admittedly subjective selection process resulted in four cases, two at the federal level, one at the state level, and one hybrid project.

SECURING THE TEXAS BORDER

Trafficking across borders has been a long concern in the nonproliferation arena, but it was the collapse of the Soviet Union in 1991 and later revelations about Pakistan's covert nuclear program that energized concerns of non-state actors acquiring WMD materials or components. The rise of global terrorism means this problem is still very much with us, and the issue of border controls remains high on the nonproliferation agenda.

Governor Rick Perry (Republican of Texas) proposed the use of crowdsourcing in anti-smuggling efforts in 2006. As a border state, Texas was concerned with illegal immigration and human smuggling. Faced with the lack of resources to surveil the 1,200-mile-long border with Mexico, the state decided crowdsourcing could provide a solution. In a press release during his reelection campaign, Perry proposed $5 million to create a virtual border watch program where "cameras

76 Michael Franklin, "The Diverse Uses of Crowdsourcing Competitions," Perspectives on Innovation, April 5, 2016, <http://blog.innocentive.com/2016/04/05/diverse-uses-crowdsourcing-competitions>.

will cover vast stretches of farm and ranchland located directly on the border where criminal activity is known to occur…."[77]

DEFINING THE PROBLEM AND SUCCESS

Conceptually, the virtual border-watch program was described by the executive director of the Texas Border Sheriffs Coalition as "very much like a high-tech neighborhood watch program."[78] The idea of human sensors is a very common one in crowdsourcing, and the effort here appears to be a reasonable use of a crowdsourcing system.[79]

As an example of an information-gathering crowdsourcing effort, project organizers could be expected to design their system with four major considerations in mind. The first of these is management control. Because the project involves citizens in activities that are traditionally considered law enforcement, there is a potential for intentional or unintentional abuse as well as liability and privacy concerns. There is also an obvious risk of bad publicity, which could undermine the project. These considerations lead to a strong requirement for management control and rapid decision making.

Next comes crowd curation. The project must be able to involve a large enough crowd to adequately monitor the large area under consideration twenty-four hours a day, seven days a week. It must also pay attention to the type of crowd it attracts. For example, should a report from an off-duty law-enforcement officer be given higher precedence than one from a regular citizen? Or, should reports from citizens of Texas receive higher priority than the citizens of other states?

Third, the project must also have a clear aggregation mechanism. Given the potential for a large volume of reports, law enforcement would need a means to separate credible reports from less than credible reports. Presumably, this mechanism would take into account any decisions made with respect to the crowd-curation problem as well as make some attempt to account for and fight against cheating and spoofing.

Finally, the project must have a simple and easily understood user interface, which would make the task of identifying and reporting information as simple as possible. This user interface would also have to provide some kind of guidance and examples of the type of information that is needed.

---

77 "Perry Authorizes More Border Security Funding, Virtual Border Watch Program," Government Technology, June 1, 2006, <http://www.govtech.com/public-safety/Perry-Authorizes-More-Border-Security-Funding.html>.
78 Brandi Grissom, "Live, Virtual Border-Watch Program Goes Back Online," El Paso Times, November 20, 2008, sec. News.
79 Weiwei Yuan et al., "Harness Human Sensor Networks for Situational Awareness in Disaster Reliefs: A Survey," IETE Technical Review, Vol. 30, No. 3 (2013), pp. 240–247; Bin Guo et al., "Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm," ACM Computing Surveys (CSUR), Vol. 48, No. 1 (2015), p. 7.

Let us look more closely at management control. Governor Perry initially pledged $5 million dollars for the project, but struggled with legislators to identify specific funding sources, and eventually awarded $2 million to the Texas Border Sheriffs Coalition to initiate the project.[80] There is little indication that anyone in the governor's office or the Sheriffs Coalition had experience with large-scale crowdsourcing projects, and project management quickly contracted with a private social-media company for assistance. Beyond a vague hope that advertising on the user website would generate enough funding to make the program self-sustaining, little thought appears to have been given to the larger questions of identifying specific tasks, aggregating resulting information, feeding that information back into a planning process, or tying the program to broader state law-enforcement objectives. Indeed, an ambitious list of fifteen program goals in 2006 was reduced to five by 2009—with only two related to law enforcement.[81]

As could be expected, this lack of management planning cascaded through the system. Although it describes itself as a social media company, the company selected to host the project–BlueServo.com–did not give much attention to user interface design and initial versions of the website were plagued by technical problems (Figure 1).[82] Many users were unclear about what they were supposed to be looking for, with dozens reporting confusion about which images were people and which were animals.[83] The problem with the user interface was summed up succinctly by one user: "Most of the time it's looking at bare ground." The same user reported that the most interesting thing he ever saw was a bird that stared directly into the camera.[84] Clearly, there was no effort to create a knowledge integration system or allow comments on observations, which would have simultaneously increased user interest and raised the quality of reports. In fact, this opportunity was never even considered as law-enforcement officers did not engage with users. Instead, BlueServo gathered user emails then forwarded them to law-enforcement officials.[85]

80 Brandi Grissom, "Border Sheriffs Coalition Announces Internet Border Watch Program," November 20, 2008, sec. News.
81 Brandi Grissom, "Border Watch Meets Lowered Expectations in Fourth Year," El Paso Times, April 20, 2010.
82 "Images Fuzzy, but Border Cams Online," El Paso Times, November 3, 2006.
83 Brandi Grissom, "Virtual Border System Ineffective, Out of Cash," El Paso Times, July 11, 2009.
84 Ibid.
85 Brandi Grissom, "Live, Virtual Border-Watch Program Goes Back Online," El Paso Times, November 20, 2008.

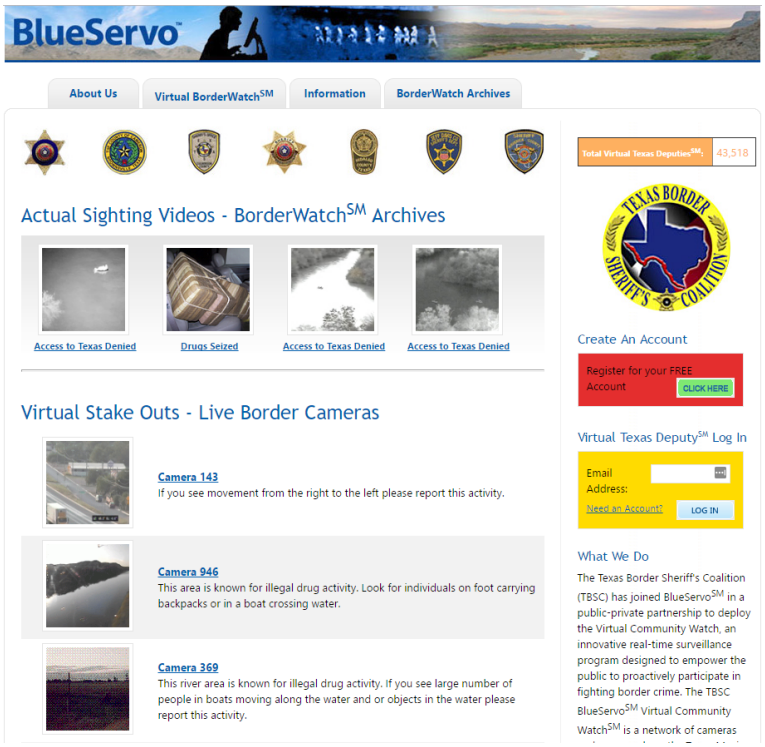

Figure 1 The Texas Border Watch Website

Source: https://web-beta.archive.org/web/20090221120144/http://www.texasborderwatch.com:80/

People participate in government projects primarily for reasons of intrinsic motivation. This means any government crowdsourcing project should attempt to capitalize on this by providing adequate user feedback and engagement. There is little evidence that this was the case for the Texas Virtual Border Watch Project. Users did not receive any indication that their reports were processed or used in an investigation or arrest, and the overall results of the project were only released when members of the press requested the information through open-records requests.[86]

Despite little attention being given to crowd curation, participation on the site was robust. During its first year of operation, it generated approximately two million hits from a global user base (ironically, several in Mexico) with approximately four thousand reports weekly.[87] This was likely due to the national publicity the project generated in the wake of Governor Perry's vocal disagreement with federal immigration policy. Unfortunately, the reports were generally of low quality. One user reported, "Cow or deer walked by; now out of screen," while another cautioned, "Just a word of warning: a moment ago I saw a spider crawl across the top of the camera."[88]

86 Brandi Grissom, "Border Cameras Net 2 Million Hits, 1 Drug Bust, 6 Illegal Entries," El Paso Times, December 17, 2008, sec. News.

87 Brandi Grissom, "Border Cameras Net 2 Million Hits, 1 Drug Bust, 6 Illegal Entries," El Paso Times, December 17, 2008.

88 Grissom, "Virtual Border System Ineffective, Out of Cash."

Reports about the project give no mention of any effort being made to control for false reporting. This is quite surprising, given that a report could presumably lead to an arrest. Rather than attempt to develop a crowd-based system of adjudication, such as user voting or member reputation, the project seemed to rely on the judgment of the individual law-enforcement officer who received the report. While there is no doubt that this is an effective system for weeding out non-credible reports, it is virtually guaranteed to overwhelm officers with irrelevant information.

By 2010, poor system design and ineffectual management led to growing recognition that the project was a failure. When asked about the reported success of the effort—twenty-six arrests and the seizure of about 7,400 pounds of narcotics in two years—Cameron County Sheriff Omar Lucio said simply, "I arrest more than that in a month."[89]

ANALYSIS

In retrospect, it seems obvious the project was doomed to failure from the start. A specific objective was never identified, and no analysis was done to see if the problem was suitable for crowdsourcing. Even assuming crowdsourcing could have provided a reasonable solution, the basic building blocks of such a system were never put into place. There were no specified objectives with respect to crowdsourcing, no effort to identify incentives or a suitable crowd, and no aggregation or voting mechanism to ensure quality results. With respect to the broader question of applicability to nonproliferation, Texas's experience with using crowdsourcing for border security leaves many questions open. A local newspaper editorial perhaps said it best: "We favor the camera system, but now it has come down to a choice: either give the system a thorough test, or ditch it and stop wasting time and money."[90]

RAPID IDENTIFICATION OF DNA SEQUENCES

The challenges of nonproliferation are not limited to nuclear materials. In recent years, technological progress has been most rapid in the fields of biotechnology and genetics. This has naturally led to increased concerns about the use of these technologies to develop evermore deadly biological weapons. In addition, global transportation networks and a burgeoning middle class with a penchant for travel have led to a rise in pandemic outbreaks and the spread of novel or unknown pathogens. And, the use of biological weapons on the battlefield by either state or non-state actors still remains a threat.

Rapid identification of unknown pathogens is essential in effectively responding to this threat. However, in 2012 the Defense Threat Reduction Agency (DTRA) determined progress on this front was falling behind. In order to accelerate research towards a better solution, DTRA turned to crowdsourcing.

89 Grissom, "Border Watch Meets Lowered Expectations in Fourth Year."
90 "EDITORIAL: Border Cameras: They Haven't Had Fair Test," El Paso Times, January 27, 2009.

DEFINING THE PROBLEM AND SUCCESS

At the time, the latest methods of DNA analysis required up to a week to identify the genetic makeup of an organism. Obviously, such a lengthy period of analysis is impractical in fast-moving situations such as a global health crisis or on the battlefield. Additionally, DTRA noted that research efforts at the time were widely scattered geographically, making coordinated efforts to solve the problem more difficult.[91]

Returning again to Brabham's description of crowdsourcing types, this project clearly falls in the broadcast search arena. Broadcast search is best used for complex problem solving where a verifiable solution is known. In this instance, the solution was unknown, but DTRA had a baseline of comparison and was well situated to identify a winning entry based on improvement over current methods. Similar to our previous example, this effort would require four key elements: a clear and detailed description of the problem; identification and recruitment of a qualified crowd (curation); identification of incentives; and a solution aggregation and voting mechanism. Also, similar to our previous example, DTRA contracted with a commercial organization, InnoCentive, to coordinate and execute the crowdsourcing effort. Notably, despite InnoCentive's admirable track record of running scientific crowdsourcing challenges, DTRA management remained fully involved in the project's task specifications, control, and determination of the winning solution.[92]

From the very beginning, the management team had a clear understanding of what would constitute success, and this understanding was translated to the requirements in the challenge description. The challenge area is the field of meta-genomics, which is the study of microbes using DNA sequencing. Participants were given a brief overview that described two objectives:[93]

- A detailed description of the proposed algorithm and how it addresses each technical requirement listed in the detailed description of the challenge. The project documentation should include a well-articulated rationale and be supported by relevant references.
- Source code implementing the proposed algorithm. Include all dependencies, packages, databases, documentation, and information to generate test results by Seeker.

DTRA further clarified in its technical guidance that proposed solutions must generate equivalent identification and characterization performance. In addition, it must achieve this and an identification timeline substantially shorter than is possible with current available technologies. Grant seekers were given nine test data sets for analysis and results were scored based on the organisms identified,

91 InnoCentive, "U.S. Defense Threat Reduction Agency," accessed August 9, 2016, <https://www.innocentive.com/files/node/casestudy/case-study-us-defense-threat-reduction-agency-dtra-challenge.pdf>.
92 Karim R. Lakhani et al., "The Value of Openness in Scientific Problem Solving," Harvard Business School Working Knowledge, February 2007, <http://www.munktellsciencepark.se/wp-content/uploads/2011/05/The-Value-of-Openness-in-Scientific-Problem-Solving..pdf>; "DTRA/SCC-WMD Announces $1 Million Algorithm Challenge," December 13, 2012, <http://www.prweb.com/releases/dtra/algorithm/prweb10233480.htm>.
93 "Identify Organisms from a Stream of DNA Sequences,"InnoCentive, accessed August 9, 2016, <https://www.innocentive.com/ar/challenge/9933138>.

the reads assigned, and the genes identified.[94] A typical solution involves millions of DNA sequence reads which must then be compared to a database containing billions of entries.[95] The magnitude of the challenge was expressed in a biotechnology discussion forum:

"I'm not sure if I missed the date or it was extended, but I still feel—as someone who works in this area—that this is an impossible task.… I would rather see this money be put towards better database design, curation, and availability."[96]

Undoubtedly, this was an opportunity to put the theoretical superiority of crowdsourcing to the test.

InnoCentive has long been known as a leader in the field of scientific problem solving using crowdsourcing and, therefore, already had a well-established pool of scientific talent from which to draw as well as a mature and tested online platform for collaboration and submission of ideas. Identification of the crowd, however, is only one half of the problem. The second half is incentivizing the crowd to participate. Recognizing that both intrinsic and extrinsic motivations are important in motivating crowds, DTRA announced a $1 million cash prize for the winning entry. DTRA also made the award open to all comers, regardless of nationality, due to its previous assessment that showed a geographically scattered research network.

ANALYSIS

The combination of careful problem and task specification, use of an established crowdsourcing platform with a mature user interface, and large monetary incentive resulted in almost 2,800 participants and seventeen completed solutions.[97] As DTRA management correctly surmised, the winning entry came from researchers outside of the United States, with the prize money going to a three-man team from Germany and Singapore. Summarizing the results, DTRA found their solution represented a "dramatic breakthrough in pathogen detection," which reduced the time of identification from days to tens of minutes and can also be performed under field conditions.[98]

FIGHTING SMALL ARMS TRAFFICKING

While security attention is rightfully focused on weapons of mass destruction, global events remind us almost daily that small arms and light weapons also pose a threat. Small arms and light weapons are not only used in terrorism, like the devastating attacks in Mumbai in 2008 and Paris in 2015,

94 Benjamin Buchfink, Daniel H. Huson, and Chao Xie, "MetaScope - Fast and Accurate Identification of Microbes in Metagenomic Sequencing Data," arXiv:1511.08753 [Q-Bio], November 25, 2015, <http://arxiv.org/abs/1511.08753>.
95 Ibid.
96 "News: Innocentive Challenge: Identify Organisms From A Stream Of DNA Sequences And Get $1,000,000," Biostars, accessed August 9, 2016, <https://www.biostars.org/p/70010/>.
97 "U.S. Defense Threat Reduction Agency."
98 ibid.

but are a major factor in civil wars, criminality, and other causes of instability in the world. UN Secretary-General Ban Ki-moon recognized this fact in a Security Council statement when he said, "Deny access to illegal weapons and ammunition, and you deny criminals, armed groups and extremists a central means to perpetrate violence, intimidation and harm."[99]

In 2014, the United Nations brought the Arms Trade Treaty (ATT) into force to regulate trade on conventional weapons.[100] It entered into force on December 24 of that year, and to date it has been ratified by eighty-eight countries.[101] Enforcement is the critical provision of any arms control treaty, and the ATT suffers from several shortfalls in this regard, most clearly the lack of an enforcement body and a centralized database to track weapons movements and monitor activities.[102] Currently, INTERPOL manages the iArms database to track illicit weapons movements, but this database is for the exclusive use of law enforcement. This prevents any deeper analysis of the data by other government agencies, as well as the academic or nongovernmental communities.

To address these shortfalls, the European Union funded iTrace, "an accessible and user-friendly global information management system on diverted or trafficked SALW and other diverted or trafficked conventional weapons and ammunition."[103]

DEFINING THE PROBLEM AND SUCCESS

The iTrace system was designed to address the basic lack of capacity in the international community to combat trafficking in SALW. This problem arises from a series of linked issues: there are few observers in the areas where weapons are trafficked; the information provided by these observers is fragmented; and monitoring and data collection efforts are uncoordinated and do not provide timely or relevant information to policy makers.[104] The goal of iTrace, therefore, is to provide officials with relevant and reliable data upon which to base counter trafficking policy decisions.

Building a usable system would require three things. First, the iTrace team had to create a database capable of containing complete details of known arms transfers, including text data, videos, photographs, and other documentary evidence (Figure 2). It also had to be integrated with the existing INTERPOL criminal transfer database (iArms) and provide a mapping and visualization capability in order to help policymakers identify key trade routes and better allocate resources. Second, the iTrace team had to gather evidence and populate the database. Third and last, the team had to

99 Remarks by UN Secretary-General Ban-ki Moon to the 7442nd UN Security Council Meeting on the "Human Cost of the Illicit Flow of SALW," in New York, NY, May 13, 2015.
100 "The Arms Trade Treaty," United Nations Office for Disarmament Affairs, accessed December 12, 2016, <https://www.un.org/disarmament/convarms/att/>.
101 Ibid.
102 Sourabhi Mukherji, "21st Century Proliferation and Tracking: Tackling Arms Proliferation in the Modern Conflict Landscape," Institute for Defence Studies & Analyses, June 24, 2015.
103 European Council, "Council Decision 2013/698/CFSP," November 25, 2013, <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013D0698>.
104 Ibid.

publicize the new database by conducting outreach to national government and UN officials, non-governmental research organizations, advocacy groups, and the international press.[105]

Unlike the other cases reviewed here, the iTrace effort does not fit neatly into Brabham's framework of crowdsourcing types. Ordinarily, gathering information to populate such a database would likely fall into the knowledge discovery and management category. In this case, however, as the data is intended to be used in an arms control monitoring context, there is a non-negotiable requirement that any monitoring data be completely accurate and wholly credible. Obviously, when dealing with situations that could lead to serious actions by nation-states, such as sanctions, data must come from a trustworthy source that has been established and agreed upon in advance. For this reason, the European Union decided to use a trusted nongovernmental organization devoted to field investigations of small arms trafficking, Conflict Armament Research (CAR), for populating the database. The EU further decided to use an existing database solution rather than designing and coding a new platform to ensure compatibility among government users.[106]

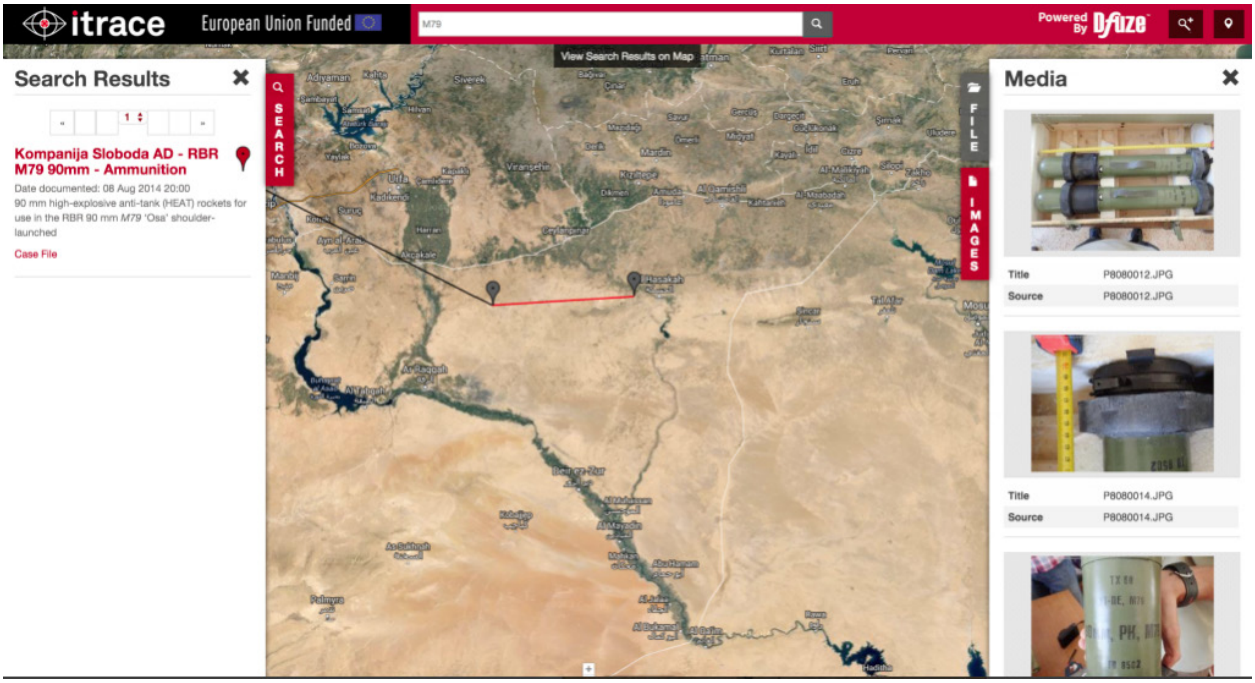

Figure 2 The iTrace Portal

Source: http://www.conflictarm.com/itrace/

105 Ibid.
106 Ibid.

While there was no explicit call for a crowdsourcing effort, the system design document implicitly references crowdsourced outcomes throughout. For example, one of the long-term outcomes is identified as "substantially increase the efficacy of international arms monitoring organizations and individuals by providing an information sharing mechanism of *continually expanding* scope," (emphasis added).[107] Additionally, the document makes frequent reference to the public accessibility of the database and states that the project will "allow any online user to … identify and deliver visual evidence of diverted or trafficked SALW and other conventional weapons and ammunition in that location, country, region or continent."[108] Moreover, the iTrace team is explicitly tasked for "the identification and support of local partners to ensure sustained data collection in support of iTrace throughout the duration of the proposed action and beyond."[109] Its outreach activities were also conceived specifically to "encourage and develop sustainable partnerships with individuals and organizations capable of *generating information that can be uploaded* onto the iTrace system. …" (emphasis added).[110]

As could be expected from the reliance on traditional outreach activities, the iTrace system has not received much use outside of CAR itself. Analysts who have attempted to access the database have noted there are very few entries.[111] Only one of the few international news reports mentioning the system provided insight into who composes the small group of other possible users.[112] Two other news reports focused on an analysis done by CAR itself on the manufacturing origins of ISIS weapons and ammunition. Neither of these reports discussed other cases.[113] Despite the lethargic growth in its user base, the European Union has decided to fund the program for an additional two years.[114]

ANALYSIS

Although iTrace was never designed as an explicitly crowdsourced solution, its execution has implications for the use of crowdsourcing in arms control. With respect to what went right, the iTrace system is indeed seen as an authoritative database by government officials and may paint a way for future efforts of multilateral arms control monitoring. At the same time, the decision to limit data collection to one organization virtually guarantees that the system will be very slow in obtaining data and may never reach the critical mass necessary to support officials in making the informed resource allocation and other policy decisions that it is intended for.

107 "Council Decision 2013/698/CFSP," para. 3 (d).
108 "Council Decision 2013/698/CFSP," para. 4.1.3 (a, b).
109 "Council Decision 2013/698/CFSP," para. 4.2.2 (e).
110 "Council Decision 2013/698/CFSP."
111 Mukherji, "21st Century Proliferation and Tracking: Tackling Arms Proliferation in the Modern Conflict Landscape."
112 "Tracing the World's Weapons, One AK-47 at a Time," Washington Post, June 25, 2014, <https://www.washingtonpost.com/news/checkpoint/wp/2014/06/25/tracing-the-worlds-weapons-one-ak-47-at-a-time/>.
113 C. J. Chivers, "ISIS' Ammunition Is Shown to Have Origins in U.S. and China," The New York Times, October 5, 2014, <https://www.nytimes.com/2014/10/06/world/isis-ammunition-is-shown-to-have-origins-in-us-and-china.html>; C. J. Chivers, "Where the Islamic State Gets Its Weapons," The New York Times, April 27, 2015, <http://www.nytimes.com/2015/04/27/magazine/where-the-islamic-state-gets-its-weapons.html>.
114 European Union, "Council Decision (CFSP) 2015/1908 – Global Reporting Mechanism on Illicit Small Arms and Light Weapons ('iTrace II')," October 22, 2015, <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3A2502_1>.

A well-designed and orchestrated crowdsourcing system could improve this effort. Rather than relying on traditional methods of outreach, such as in-person presentations at meetings, the iTrace team could have incentivized the data-gathering process in the same way that many other government agencies have.[115] As the DTRA case demonstrated, it is entirely possible to recruit from preselected expert communities, thus guaranteeing a desired level of confidence and reliability without sacrificing participation. Beyond data gathering, the iTrace team could have used the crowd to create and publicize research based on the data already collected. It could have also created a series of human intelligence tasks to vet and verify the data, thereby improving its quality and increasing the speed with which it was entered into the database.

Projects like iTrace show the clear potential for crowdsourcing solutions to be applied even in sensitive areas of government cooperation. With the proper management structure and design concepts, tapping into the wisdom of the expert crowd is not only possible, but likely to improve existing efforts.

WIKICRIMES

Longevity is not commonplace among crowdsourcing projects, which tend to be created to solve specific problems at specific points in time. One notable exception is WikiCrimes, an online collaborative project devoted to mapping criminal incidents in South America since 2008.[116] Crime mapping is an established law-enforcement technique to analyze patterns of criminal activity and suggest possible means of intervention.[117] Crime mapping projects face many of the challenges that would be found in a nonproliferation crowdsourcing system. For example, information veracity is extremely important, and witnesses must be protected against retaliation. Also, there is an interplay between witnesses and analysts (detectives), and the evidence is adjudicated by an outside body. For these reasons, crime mapping represents an intriguing case for crowdsourcing.

A common objection raised against efforts to involve the crowd in nonproliferation monitoring is the concern about false reporting.[118] This concern is not exclusive to nonproliferation questions and has also been raised as a factor in disaster response efforts as well as more mundane challenges, such as public design contests.[119] Much of the work examining this question has been theoretical or related

115 Mitchell et al., "The Craft of Incentive Prize Design."
116 Vasco Furtado et al., "Collective Intelligence in Law Enforcement – The WikiCrimes System," Information Sciences, Vol. 180, No. 1 (January 2, 2010), pp. 4–17.
117 Kent A. Harries, "Mapping Crime: Principle and Practice" (Department of Justice, 1999), <https://trid.trb.org/view.aspx?id=654390>.
118 Bryan Lee and Margarita Zolotova, "New Media Solutions in Nonproliferation and Arms Control: Opportunities and Challenges," James Martin Center for Nonproliferation Studies, 2013, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.666.2973&rep=rep1&type=pdf>.
119 Gao et al., "Harnessing the Crowdsourcing Power of Social Media for Disaster Relief"; Brabham, "Motivations for Participation in a Crowdsourcing Application to Improve Public Engagement in Transit Planning"; Hutter et al., "Machiavellianism or Morality."

to the expert-layperson divide.[120] The WikiCrimes system had to address this particular issue along with the traditional questions of crowdsource system design at the outset.

DEFINING THE PROBLEM AND SUCCESS

According to the United Nations Office on Drugs and Crime, Brazil ranks fifteenth in the world for violent crimes.[121] It also ranks poorly on corruption indices, which has led to a situation where the population has low levels of trust in law-enforcement authorities.[122] This is exacerbated by the fact that authorities share little crime information with the public. This has created an environment where a large majority of crimes go unreported because the victims do not feel that this reporting will lead to any resolution.[123]

To address this problem, a team of researchers created the WikiCrimes system with three goals in mind. First, they wanted to create a system with high levels of transparency. Next, they wanted a system based on citizen involvement in crime prevention. And finally, they wanted to create a system that would be reliable enough that it could ultimately address the issue of underreporting.[124]

Returning briefly to Brabham's typology of crowdsourcing systems, the creators of WikiCrimes were dealing with a knowledge discovery and management problem. Somewhat unusual for a crowdsourcing project in the government space, the team also carefully mapped out the functions of their system and aligned those functions with their managerial intent at the start of the process. To achieve the three goals of the project, they had to: activate and motivate the crowd; create an easy to use and intuitive interface; and establish a fraud prevention and detection system that could root out false reporting without penalizing honest but casual users.

As noted earlier, public-sector projects tend to have a motivation advantage over non-public-sector ones. The creators of WikiCrimes took advantage of this fact: "Here, we are appealing to the feelings of sharing that is so usual in victims of violence."[125] However, the team soon realized this would not be enough and organized a multilayered system of incentives to activate the crowd and maintain their participation. This included formal partnership efforts with numerous nongovernmental organizations, media operations, municipal governments, and representatives from businesses, such as insurance agents.

---

120 Andrew J. Flanagin and Miriam J. Metzger, "The Credibility of Volunteered Geographic Information," GeoJournal, Vol. 72, No. 3–4 (2008), pp. 137–148; Linda See et al., "Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts," PloS One, Vol. 8, No. 7 (2013): e69958.
121 UNODC, "Global Study on Homicide" (Vienna, Austria: United Nations, 2013).
122 "Corruption Perceptions Index 2015," Transparency International, 2015, <http://www.transparency.org/cpi2015>.
123 Furtado et al., "Collective Intelligence in Law Enforcement – The WikiCrimes System."
124 Ibid.
125 Ibid, p. 5.

In addition to a traditional advertising campaign, the researchers also invested in social network outreach and integration. They also cleverly tapped into the potential of their information to provide a consumer service by creating a "crime alert" software application that could be used both to receive and to report information. They replicated this service for the media market by providing a website "widget" that online newspapers could embed on their websites as reference material for crime scene reporting. All of these efforts led to high rates of participation, with 100,000 visits from 156 countries within the first eighteen months of operation.[126]

The WikiCrimes researchers opted to create a user interface themselves rather than use an established crowdsourcing platform (Figure 3). This can frequently lead to suboptimal results, but in this case researchers were acutely aware of the importance of the user interface to the success of their project. Using an interface based on the tenets of eXtreme Communication-Centered Design, the team focused on creating a system that was easy to use.[127] This included visual clustering of features on the maps, a multilevel help system, including contextual pop-ups and an online user manual, and a chat function that allows users to receive more specific guidance than what is provided by the written manuals.

126 Ibid.

127 Bruno Santana da Silva, Viviane Cristina Oliveira Aureliano, and Simone Diniz Junqueira Barbosa, "Extreme Designing: Binding Sketching to an Interaction Model in a Streamlined HCI Design Approach," in *Proceedings of VII Brazilian Symposium on Human Factors in Computing Systems*, IHC '06 (New York, NY, USA: ACM, 2006),pp. 101–109; S.D.J. Barbosa, M.G. Paula, C.J.P. Lucena, "Adopting a Communication-centered Design Approach to Support Interdisciplinary Design Teams," in *Proceedings of "Bridging the Gaps II: Bridging the Gaps Between Software Engineering and Human–Computer Interaction*," workshop at the International Conference of Software Engineering, Scotland, May 2004.

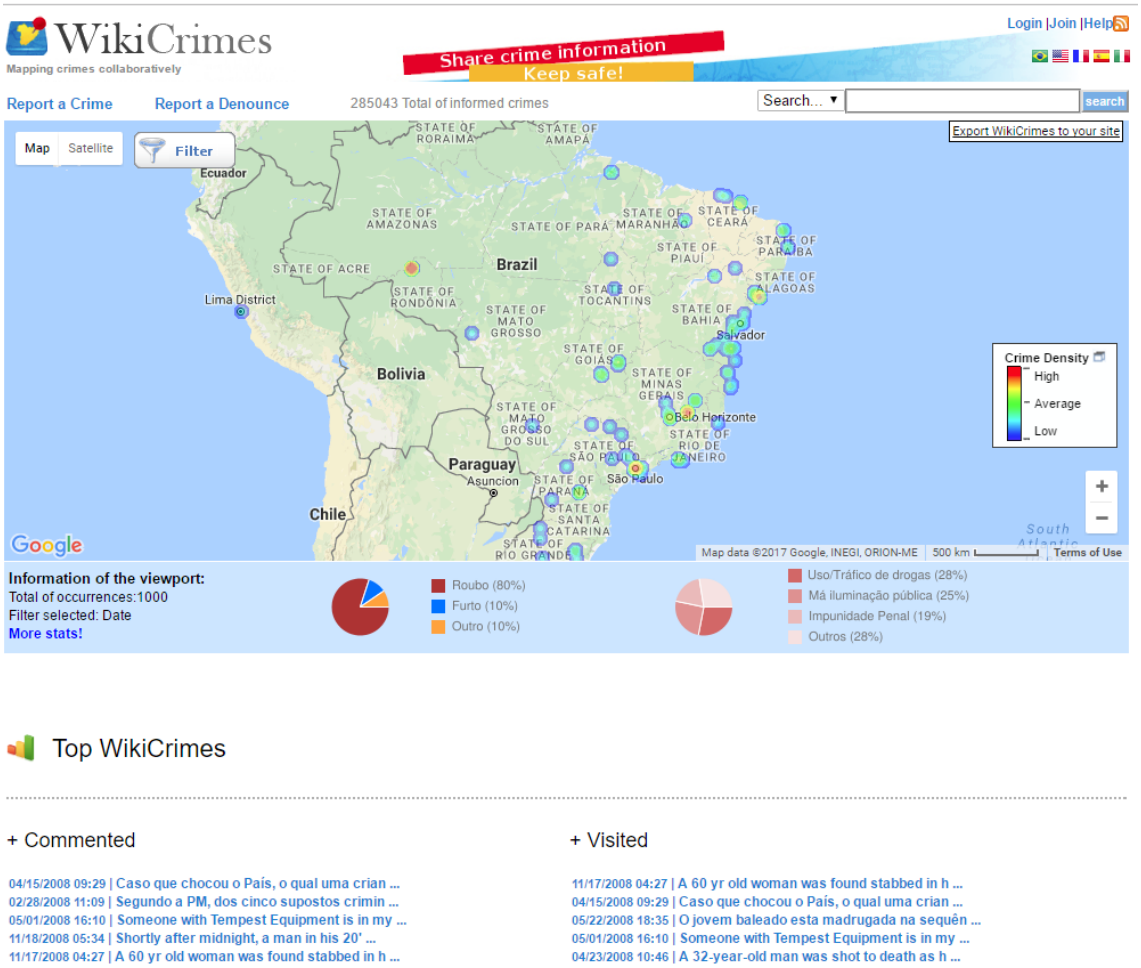

Figure 3 The WikiCrimes Website

Source: http://www.wikicrimes.org/main.html

It is also important to note that the design choices balance the needs of both active and casual users. For example, the contextual help system is mostly geared towards those reporting crimes, while features such as a prominent search window, a newsfeed, and a "hotspot" identification button are clearly geared towards browsers and others only seeking information.

Fraud detection is a critical component of the WikiCrimes system and was fundamental in building a site that is accepted as reliable and credible by users. Restricting access to allow only preapproved users, however, would prevent the bulk of the crime victims from reporting. To solve this conundrum, the WikiCrimes team did two things. First, they designed their site as a "multi-agent" system; that is, the research team created defined roles for each type of user and specified the functions of

each role in advance.[128] Echoing the finding that crowdsourcing systems are management intensive, five of the nine user types were assigned to some kind of administrative role, such as database administration, user registration, or community monitor.[129] Non-administrative users were further broken down into four types: browsers, who only have the ability to review posted information; reporters, who can post information about crime; certified entities, which are credible institutions such as government agencies or media; and invited users who are users who can verify information submitted by reporters. Crucially, non-administrative user registration only required a name and an email address.

The second thing that the WikiCrimes team did to resolve the credibility question is a variation of the traditional crowdsourcing practice of voting. In this instance, however, the team created a reputation algorithm that measured several factors, including adherence to community norms (confirmed by the community monitor agents), whether user reports are verified by an invited outside witness, and the outcome of an analysis of several network analytical measures. These scores were then aggregated to create a dynamic reputation score that is maintained in the site database.[130]

The social-network analytical measures are particularly interesting because of their simple calculation and their power to highlight potentially fraudulent entries. A common property of social networks is the so-called "small world" phenomenon, whereby any node in the network can contact any other node through a small number of intermediary steps.[131] This property gives rise to a network structure with numerous hubs centered around a few key nodes. Not unexpectedly, the super nodes in the WikiCrimes network turn out to be the certified entities. By generating a network graph where users are linked by a common crime report, it is possible to identify suspect entries quickly because those users appear isolated and do not share many links with the super nodes. These suspect reports can then be easily subjected to further review.[132]

ANALYSIS

In evaluating the success of the WikiCrimes system, one can simply ask whether it met its original goals of transparency, citizen involvement, and reliability. Perhaps the clearest indicator of success is that the site is still active today in 2016 and has built up a database of more than 20,000 reported

128 Marcos De Oliveira et al., "Open Collaborative Systems as Institutions of Agents," published in *WI-IAT'08. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008*, Vol. 2, pp. 100–103 (IIEE, 2008), delivered in Sydney, Australia, December 9–12, 2008, <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4740605>.

129 Furtado et al., "Collective Intelligence in Law Enforcement – The WikiCrimes System."

130 Ibid.; Carlos Caminha and Vasco Furtado, "Modeling User Reports in Crowdmaps as a Complex Network," in *Proceedings of 21st International World Wide Web Conference* (Citeseer, 2012), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.309.9555&rep=rep1&type=pdf>.

131 Duncan J. Watts and Steven H. Strogatz, "Collective Dynamics of `Small-World' Networks," *Nature* 393, no. 6684 (June 4, 1998): 440–42.

132 Caminha and Furtado, "Modeling User Reports in Crowdmaps as a Complex Network."

crimes.[133] And, while the level of activity on the site appears to have tapered off, the model has been adopted for use in other countries and regions of the world suffering from similar problems.[134]

The WikiCrimes project has demonstrated that crowdsourcing is a viable method for aggregating and validating sensitive information. It also shows how simple network analysis techniques can be applied to identify misleading or false reports and how common sense information confirmation increases transparency and reliability. Through careful planning and the deliberate alignment of goals and resources, the WikiCrimes system was able to improve the situation with respect to underreported crime and is now under evaluation by the Brazilian Ministry of Justice as a possible tool in its efforts to improve population security.[135]

DISCUSSION

The essential question of the study has been to determine if crowdsourcing is applicable within the field of nonproliferation and arms control. As the case studies demonstrate, the answer would seem to be yes, within reasonable limits. The key to making crowdsourcing work in this nontraditional space is a better understanding of how to structure the effort as well as setting reasonable expectations.

A review of the case studies covered thus far suggests several lessons to be learned. The first is perhaps obvious but bears repeating. Crowdsourcing is not a panacea. There are many challenges and problems that are best approached using traditional means. This is especially the case for challenges in the national-security arena due to the requirements for secrecy and the potential for physical harm to come to participants in the process.

Nevertheless, once managers have determined that crowdsourcing is appropriate, the first consideration must be given to aligning the crowdsourcing system with the organization's mission. By doing this, managers can anticipate potential outcomes and harmonize stakeholder expectations. For example, senior executives may be satisfied with the number of participants a challenge attracts because it highlights the importance of the work done by the agency. Program managers, on the other hand, may be seeking tangible solutions that can be applied to improve program outcomes. Financial managers, however, may be only interested in solutions that substantially reduce costs.

133 WikiCrimes," 2016, <http://www.wikicrimes.org/estatistica.html;jsessionid=206D460101A8B0C894301BFB6383EAB>.
134 Linda Geddes, "Wiki Map Aims to Help Rebuild Shattered Zimbabwe," *New Scientist*, Vol. 201, No. 2700 (2009), pp. 16–17.
135 "Governo Estimula Uso de Tecnologia de Ponta No Combate À Violência," Portal Brasil, October 23, 2015, <http://www.brasil.gov.br/defesa-e-seguranca/2015/10/governo-estimula-uso-de-tecnologia-de-ponta-no-combate-a-violencia>.

Aligning goals and potential outcomes with agency objectives simplifies crowdsourcing management because it provides a roadmap for execution. When all levels of the organization involved understand the intended outcomes and the reasons behind them, it is much easier to design effective incentives to encourage participation from the crowd. Finally, proper alignment ensures a focus on the actual problem, which makes it easier to identify cheating and also to identify and encourage promising solutions.

Once alignment issues are resolved, the next consideration should be incentives. Government efforts should have an advantage here because of the acknowledged importance of intrinsic motivation and the documented connections between this type of user motivation and the desire to assist with public-sector problems. Accordingly, appeals to goodwill and public-sector responsibility can be powerful incentives at all stages of user engagement.

This is not to say, however, that monetary rewards are not important for public-sector actions, but rather they should always be used in some combination with intrinsic appeals. A monetary reward is perhaps most important in the activation phase, when a task owner is attempting to identify as large a crowd as possible and encourage it to participate. Once the reward has been established, intrinsic motivation gains in importance during the maintenance phase of the effort. Feedback from the task sponsor is especially important at this stage and serves the dual purpose of increasing motivation and helping to identify optimal solutions. Feedback from other users, such as "Liking" comments or awarding status badges for community efforts, is another way to maintain levels of participation. Some of this may be automated, but the most successful crowdsourcing efforts are those which have anticipated and prepared for a high level of effort from, and interaction with, the sponsor.

With respect to automation, there is a tendency to ignore the user interface when creating crowdsourcing systems. Crowdsourcing takes place online, and today's users are increasingly sophisticated and demanding. If the online workspace is not in tune with contemporary tastes and expectations, the project will suffer. Fortunately, there are a number of vendors in the private sector that specialize in hosting crowdsourcing efforts. When it is at all feasible, preference should be given to using an established and proven user interface rather than attempting to create something from scratch.

With all of the publicity given to public and private crowdsourcing efforts, it is sobering to realize that the majority of projects on the government's crowdsourcing site, Challenge.gov, are focused on information campaigns and education efforts and only rarely seek solutions to hard technical or policy problems.[136] Part of this, of course, comes from the fact that crowdsourcing is new and much of the implementation is still experimental. But part of it comes from the fact that task owners were frequently unprepared to actually implement solutions derived from crowdsourcing.[137] We saw, for example, in the Texas border watch case that Texas authorities did not have the manpower or other resources available to sort through the hundreds of reports they received. In the same manner, the

136 Ines Mergel et al., "The Challenges of Challenge. Gov: Adopting Private Sector Business Innovations in the Federal Government," in *2014 47th Hawaii International Conference on System Sciences* (IEEE, 2014), 2073–2082, <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6758860>.

137 Kevin C. Desouza, "Challenge. Gov: Using Competitions and Awards to Spur Innovation," *IBM Center for the Business of Government*, retrieved October 17, 2012.

iTrace system has not been of much apparent benefit to its intended user base because data collection has been so limited. Failing to create an implementation plan not only results in a wasted effort, but it hinders future attempts at crowdsourcing because it dampens motivation on both the user and agency sides.[138]

CONCLUSION

Reading through tech company websites or popular science journals, it is easy to get the impression that crowdsourcing is a solved problem. We hope this paper has shown that crowdsourcing is still very much in its infancy and no single technique or method can consistently deliver results. Nevertheless, new does not mean ignored, and there is a growing consensus around the minimum elements required in a successful system.

First are management systems to address the internal functioning of the initiative. This includes functions to address and manage the project's users, objectives, contributions, and overall workflow. Once those are in place and understood, a crowdsourcing system must address the less tangible aspects of incentive design and crowd curation. Unfortunately, there are no set rules here, and the ideal combinations of incentives and users are likely to be unique for each project.

That is not to say that managers are required to start from scratch each time. The US government in particular has embraced the crowdsourcing concept and is starting to generate its own set of principles and guidelines. Efforts such as NASA's Center of Excellence for Collaborative Innovation (CoECI) will likely continue to bear fruit as formal efforts to improve performance percolate through the government. In the meantime, observational studies, such as those described in this paper, provide a useful shortcut to best practices for design and management.

The four case studies described here demonstrate that putting these best practices into use is often difficult. This is especially the case in fields such as nonproliferation and arms control, which rely on multiple sources of information, interagency and international cooperation, and high degrees of trust and reliability. Success in this arena is clearly possible but will require an even higher degree of management involvement than is typical in the average crowdsourcing project. We hope this paper has identified the tools and concepts to guide future crowdsourcing project managers in that task.

138 Ibid.

OCCASIONAL PAPERS AVAILABLE FROM CNS

online at http://nonproliferation.org/category/topics/cns_papers

#33 **Evaluating WMD Proliferation Risks at the Nexus of 3D Printing and Do-It-Yourself (DIY) Communities**
Robert Shaw, Ferenc Dalnoki-Veress, Shea Cotton, Joshua Pollack, Masako Toki, Ruby Russell, Olivia Vassalotti, Syed Gohar Altaf • 2017

#32 **Taiwan's Export Control System: Overview and Recommendations**
Melissa Hanham, Catherine Dill, Daniel Salisbury, P. Alex Kynerd, Raymond Wang • 2017

#31 **Revisiting Compliance in the Biological Weapons Convention**
James Revill • 2017

#30 **Crowdsourcing Systems and Potential Applications in Nonproliferation**
Bryan Lee • 2017

#29 **The Verification Clearinghouse: Debunking Websites and the Potential for Public Nonproliferation Monitoring**
Bryan Lee, Kyle Pilutti • 2017

#28 **Geo4nonpro.org: A Geospatial Crowd-Sourcing Platform for WMD Verification**
Melissa Hanham, Catherine Dill, Jeffrey Lewis, Bo Kim, Dave Schmerler, Joseph Rodgers • 2017

#27 **Searching for Illicit Dual Use Items in Online Marketplaces: A Semi-Automated Approach**
Bryan Lee, Margaret Arno, Daniel Salisbury • 2017

#26 **2016 Symposium Findings on Export Control of Emerging Biotechnologies**
Steven Fairchild, Caroline R. M. Kennedy, Philippe Mauger, Todd J. Savage, Raymond A. Zilinskas • 2017

#25 **Outlawing State-Sponsored Nuclear Procurement Programs & Recovery of Misappropriated Nuclear Goods**
Leonard S. Spector • 2016

#24 **Strengthening the ROK-US Nuclear Partnership**
Miles A. Pomper, Toby Dalton, Scott Snyder, Ferenc Dalnoki-Veress • 2016

#23 **Replacing High-Risk Radiological Materials**
George M. Moore, Miles A. Pomper • 2015

#22 **A Blueprint to a Middle East WMD Free Zone**
Chen Kane, PhD • 2015

#21 **Biotechnology E-commerce: A Disruptive Challenge to Biological Arms Control**
Raymond A. Zilinskas, Philippe Mauger • 2015

#20 **Countering Nuclear Commodity Smuggling: A System of Systems**
Leonard S. Spector, Egle Murauskaite • 2014

#19 **Alternatives to High-Risk Radiological Sources**
Miles Pomper, Egle Murauskaite, Tom Coppen • 2014

#18 **Stories of the Soviet Anti-Plague System**
Casey W. Mahoney, James W. Toppin, Raymond A. Zilinskas, eds. • 2013

#17 **Ugly Truths: Saddam Hussein and Other Insiders on Iraq's Covert Bioweapons**
Amy E. Smithson, PhD • 2013

#16 **Rethinking Spent Fuel Management in South Korea**
Ferenc Dalnoki-Veress, Miles Pomper, Stephanie Lieggi, Charles McCombie, Neil Chapman • 2013

Older Papers

#15 Engaging China and Russia on Nuclear Disarmament • *2009*

#14 Nuclear Challenges and Policy Options for the Next US Administration • *2009*

#13 Trafficking Networks for Chemical Weapons Precursors: Lessons from the 1980s Iran-Iraq War • *2008*

#12 New Challenges in Missile Proliferation, Missile Defense, and Space Security • *2003*

#11 Commercial Radioactive Sources: Surveying the Security Risks • *2003*

#10 Future Security in Space: Commercial, Military, and Arms Control Trade-Offs • *2002*

#9 The 1971 Smallpox Epidemic in Aralsk, Kazakhstan, and the Soviet Biological Warfare Program • *2002*

#8 After 9/11: Preventing Mass-Destruction Terrorism and Weapons Proliferation • *2002*

#7 Missile Proliferation and Defences: Problems and Prospects • *2001*

#6 WMD Threats 2001: Critical Choices for the Bush Administration • *2001*

#5 International Perspectives on Ballistic Missile Proliferation & Defenses • *2001*

#4 Proliferation Challenges and Nonproliferation Opportunities for New Administrations • *2000*

#3 Nonproliferation Regimes at Risk

#2 A History of Ballistic Missile Development in the DPRK • *1999*

#1 Former Soviet Biological Weapons Facilities in Kazakhstan: Past, Present, and Future • *1999*

nonproliferation.org

Middlebury Institute *of*
International Studies at Monterey
James Martin Center for Nonproliferation Studies