

## **SANDIA REPORT**

SAND201X-XXXX

Unlimited Release

Printed October 2017

# **FY2017 Final Report Power of the People: A Technical, Ethical and Experimental Examination of the Use of Crowdsourcing to Support International Nuclear Safeguards Verification**

Zoe N. Gastelum  
Kari Sentz  
Meili C. Swanson  
Cristina Rinaudo

Prepared by  
Sandia National Laboratories  
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.



**Sandia National Laboratories**

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

**NOTICE:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from  
U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831

Telephone: (865) 576-8401  
Facsimile: (865) 576-5728  
E-Mail: [reports@osti.gov](mailto:reports@osti.gov)  
Online ordering: <http://www.osti.gov/scitech>

Available to the public from  
U.S. Department of Commerce  
National Technical Information Service  
5301 Shawnee Rd  
Alexandria, VA 22312

Telephone: (800) 553-6847  
Facsimile: (703) 605-6900  
E-Mail: [orders@ntis.gov](mailto:orders@ntis.gov)  
Online order: <https://classic.ntis.gov/help/order-methods/>



SAND201X-XXXX  
Printed October 2017  
Unlimited Release

# **FY2017 Final Report**

## **Power of the People: A Technical, Ethical and Experimental Examination of the Use of Crowdsourcing to Support International Nuclear Safeguards Verification**

Zoe N. Gastelum  
Meili C. Swanson  
International Safeguards & Engagements  
Sandia National Laboratories  
P. O. Box 5800  
Albuquerque, New Mexico 87185-MS1371

Kari Sentz  
Christina Rinaudo  
Risk Analysis and Decision Support Systems  
Los Alamos National Laboratory  
P.O. Box 1663  
Los Alamos, NM 87545

### **Abstract**

Recent advances in information technology have led to an expansion of crowdsourcing activities that utilize the “power of the people” harnessed via online games, communities of interest, and other platforms to collect, analyze, verify, and provide technological solutions for challenges from a multitude of domains. To related this surge in popularity, the research team developed a taxonomy of crowdsourcing activities as they relate to international nuclear safeguards, evaluated the potential legal and ethical issues surrounding the use of crowdsourcing to support safeguards, and proposed experimental designs to test the capabilities and prospect for the use of crowdsourcing to support nuclear safeguards verification.

## **ACKNOWLEDGMENTS**

This work was sponsored by the Office of Nonproliferation and Arms Control (NPAC, NA-24), Office of International Nuclear Safeguards Concepts & Approaches portfolio. Thank you to Melissa Einwechter for her support of this work.

Tucker Boyce (formerly SNL) supported early research in support of this project and Laura Matzen (SNL) contributed to experimental design brainstorming.

## TABLE OF CONTENTS

1.	Introduction to the Technical Examination of the Use of Crowdsourcing .....	11
2.	Crowdsourcing Safeguards .....	12
3.	Undertaking a Crowdsourcing Experiment .....	13
3.1.	Planning Considerations for Safeguards Crowdsourcing Experiment.....	13
3.2.	The Importance of Validation and Verification of Information .....	17
4.	Conclusions and Future Work.....	18
5.	Introduction to the Legal and Ethical Examination.....	20
6.	Legal Collection of Crowdsourced Data .....	21
6.1.	93+2 Part I Measures .....	21
6.2.	Terms of Service from Individual Platforms.....	23
7.	Legal Use of Crowdsourced Data .....	24
8.	Ethical Collection of Crowdsourced Data .....	25
8.1.	Respect for persons .....	25
8.2.	Beneficence .....	25
8.2.1.	Do No Harm .....	26
8.2.2.	Maximize Benefits and Minimize Harms .....	26
8.3.	Justice .....	27
9.	Ethical Use of Crowdsourced Data .....	28
10.	Other Best Practices.....	29
10.1.	Independent Verification.....	29
10.2.	Protecting Sensitive Data .....	29
10.3.	Avoiding Undue Burden .....	30
10.4.	Protection of IAEA Interests .....	30
10.5.	Optics .....	30
11.	Conclusions and Future Work.....	31
12.	Introduction to Experimental Design .....	33
13.	Experiment 1 (SNL): “Tag the Tower” – Classification of Cooling Tower Images.....	34
13.1.	Research Question .....	34
13.2.	Methodology.....	34
13.3.	Platform.....	34
13.4.	Data Collection and Analysis Plan .....	34
13.5.	Participants .....	35
13.5.1.	Recruitment .....	35
13.5.2.	Consent.....	35
13.5.3.	Privacy .....	35
14.	Experiment 2 (SNL): Personnel Patterns of Life .....	36
14.1.	Research Question .....	36
14.2.	3.2 Methodology.....	36
14.3.	3.3 Platform .....	36
14.4.	3.4 Data Collection and Analysis .....	36
14.5.	Participants .....	37

14.5.1.	Recruitment .....	37
14.5.2.	Consent.....	37
14.5.3.	Privacy .....	37
15.	Experiment 3 (SNL): Audio and Video Transcription .....	38
15.1.	Research Question .....	38
15.2.	Methodology.....	38
15.3.	Platform.....	38
15.4.	Data Collection and Analysis .....	38
15.5.	Participants .....	38
15.5.1.	Recruitment .....	38
15.5.2.	Consent.....	39
15.5.3.	Privacy .....	39
16.	Experiment 4 (LANL): Eliciting Topics: Concepts and Key Words .....	40
16.1.	Research Question .....	40
16.2.	Methodology.....	40
16.2.1.	Gamified Version.....	40
16.2.2.	Regular Version .....	40
16.3.	Platform.....	40
16.4.	Data Collection and Analysis .....	41
16.5.	Participants .....	41
16.5.1.	Recruitment .....	41
16.5.2.	Consent.....	41
16.5.3.	Privacy .....	41
17.	Experiment 5 (LANL): Establishing Relationships between Concepts and Key Words for Knowledge Structuring .....	42
17.1.	Research Question .....	42
17.2.	Methodology.....	42
17.3.	Platform.....	42
17.4.	Data Collection and Analysis .....	42
17.5.	Participants .....	42
17.5.1.	Recruitment .....	42
17.5.2.	Consent.....	43
17.5.3.	Privacy .....	43
18.	Experiment 6 (LANL): Validation of the Knowledge Structure .....	44
18.1.	Research Question .....	44
18.2.	Methodology.....	44
18.3.	Platform.....	44
18.4.	Data Collection and Analysis .....	44
18.5.	Participants .....	44
18.5.1.	Recruitment .....	44
18.5.2.	Consent.....	44
18.5.3.	Privacy .....	45
19.	Next Steps .....	46
References	47	



## EXECUTIVE SUMMARY

This report reflects research conducted by Sandia National Laboratories (SNL) and Los Alamos National Laboratory (LANL) in assessing the feasibility of incorporating crowdsourcing to support international nuclear safeguards verification activities. The report was written in three parts, each part serving as a mid-term report on the evaluation of the use of crowdsourcing:

- Part 1 was led by LANL, and examines a taxonomy of crowdsourcing activities, breaking down into hierarchical structure related to purpose, players, type, motivation, and disclosures.
- Part 2 was led by SNL, and examines legal and ethical considerations of crowdsourcing activities in the context of international safeguards, examining from the perspectives of both collecting and using crowdsourced data.
- Part 3 was jointly written, with each lab developing its respective crowdsourcing for safeguards experiments. LANL experiments focused on expert communities, and the SNL experiments focused on non-experts. Each lab took a different perspective on administering the experiments, with SNL focusing on micro-tasks completed via online platforms, and the LANL team focusing on in-person engagement and gamification.

## NOMENCLATURE

<b>Abbreviation</b>	<b>Definition</b>
BOG	Board of Governors
CSA	Comprehensive Safeguards Agreement
FY	Fiscal Year
IAEA	International Atomic Energy Agency
LOF	Locations Offsite a Facility
SDT	Self Determination Theory
TOS	Terms of Service
VOA	Voluntary Offer Agreement

PART I  
A Technical Examination of the Use of Crowdsourcing to  
Support International Nuclear Safeguards Verification Activities

## **1. INTRODUCTION TO THE TECHNICAL EXAMINATION OF THE USE OF CROWDSOURCING**

While crowdsourcing seems like a recent phenomenon, the idea of crowdsourcing has been around for hundreds of years. Starting in the late 1500s major European seafaring nations were offering large cash prizes to inspire a solution to finding longitude at sea. The prize that was finally claimed was posted by the British Parliament for what would be close to \$20 million today after nearly two thousand sailors were lost when four British warships ran aground. The successful claimant was John Harrison, an English clockmaker for his marine timekeeper H4 in 1759 [Sovel (1994)]. In the 19<sup>th</sup> century, US Naval Officer Matthew Fontaine Maury provided free wind and current charts to sailors on the condition of the return of standardized logs of their voyages so that he could collect the experience of different navigators in different seasons and different vessels traveling the same routes that could serve as guide to future navigators. [Pinsel (1981)] In the last decade, crowdsourcing has become an emergent technology because of the recognition of the value of human sourced data, analysis, or ingenuity and the coincidence with technological enablers such as the internet, mobile technology, social media, and ways to motivate crowds.

Because of the enormous resources it can bring to wicked and data impoverished problems, crowdsourcing technology attracts many researchers in the safeguards community [Lee, Zolotova (2013); Hartigan, Hinderstein (2013); Gerami (2013); Hinderstein et al. (2014)]. We all come with the question of how might we leverage the power of the people in crowdsourcing for nuclear safeguards? In this midterm report, we explore how to design a crowdsourcing experiment that satisfies safeguards objectives and goals, protects nuclear industry professionals, crowd participants, and stakeholders and addresses the standards of quality necessary for safeguards verification.

## 2. CROWDSOURCING SAFEGUARDS

We start with the high-level objectives of the International Atomic Energy Agency (IAEA) with regards to the implementation of safeguards. As summarized in [Board of Governors, IAEA (2014)], the generic objectives of safeguards are\*:

*For States with [comprehensive safeguards agreements] CSAs:*

- *To detect any diversion of declared nuclear material at declared facilities or locations outside facilities where nuclear material is customarily used (LOFs)*
- *To detection an undeclared production or processing of nuclear material at declared facilities or LOFs; and*
- *To detect any undeclared nuclear material or activity in the State as a whole.*

*For States with item-specific safeguards agreements:*

- *To detect any diversion of nuclear material subject to safeguards under the safeguards agreement; and*
- *To detect any misuse of facilities and other items subject to safeguards under the safeguards agreement.*

*For States with [voluntary offer agreements] VOAs:*

- *To detect any withdrawal of nuclear material from safeguards in selected facilities or parts thereof, except as provided for in the agreement*

We can imagine this in terms of more specific detections such as discussed in [Hinderstein, Hartigan (2012)] looking for indicators of:

- Acquisition of or attempts to acquire specialized equipment
- Acquisition of or attempts to acquire materials through trade or diversion
- Transportation of specialized equipment and/or materials
- Production of fissile material
- Manufacturing of warhead
- Preparations for a nuclear test or missile launch
- Nuclear test or missile launch

With the expansion of the use of nuclear technologies for peaceful purposes, the opportunities for diversion are increased and the challenge of distinguishing declared from undeclared activities becomes more arduous.

---

\* taken verbatim with acronyms explained, p.4

### 3. UNDERTAKING A CROWDSOURCING EXPERIMENT

We characterize this as a *crowdsourcing experiment* to highlight the meaningful commonalities between these activities: both involve aspects that are controlled and uncontrolled and the management, the outcomes, and positive and negative consequences are very much determined by the tensions between these two types of factors. Experimentalists call out these factors explicitly in design of a scientific experiment and accordingly, a crowdsourcing investigator should carefully consider what is controllable and uncontrollable in the experimental setting and the range of possible consequences up front. In any crowdsourcing experiment involves key steps [from Grier (2017)]:

1. Design the job in accordance with the goal and objectives and identify the crowd selection criteria
2. Write clear instructions
3. Choose a platform to serve as the crowdmarket
4. Release the job and recruit the crowd
5. Listen to the crowd and manage the job
6. Assemble the work and create the final product

While certainly these steps figure into the planning of a safeguards design of a crowdsourcing experiment, they fall short of the full breadth of things we need to consider.

#### 3.1. Planning Considerations for Safeguards Crowdsourcing Experiment

To help guide the design of a crowdsourcing experiment, we provide a framework summarized in Table 1. This is a taxonomy for crowdsourcing experimental design that endeavors to walk through different aspects of the planning considerations with an emphasis on specific concerns for safeguards. We contrast this with previous efforts in more generally applicable crowdsourcing taxonomies such as [Rouse (2010)] that organizes crowdsourcing by distribution of benefits and capabilities as well as [Cullina et al. (2015)] that are focusing on metrics.

Table 1: Taxonomy for the Design of a Safeguards Crowdsourcing Experiment

<b>PURPOSE</b>				
<b>Goal</b>				
<b>Objectives</b>				
<i>Data Collection or Content Generation</i>	<i>Data Labeling or Analysis</i>	<i>Data or Analysis Verification/Validation</i>	<i>Technological Solution</i>	
<b>Tasks</b>				
<i>Crowdcontests</i>	<i>Self-organized crowds</i>	<i>Macrotasks</i>	<i>Microtasks</i>	<i>Crowdfunding</i>
<b>CROWDSOURCING TYPE</b>				
<i>Active</i>			<i>Passive</i>	
<b>POTENTIAL PLAYERS</b>				
<b>Stakeholders</b>				
<i>IAEA</i>	<i>State</i>			<i>Other</i>

Crowdsourcing Experiment Administrator				
IAEA	State	Other	Third Party	
<i>General</i>		<b>Crowd Selection</b>		<i>Expert</i>
<b>CROWD MOTIVATION</b>				
<b>Innate Needs</b>				
<i>Competence</i>	<i>Sense of Belonging</i>		<i>Psychological Relatedness</i>	
<b>Main Motivators</b>				
<i>Extrinsic</i>		<i>Intrinsic</i>		
<i>Quid pro quo</i>	<i>Peer Recognition</i>	<i>Altruism</i>	<i>Challenge</i>	<i>Enjoyment</i>
<b>DISCLOSURE CHOICES</b>				
<i>Disclose the Goal</i>	<i>Disclose the Objective</i>	<i>Identification of Crowd Participants</i>	<i>Identification of the Stakeholder</i>	

We think of the **purpose** as the reason for the crowdsourcing experiment and the **goal** as what we are trying to achieve with the crowdsourcing experiment. The **objectives** are concrete actions in how the goal will be achieved. Here we identify four types of crowdsourcing objectives: *data collection or content generation*; *data labeling or analysis*; *data or analysis verification/validation*; and a *technological solution*.

*Data collection or content generation* can take many conceivable forms such as the collection of images, generation of text, generation of audio, collection of radioactivity data, seismic data, environmental data, etc. With the increasing sophistication of mobile devices, the reduction in footprint of memory stores as well as the diminishing size of sensors, the amount and type of data that can be collected and transmitted continues to increase.

*Data Labeling or Analysis* is another activity that can take on a wide range of possible forms. Examples include labeling and searching image data, interpreting context, fusing disparate sources of information, anomaly detection, change detection, same detection.

*Data or analysis verification/validation* may be the most important task a crowd can do for building confidence in both the experimental process and the outcome. This can stand-alone or be built into a *data collection* task or *data labeling/analysis* task. For example, many crowdsourcing experiments will require corroboration of data collects or findings.

*Technological solution* crowdsourcing is also a very common and very successful form of crowdsourcing as metrics of success are frequently clear cut and verifiable through testing. In addition to challenges like the longitudinal problem in the introduction or the IAEA Technology Challenge [Createc (2016)] Open source code libraries are common example of successfully crowdsourced technical solutions.

The specific **tasks** are what we ask the crowdworkers to do and frame how to do it. We include the tasks and descriptions as suggested by Grier in [Grier (2017)]. While these are broader than might apply to safeguards we include them here for comprehensiveness:

*Crowdcontests*: Challenge or single job description with many people proposing or answering the challenge and competing for a singular reward

*Self-organizing crowds*: A crowdcontest where the crowd organizes itself into a team and teams compete for a singular reward

*Macrotask*: A specialized single task that can be done independently in a fixed amount of time that requires special skills of a worker

*Microtask*: A small or simple task, often a part of a larger more complicated job where members of the crowd can do tasks and all are rewarded.

*Crowdfunding*: Ways of raising funds through crowdsourcing, often for humanitarian purposes or venture capitalism.

All of these previous tasks presume a specific **type of crowdsourcing** is being invoked, namely *active crowdsourcing* where the stakeholder or administrator plays an active role by posing a particular problem, soliciting information or solutions etc. Another type of crowdsourcing is *passive crowdsourcing* where the stakeholder or administrator assumes a passive role and collects and analyzes content on a specific topic that has been freely generated by citizens in various sources [Loukis, Charalabidis (2015)].

We deliberately separate the roles of **Stakeholders** and **Administrators** though often the players are conflated. The **Stakeholder** is the party who wants the product of the experiment and formulates the **Purpose** and the **Goal**. The **Stakeholder** can be an agency such as the *IAEA*, a *State*, or *Other* organizations that work in the area of safeguards. The **Administrator** is the party who performs and manages the crowdsourcing experiment. It can be useful to separate these parties for the purpose of logistics and qualifications but also public perception. For example a *State* can perform a crowdsourcing experiment on itself and share the results with other **Stakeholders** such as the *IAEA*. In the *Passive* crowdsourcing example, a commercial social media entity can perform a crowdsourcing experiment with no goal of safeguards but obtains relevant information that is freely available and of interest to a *State* or the *IAEA*. This relationship can be captured by the **Stakeholder**: *State* or *IAEA* and **Administrator**: *Third Party*. Second to the objective and task formulation, **Crowd Selection** is one of the more critical choices to the management and outcome of the experiment as it impacts the potential number of participants, the scale of the project to manage but also the quality of the task execution and the need for *post hoc* validation and verification. Here we consider three different options: *General* crowd (unrestricted); *Proximal* crowd (the people close in proximity to a location of interest); *Expert* crowd (people with specialized knowledge).

With *General* crowd recruitment and compelling motivation, we can see enormous numbers of people worldwide coming together to solve problems. For example, the Tomnod online search party for the Malaysian airline flight MH370 recruited 2.3 million people who scanned every pixel of 750,000 images at least 30 times within 5 days [Fishwick (2014)]. While this effort was not successful in recovering the lost airline, it is

an astounding crowdsourcing design involving huge numbers of crowdworkers performing microtasks with multiple validations quickly.

*Proximal* crowds require either a pre-selection criteria for participation, a geolocation filter on voluntarily collected data, or a voluntary disclosure of location through mobile devices. Good examples include Safecast (<http://blog.safecast.org/>) where radioactivity data is geolocated and mapped or the use of the Ushahidi platform to aid in emergency response after the Haitian earthquake in 2010 [Gerami (2013)].

*Expert* crowds will also require a pre-selection criteria or a success in a specialized task, competition, or challenge. While the expertise may help alleviate the burden on verification of results, that burden may simply be shifted to the assessment of expertise up front.

**Crowd Motivation:** Sustaining crowd participation and interest over time is critical to a successful crowdsourcing experiment. This requires analysis into what motivates the crowd to engage in crowdsourcing activities and how to create an incentive mechanism to attract participants. A useful framework to understanding people's willingness to participate in crowdsourcing comes from Self Determination Theory (SDT), formulated by Ryan and Deci (1985). SDT has been used extensively in gamification research and game design and frames motivation as one of the fundamental bases of human behavior: it starts with a particular need and activates a consequent behavior aimed to reach a goal. Deci and Ryan describe three innate and psychological Needs: *Competence* (being competent in performing a task), *Autonomy* (being in control of own behaviors and goals) and *Psychological Relatedness* (experiencing a sense of belonging). When people experience these innate needs, they become self-determined and intrinsically motivated to pursue certain behaviors.

We identify two main motivators: *Intrinsic* and *Extrinsic*:

*Intrinsic* motivators are innate to humans and refer to the natural predisposition to explore, learn, and master new skills and abilities that are essential to cognitive and social development and the satisfaction provided by completing a task.

*Extrinsic* motivators are related to attaining a goal (a promotion), or some kind of external outcome (a monetary reward) and are not usually related to the satisfaction derived from completing an activity.

According to [Roberts et al. (2006)] internal and external motivation likely interact and both affect participation. For these reasons, a crowdsourcing experiment may carefully design intrinsic and extrinsic motivators in combination and perhaps to include changes over time in an attempt to sustain engagement. Categories of ways of motivating such as *quid pro quo* (rewards or compensation), *peer recognition* through point systems, contests and *challenges*, gamification, *altruism*.

The next planning consideration in the taxonomy deals with the choices in **Disclosure** that we explore from the perspective of the **Goal**, the **Objective**, the **Stakeholder** and the **Crowd**

**Participants.** The point over disclosures is to capture in the idea of “experimenter effect” that the knowledge of the goal of the experiment or the **Stakeholder** of the experiment can bias the outcome in both positive and negative ways. There may be an advantage to non-disclosed **Goals** and non-disclosed **Stakeholders** administrated through commercial third parties just to circumvent the disinformation that could be imagined in the fully disclosed case. The concern over disclosing the **Crowd Participants** relates to their protection. The anonymity of the crowd can yield better information on one hand, it allows for disinformation with impunity on the other.

### **3.2. The Importance of Validation and Verification of Information**

The challenge of crowdsourcing is the validation and verification of information and the discrimination of valid information from disinformation and misinformation. We have seen an example of built-in validation in the Tomnod Malaysian airliner search of cooperation and consensus over many crowdworkers before an area was tagged for follow up. We also identify *Data or Analysis Verification/Validation* as its own crowdsourcing task. Hinderstein et al. (2014) voice optimism on the successes of the private sector with Amazon Turk with feedback loops and Wikipedia moderators as well as tradecraft on vetting data and detecting bias. However with the number of intentional disclosures of disinformation seen over the internet in the last year, we can only expect that the validation and verification of information culled from crowdsourcing will be harder and previously successful methods may have to be evolved.

#### **4. CONCLUSIONS AND FUTURE WORK**

In this Midterm report, we outline a taxonomy for the design of a safeguards crowdsourcing experiment with a detailed discussion on each element of the taxonomy. In forthcoming work, we will discuss the legal and ethical considerations for crowdsourcing safeguards and the types of crowdsourcing experimental designs that positively adhere to those considerations. Finally we will discuss candidates for active crowdsourcing experiments for Phase 2. This is the first installment of the new collaborative effort on crowdsourcing safeguards with Los Alamos National Laboratory and Sandia National Laboratories.

PART II  
A Legal and Ethical Examination of the Use of Crowdsourcing to  
Support International Nuclear Safeguards Verification

## 5. INTRODUCTION TO THE LEGAL AND ETHICAL EXAMINATION

The growth of the Internet and the Information Age has allowed the International Atomic Energy Agency (IAEA) Department of Safeguards to significantly expand its open source information collection, including new formats (such as multimedia [Barletta et al. (2016)]) and sources (such as social media [Lorenz, Feldman (2014) and Fowler et al. (2016)]) for potentially safeguards-relevant information. The Agency's exploration of social media data has led some researchers in the nuclear nonproliferation community to consider the potential for the use of societal mobilization, also known as crowdsourcing, to support information collection and analysis efforts. For the purposes of this research, societal mobilization "refers to an appeal that is broadcast to the public...requesting information, analysis, or opinion" [Gastelum (forthcoming 2017)].

In this paper, we discuss the legal and ethical implications of the spectrum of crowdsourcing activities that may be available to support international nuclear safeguards. We define legal implications as the legal rights and permissions to collect and use information (which includes raw data, analyses, or even technologies) from crowdsourced activities. We examine ethics within the framework of the three basic principles defined in the Belmont Report [The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978)] regarding research involving human subjects. While we recognize that crowdsourcing for international safeguards is distinct from biomedical or behavioral research using human subjects, we find that the protections for human life and well-being proscribed in the Belmont report are highly applicable to the potential collection or use of crowdsourced data to support international safeguards. We conduct this analysis with separate considerations for the *collection* and *use* of crowdsourced information (or technologies), and incorporate the taxonomy of the design of a safeguards crowdsourcing experiment presented in this project's midterm report (Part I) as framework for our examination. To better illustrate the connections between this analysis and the taxonomy presented in the midterm report, we **bold** the concepts defined in that paper.

During the course of this research, the team also became aware of a number of best practices and considerations for the IAEA's use of crowdsourcing activities to support international safeguards verification. While these were neither legal nor ethical per se, they are critical to any IAEA implementation of crowdsourcing for safeguards and thus included in the last section of this report.

## 6. LEGAL COLLECTION OF CROWDSOURCED DATA

There are two means to analyze the IAEA's legal rights and boundaries for collecting crowdsourced data. The first includes how the IAEA *collects* and *evaluates* all relevant information for its safeguards verification activities under Part I of the strengthened safeguards measures established by the 93+2 Commission. The second relates to how the crowdsourced data itself is collected from the various crowdsourcing platforms, for which legal terms regarding the collection and use of the data are defined in each platform's respective Terms of Use or Terms of Service.

### 6.1. 93+2 Part I Measures

The IAEA has been routinely collecting and analyzing open source information in support of international nuclear safeguards verification since at least 1991 with the establishment of an open source information analysis unit within the Department of Safeguards [IAEA (2007)]. The Agency's collection and analysis of open source information was formalized in the 93+2 Part I measures [International Atomic Energy Agency (1995)]:

*The proposed approach to a strengthened and more cost-effective safeguards system builds on the current system of material accountancy and control by integrating...[e]lements of increased access to information and its effective use by the Agency, including...improved analysis and evaluation of all relevant information available to the Agency.*

GOV/2784 describes three sources of that increased information as information provided by the state under an Expanded Declaration, information from strengthened safeguards measures, and "information from all sources available to the Agency, including the public media, scientific publications and existing Secretariat databases...as well as other information made available by Member States" [IAEA (1995) pp. 21].

Today, open sources of data collected and analyzed as part of the State Level Concept include, for example, press releases, news media, government and academic websites, trade information, scientific and technical journal publications, and satellite and ground-based imagery. In the past two decades, the visibility of open source information analysis in support of nuclear safeguards verification has grown considerably, providing context for traditional in-field safeguards activities and complementary access visits, raising important questions about state declarations or activities, and in some cases such as for states with Small Quantities Protocols, serving as the primary source on safeguards-relevant information.

GOV/2784 offers the following legal analysis regarding "improved analysis of information" including open source information [IAEA (1995) pp. 23]:

*"Comprehensive safeguards agreements require the Agency to draw conclusions from its verification activities (INFCIRC/153, para. 90), which presupposes the analysis and evaluation of the results of such activities. Improvements in the Agency's analytical techniques would therefore be consistent with the overall objective of a strengthened and more cost-effective safeguards system, and can be pursued within the Agency's existing legal authority."*

This appears to legally justify at least *some* use of some crowdsourced data. However, the means by which the IAEA gains access to that data may be interpreted differently. For this, must consider the **crowdsourcing type**: passive or active. Passive crowdsourcing is defined in the midterm report (Part I) as an activity in which “the stakeholder or administrator assumes a passive role and collects and analyzes content on a specific topic that has been freely generated by citizens in various sources.” This could be interpreted as constituting any open source information, including information on the Internet and social media platforms. Active crowdsourcing is defined as an activity in which “the stakeholder or administrator plays an active role by posing a particular problem [and] soliciting information or solutions.”

Assuming the consideration only of active crowdsourcing activities for this research,<sup>†</sup> we must consider both the **stakeholders** and the **administrators** of the crowdsource activity. A stakeholder is defined in our mid-term report (Part I) as “the party who wants the product of the experiment and formulated the purpose and goal.” In cases where the IAEA is the primary stakeholder, the Agency would initiate a crowdsourcing activity and have direct interest in the resulting data (this type of activity may be either directly administered by the IAEA, or by an external administrator on behalf of the IAEA). However, given the various means by which the Agency might gain access to crowdsourced data,<sup>‡</sup> the Agency may also be a secondary stakeholder, in which the data was collected for another purpose or stakeholder but also interest to the IAEA. Examples of instances in which the IAEA is a secondary stakeholder include, for example, the IAEA accessing open source crowd-sourced data from a platform that collected the information for non-safeguards purposes, or a non-governmental organization or third-party providing information collected via crowdsourcing from their own activities to the IAEA if it was later determined that the results would be of interest to the Agency.

In cases where the IAEA is a secondary stakeholder and accesses via open sources or is otherwise provided access to crowdsourced data, there appear to be no legal barriers for the Agency to treat that data any differently than it would other potentially safeguards-relevant information.

However, in cases where the IAEA is a primary stakeholder and is therefore responsible for initiating the collection of crowdsourced data, Member State buy-in would likely be required in any state for which data was being collected or analyzed in order to be considered “other information made available by the Member State” [International Atomic Energy Agency, (1995) pp. 21] If the IAEA were to be the primary stakeholder in a crowd-sourced activity that did not have buy-in from the host state, this could be considered espionage.

---

<sup>†</sup> Due to the Agency’s current use of “passive crowdsourced data”, and some debate as to whether such data is really crowdsourced, this project focuses on the use of actively crowdsourced data.

<sup>‡</sup> In the midterm report we define four potential crowdsourcing activity administrators: 1) The IAEA, 2), A Member-State that would provide data regarding its own activities to the IAEA; 3) a Third-Party, or 4) an other administrator such as an NGO or commercial organization.

## **6.2. Terms of Service from Individual Platforms**

Terms of Service (TOS) are the legal agreement between platforms and their users regarding, among other things, how the data on their platform can be collected or used. TOS differ by platform, country,<sup>§</sup> and by the crowdsourcing activities on that platform. If the IAEA were the primary stakeholder for a crowdsourcing activity, it would presumably administer the activity on its own site (in which it could define favorable terms of use that would enable the IAEA to collect and analyze data for safeguards purposes) or work with a platform (administrator) that had TOS that would be favorable to how the Agency wanted to collect the data. Most platforms include in their TOS that data harvesting should not be conducted in a way that results in a Denial of Service for the site. In addition to specifications regarding how the data from a specific platform is collected, some platforms terms of service explicitly prohibit the misrepresentation of an individual's identity for the purpose of data collection. This would prohibit the IAEA from collecting data under the guise of a non-IAEA identity.

---

<sup>§</sup> Most countries have country-specific legislation regarding the privacy of personal data that social media research must comply with. There may also be regional legal limitations, for example the European Union's Regulation 2016/679 and Directive 220/58/EC which govern the processing, storage, and deletion of personal data, and protection of privacy.

## **7. LEGAL USE OF CROWDSOURCED DATA**

As noted above, the IAEA's collection of crowdsourced data through and IAEA-administered or IAEA-stakeholder activity may face legal limitations via Terms of Service agreements and avoiding the appearance of espionage. However, once the data is available (for example, as provided by a third party and not collected directly by or for the IAEA), the legal use of that data to support safeguards verification originates from the same Board of Governors report which authorizes the collection and analysis of open source and other Member-State supplied information [International Atomic Energy Agency (1995)].

For cases in which the IAEA uses data collected from a commercial platform not administered by the Agency itself, care must be taken to follow the terms of use of the platform regarding use of the data. For instance, many platforms prohibit users from copying materials under an intellectual property rights clause. Additionally, specific countries have TOS agreements that prohibit users from copying data without authorization. However, since the IAEA would be unlikely to publish the data or use it for commercial gain, it would likely be able to use the data under a "fair use" clause without violating intellectual property regulations.

---

\*\* Fair use refers to a limitation on copyright laws in which research, scholarship, commentary, and other uses of copyrighted information is permitted.

## 8. ETHICAL COLLECTION OF CROWDSOURCED DATA

For the purposes of this research, we found the ethical principles on the use of human subjects for biomedical and behavioral research from the 1978 Belmont Report to be highly applicable as ethical considerations for crowdsourcing. The Belmont Report outlines three main ethical principles: 1) respect for persons; 2) beneficence; and 3) justice, explained in the context of crowdsourcing below.

### 8.1. Respect for persons

According to the Belmont report, respect for persons consists of two ethical principles: first, that individuals should be treated as autonomous agents (i.e., they can choose to participate or cease participating at any time), and second, that some individuals have diminished autonomy and must be protected. For crowdsourcing activities, this pertains to **crowd selection** and **disclosure choices**.

Crowd selection can consist of the general public, proximal crowds (those in close physical proximity to an area of interest) or expert crowds (those with specialized knowledge). For crowdsourcing for safeguards, respect for persons means that individuals need to be given the explicit choice to participate in the verification activity (which could be terminated at any time), and that some individuals who are not able to make that choice should not be included as they may not have the full mental capacity to determine their participation (for example, children, those with reduced mental capacity or illness, or incarcerated persons). Any crowd selection activity should make explicit that participation is completely voluntary. For selection among the general public or proximal crowds, criteria for participation could include minimum age restrictions in order to protect children. Protecting those in other compromised situations such as incarcerated persons or those with mental illness is much more challenging, but efforts may be made in the design of a crowdsourcing activity to neither target nor discriminate against these populations.

The disclosure of a crowdsourced activity's stakeholder and goal can be an important determinant for participants to decide their own level of participation. For example, if an individual agrees with the mission of IAEA safeguards, disclosure of the IAEA as a stakeholder may increase the person's intrinsic motivation to participate. Alternatively, an individual who disagrees with the IAEA's mission may choose not to participate. In cases when the IAEA is a secondary stakeholder, the stakeholder and goal change from the point of original collection to the secondary use, and may be resource prohibitive to retroactively disclose.

### 8.2. Beneficence

Under beneficence, people are treated ethically by respecting their decisions, protecting them from harm, and making efforts to secure their well-being. The two general rules under beneficence are: "(1) do not harm and (2) maximize possible benefits and minimize possible harms" [National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978)].

Abiding by the principles of beneficence for a crowdsourcing activity in support of safeguards would impact the activity's **objectives** (some objectives carry more risk for participants than

others) and **disclosures** (specifically identification of crowd participants, which could threaten participants' identity or personal information).

### **8.2.1. Do No Harm**

The principle of doing no harm for human research subjects in biomedical and behavioral research stems from a history of experimentation in the 20<sup>th</sup> century in which participants were exposed to infectious disease, experienced psychological impacts of high-stress experiments, or experienced other side effects from participating in experiments that were ethically unsound.

For a safeguards crowdsourcing activity's objectives, it is unlikely that participants would experience physical or psychological harm. However, depending on the activity, data collection or content generation may pose potential risk to participant safety. For safeguards verification purposes, we assume that collection refers to information or sample collection. Crowd-sourced data generation leverages the ubiquitous presence of sophisticated camera and monitoring equipment in the hands of citizens to provide data streams to the IAEA or societal mobilization platforms that could then be analyzed for safeguards-relevance. For safeguards, data collection or content generation may include taking photos, taking radiation measurements or environmental samples (if deemed acceptable by the BOG as part of wide area environmental monitoring), sending observations regarding facility status (presence/absence of a steam plume, presence/absence of equipment, or patterns of life such as fullness of the parking lot). In order to protect participants from safety hazards, the administrator should define the task so that it is not inherently dangerous, and so that participants are aware of potentially hazardous conditions. For example, if a crowdsourcing activity includes taking photos of a nuclear facility, participants should be advised to take photos only from publically accessible areas (i.e. not within a restricted area of the site).

### **8.2.2. Maximize Benefits and Minimize Harms**

To maximize benefits, only crowdsourcing activities for which the data can be reasonably expected to be used in support of safeguards verification activities should be collected (assuming IAEA as the primary **stakeholder**). That is to say, collecting data from crowdsource activities should not be collected for the sake of collecting data – there should be a defined data collection plan in which the collected data can be applied for a specific verification activity.

To minimize harm, the most significant consideration is the disclosure of crowd participants' identities to protect them from potential retaliation (from the host state, from anti-nuclear groups, etc.). The fear is that participants in a crowdsourcing activity without Member State consent<sup>††</sup> could be viewed as whistleblowers or even as spies. As such, protection of participants' personal information and identities is an important consideration. For an IAEA-administered crowdsource activity, this would be the direct responsibility of the IAEA (and presumably would be described in the TOS). For activities administered by non-IAEA entities, it would be important that the IAEA only use, store, protect, and transmit the data in accordance with the original

---

<sup>††</sup> This could happen a few ways. For example, a participant could take part in an activity that was not explicitly focused on safeguards (IAEA as a secondary stakeholder) that was later used to support verification efforts – if the Member State doesn't like the outcome, that could be bad. Or, if the Member State technically agrees to support the activity for the IAEA (as a primary stakeholder), but in practice views participants as spies.

administrator's or stakeholder's specifications per their activity so as not to compromise any of the participants' privacy.

### 8.3. Justice

In regards to the Belmont report and the ethical use of human subjects, justice refers to the distribution of the benefits of the research compared to who bears the burden. For crowdsourcing related to international safeguards, the benefit could be considered global (for those who consider nuclear proliferation to be a global good/benefit). However, the risks would likely be concentrated in specific populations depending on **crowd selection** and would vary depending on the **objective**. Understanding the potential burden of a crowdsourcing activity for safeguards also calls into question the **crowd motivation**.

If the crowdsourcing activity was global in nature (e.g., send us pictures of all the cooling towers in the world), the burden would be fairly distributed, as there are cooling towers in countries with nuclear energy, as well as those used for coal power plants. However, if the activity was focused on a specific country or site of interest, the **proximal crowd** would bear a much larger role (and therefore risk, burden). In crowd analysis activities, an **expert crowd** could theoretically be unduly burdened by excessive requests on a highly specialized topic. However, these impacts are expected to be minimal given the non-discrimination principles imposed by IAEA safeguards as well as the voluntary nature of crowdsourcing activities.

Regarding crowd motivation, the Mid-Term report describes **intrinsic motivation** (i.e. satisfaction derived from completing an activity) and **extrinsic motivation** (related to attaining a goal or external outcome such as a reward). Intrinsically motivated crowds who participate in crowdsource activities based on their own personal motivations pose little ethical challenge. However, there is significant debate in the crowdsourcing community regarding the use of monetary payment (an archetype of extrinsic motivation). Some crowdsource platforms, for example Amazon's Mechanical Turk, use small payments to motivate participants. While this was likely intended as a motivational tool and not as an employment opportunity, some have claimed that the use of small payments could exploit certain types of workers who rely on the payments as a source of income (these populations may also be in a state of diminished autonomy, for example undocumented immigrants).

The issue of payment may be exacerbated under the **technological solution** objective, in which the crowd develops equipment, computer codes, or other solutions that can require significant investments of infrastructure, time and resources on behalf of the participants. While many crowdsource activities with technological solution objectives do offer a reward for the best solution,<sup>##</sup> many participants receive no compensation for their time or supplies.

---

<sup>##</sup> For safeguards relevant examples, see the Underhanded C Contest ([http://www.underhanded-c.org/page\\_id\\_5.html](http://www.underhanded-c.org/page_id_5.html)) and the IAEA Robotics Contest (<https://challenge.iaea.org/robotics/about>)

## **9. ETHICAL USE OF CROWDSOURCED DATA**

Once crowdsourced data has been collected, its use poses minimal ethical concerns. However, as with ethical collection principles, participant identities and personal information should continue to be protected. In addition, it could be argued that the IAEA also has the responsibility to verify the data and to utilize the information in a non-discriminatory manner. However, because these responsibilities do not clearly fall within the definition of ethics, they will be discussed in more detail in the Best Practices section below.

For instances in which the IAEA is the secondary stakeholder, they might consider if the activity was conducted ethically before determining if they will use the data. Data generated or collected in an unethical way, even if not initially for safeguards use, might be better avoided.

## 10. OTHER BEST PRACTICES

In the process of analyzing legal and ethical issues for incorporating crowdsourcing activities into international nuclear safeguards verification, several concerns, issues, or recommendations for best practice emerged. While many of these were not explicitly legal or technical, the research team found them sufficiently pertinent to include here.

### 10.1. Independent Verification

Recent high-profile errors made in crowdsourcing efforts<sup>§§</sup> call out the potential fallibility of “the crowd.” Furthermore, the potential high visibility of a crowdsourcing activity in support of IAEA safeguards may have increased susceptibility to sabotage<sup>\*\*\*</sup> or misinformation. As with all open source data the Agency collects, they have the responsibility to independently verify that information. This is currently standard practice in the Agency’s analytical due diligence, and would need to extend to crowd-sourced data. Fortunately, many **data collection/generation** and **data analysis** crowdsourcing activities already include verification and validation within the activities themselves, requiring multiple users to submit similar responses before the data would be considered for use.

A **technological solution** developed for safeguards via a crowdsourcing activity would be subject to the same vulnerability analysis and technology approval process used for any safeguards equipment.

### 10.2. Protecting Sensitive Data

**Data collection and generation** about a state’s nuclear facilities has the potential to generate or expose commercial proprietary, security, and safeguards information. Furthermore, **data labeling or analysis** activities could unintentionally expose, via the mosaic effect, states’ sensitive information. As Oboler et al point out, “the potential damage of multiple individually benign pieces of information being combined to infer, or a big dataset being analyzed to reveal, sensitive information” is difficult to predict [Oboler et al. (2012)].

As with all safeguards activities conducted by the IAEA, the Agency has the responsibility to protect that data. Considerations regarding what information about a state would be exposed, collected, or analyzed in the course of a crowdsourcing activity to support safeguards verification would have to be carefully considered prior to the launch of an activity to ensure that the IAEA maintains its high standards for protecting sensitive information. This will involve careful cooperation between the IAEA and the state to ensure no sensitive data is being exposed, and that data collected is stored and analyzed according to the information security practices required for other safeguards data.

---

<sup>§§</sup> Such as the misidentification of potential suspects in the Boston Marathon Bombing (“New York Post Faces Suit Over Boston Bomb Article.” Christine Aughney, *The New York Times*, 06 June 2013) and of participants in white nationalist protests in Charlottesville (“Amateur Slieths Aim to Identify Charlottesville Marchers, but Sometimes Misfire.” Daniel Victor, *The New York Times*, 14 August 2017)

<sup>\*\*\*</sup> For a recent example, see: “DOS Attack Crashes Website Monitoring North Korea’s Nuclear Test Site.” Eric Niiler, *Wired*, 09 September 2016. <https://www.wired.com/2016/09/dos-attack-crashes-website-monitoring-north-koreas-nuclear-test-site/> Accessed 15 August 2017.

### **10.3. Avoiding Undue Burden**

Though we have established that crowdsourced data may be legally permissible for the IAEA to use in support of its safeguards verification activities, any collection, generation, or analysis of that data must be done so in a way that does not unduly burden states or nuclear facility operators. Indeed, the IAEA's report GOV/2784, which describes strengthened safeguards measures under the 93+2 activities, notes that while the IAEA can collect and analyze additional information to support verification, it should not burden states with "excessive costs or by cumbersome measures to facilitate verification" [International Atomic Energy Agency (1995) pp. 1-2]. Regarding the application of crowd sourcing, this could be interpreted to include activities administered by, or with the primary stakeholder as, the IAEA. Such undue burdens resulting from an IAEA crowdsourcing activity could include, for example, crowding around a nuclear facility perimeter disrupting physical protection operations or trespassing by overeager participants wanting to collect data.<sup>†††</sup>

### **10.4. Protection of IAEA Interests**

While it may seem self-evident, a crowdsourcing activity undertaken by the IAEA to support safeguards verification should protect the Agency's own interests. In order to prevent potential obfuscation, the IAEA does not disclose details regarding areas of potential proliferation concern directly to a state, other than to ask follow-up questions or perhaps request a visit via complementary access. Thus, care must be taken in the construction of a crowdsourcing activity to ensure that IAEA safeguards questions regarding potential undeclared activities are not disclosed.

### **10.5. Optics**

Finally, it is recognized that the IAEA safeguards mission operates in a politically sensitive environment. The Agency, if it chooses to adopt the use of crowdsourced data or analysis, should do so only in a manner that cannot be interpreted as intelligence collection or espionage. Any crowdsourcing activity conducted for the IAEA as the primary stakeholder should include Member State buy-in from all states which may be subject to information collection or analysis activities to alleviate this concern.

---

<sup>†††</sup> For a recent example of trespassing concerns related to crowdsourcing and online gaming, see: "Pokémon Go makers sued for encouraging trespassing." Cara McGoogan, *The Telegraph*. 03 August 2016. <http://www.telegraph.co.uk/technology/2016/08/03/pokemon-go-makers-sued-for-encouraging-trespassing/> Accessed 10 August 2017.

## **11. CONCLUSIONS AND FUTURE WORK**

Our analysis indicates that there are ways for the IAEA to utilize data from crowdsourcing activities to support safeguards verification. Some implementations of crowdsourcing for safeguards are legally or ethically uncertain, and must be carefully considered prior to adoption. In addition to compliance with legal and ethical norms for the use of crowdsourcing, there are other best practice considerations that would need to be accounted for in any IAEA-sponsored (i.e. as primary stakeholder) crowdsource activity.

While crowdsourcing could theoretically provide data useful for the analysis of a state's nuclear activities, there has not been sufficient testing to conclude that 1) sufficient quantities of data could be collected; and 2) quality and veracity of data would be sufficient for safeguards use. As such, experimental testing is required to further assess the use of crowdsourcing for safeguards. A conceptual experimental plan will be delivered with this project's year-end report, with experiments to be conducted in FY18.

PART III  
Experimental Designs to test Crowdsourcing in support of  
International Nuclear Safeguards

## **12. INTRODUCTION TO EXPERIMENTAL DESIGN**

In FY17, the Power of the People research team defined a taxonomy of crowdsourcing activities, and conducted a legal and ethical examination of how such activities could be conducted to support international nuclear safeguards verification as a foundation for a more in-depth examination of the potential to use crowdsourcing to support safeguards. The team plans to conduct a series of experiments in FY18 to more precisely evaluate the ability of the public or expert groups to provide relevant insight for safeguards-like problems. In this report, we describe conceptual experimental designs of activities to be conducted by Sandia National Laboratories and Los Alamos National Laboratory in FY18.

In our FY17 work, we identified several types of crowdsourcing activity, including data collection/generation, data analysis and labeling, and the development of technological solutions for safeguards. The team decided to focus on analysis and labeling tasks, due to implementation challenges expected from other crowdsourcing activity types. In addition, analysis and labeling tasks appeared more feasible for the support of international safeguards based on the team's ethical and legal analysis. In order to test a wide range of types of analysis, the team split into expert analysis activities (led by LANL) and non-expert communities (led by SNL). Each laboratory designed a series of experiments, described in more detail below.

### **13. EXPERIMENT 1 (SNL): “TAG THE TOWER” – CLASSIFICATION OF COOLING TOWER IMAGES**

In this experiment, we will test non-expert crowds using a basic information analysis experiment in which participants will classify digital images based on whether or not they contain a hyperbolic cooling tower. Classifying cooling towers serves as a simplified proxy for recognition and tagging of photos of nuclear facilities, which could be used to draw analyst attention to a site of interest, or to train an algorithm to recognize them.

This proxy problem will be used to determine if we can use non-experts with minimal training to classify photos of potential safeguards interest. If we are successful, more complex experiments such as assessments of multiple photos or determining the geolocation of a photograph may be further explored.

#### **13.1. Research Question**

This experiment seeks to test the question: Can crowdsourcing be reliably used to classify photographs of safeguards interest among a non-expert community?

To determine success, we will measure:

- Accuracy (compared to existing image labels);
- Timeliness (time to reach the desired number of labels per image on the complete image set);
- Participant completion (the number of tags a unique user completes)
- Participant diversity (number of unique participants); and
- Participant agreement (% agreement on tagging, for data validation).

#### **13.2. Methodology**

In this experiment, participants will be presented with a set of digital photographs, and asked to determine:

1. If there is a concrete, hyperbolic cooling tower such as those used at nuclear power stations and coal fired power plants; and
2. Whether the cooling tower (if present) has a steam plume.

Participants will be presented with a set of images, consisting of both cooling tower (steam plume and not) and not cooling tower images that have been collected from the Flickr site in accordance with the Flickr API user agreement. The images were manually labeled as part of a NA-22 funded research project in FY17.

#### **13.3. Platform**

We plan to conduct this experiment using the open source crowdsourcing platform Zooniverse. Because the Zooniverse platform selects which crowdsourcing activities to make public, we will be prepared to use an alternate platform if this experiment is not selected.

#### **13.4. Data Collection and Analysis Plan**

We plan to collect three to five labels (from unique participants) for each image in our dataset. We intend to collect the following information in the course of the experiment:

- Image labels

- Unique identifier of the participant who provided each label
- Timestamp associated with the label

Given those data points, we will be able to assess the five measures (accuracy, timeliness, participant completion, participant diversity, and participant agreement) described above.

## **13.5. Participants**

### **13.5.1. Recruitment**

Participants will be recruited via standard practices on the selected crowdsourcing platform. No targeted recruitment is anticipated. All potential participants will be provided with a short description of the task, and an estimate of the time required to complete the task. Due to the online, open, and voluntary nature of the platform, we will neither target nor discriminate against vulnerable populations.

### **13.5.2. Consent**

Users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the ease in discontinuing use, especially for activities which offer no or limited compensation, continued participation can presume continued consent, and users can stop at any time.

### **13.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with the labels they assigned to each image. This will be done through the collection of a user hash, user name, IP address, or other identifying information depending on the crowdsourcing platform. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **14. EXPERIMENT 2 (SNL): PERSONNEL PATTERNS OF LIFE**

The personnel present at a nuclear facility (just as any location) can provide a secondary indicator of the type, or scale, of activity taking place at that site. The presence of construction workers, office workers, military personnel, first responders, or other types of individuals – and their respective density at a facility – are an important aspect of “patterns of life” analysis.

In this experiment, we will test the ability of a non-expert crowd to identify:

1. whether an image contains people, and if so,
2. if an image includes people wearing firefighter uniforms, and
3. the count of people (uniformed, and total count) in an image.

The purpose of this activity is two-fold. First, it establishes a more complex level of analysis to test the effectiveness of crowdsourcing activities with potential safeguards implications. Second, this activity will share data and results with an on-going NA22 project that will use the image labels to train an algorithm to automate the counting/density estimates of people in various uniforms for patterns of life analysis. The NA22 project will provide both the data and the funding for participant compensation. The roll-out of the experiment will be a shared-cost between the two projects, with this NA-241 project having access to the full dataset collected by the activity in order to assess crowdsourcing effectiveness.

### **14.1. Research Question**

This experiment will test the question: Can the public reliably identify uniforms and count the number of people in a photograph, to assist in patterns of life assessments?

To determine success, we will measure:

- Participant agreement (% agreement, to approximate accuracy);
- Timeliness (time to reach the desired number of labels per image on the complete image set);
- Participant completion (the number of tags a unique user completes); and
- Participant diversity (number of unique participants)

### **14.2. 3.2 Methodology**

Users will be presented with a set of photographs collected from the open source photography sharing website Flickr, that have been selected based on a series of search terms developed under the NA22 project described above to target people and firefighters. For each photograph, users will be asked:

1. Does the photograph contain people? (if not, this will complete the task)
2. Does the photograph contain images of people in a firefighter uniform? If yes, how many?
3. How many people total are in the photograph?

### **14.3. 3.3 Platform**

The experiment will utilize Amazon’s Mechanical Turk platform.

### **14.4. 3.4 Data Collection and Analysis**

This experiment intends to provide approximately 15,000 images, with a goal of collecting three to five complete labels (i.e. answering the full question set above) for each. A user identifier such as an IP address or user name will be collected with each question set, along with a timestamp on the completion of the activity.

## **14.5. Participants**

### **14.5.1. Recruitment**

Participants will be recruited via standard practices on the selected crowdsourcing platform. All potential participants will be provided with a short description of the task, and an estimate of the time required to complete the task. Due to the online, open, and voluntary nature of the platform, we will neither target nor discriminate against vulnerable populations. The research team intends to provide minimal compensation (\$0.01 to \$0.10 per image) to encourage user completion.

### **14.5.2. Consent**

Users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the ease in discontinuing use, continued participation can presume continued consent, and users can stop at any time. Per the terms of Mechanical Turk, compensation will be provided on an image-by-image basis that a user completes. The minimal nature of compensation is not expected to inflict undue influence for users to continue participation if they are no longer willing.

### **14.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with the labels they assigned to each image. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **15. EXPERIMENT 3 (SNL): AUDIO AND VIDEO TRANSCRIPTION**

Time and resources permitting, Sandia would like to run a third experiment in which we evaluate a non-expert crowd's ability to transcribe audio and video files, from a variety of voices and accents. This will directly address a challenge currently being faced by the IAEA's State Factors Analysis section to collect and integrate multimedia information into their assessments. While efforts are currently underway to evaluate a suite of tools available for transcription, this activity will assess the ability of crowdsourcing to support that goal. While multi-lingual assessments would potentially provide the most value added for a safeguards use case, this assessment will start with English-language videos due to their availability to the research team.

### **15.1. Research Question**

This experiment will test the question: Can the public reliably transcribe English-language video and audio files across a number of voices and accents?

To determine success, we will measure:

- Accuracy (scoring mechanism is pending);
- Timeliness (time to reach the desired number of transcriptions per audio or video file);
- Participant completion (the number of transcriptions a unique user completes); and
- Participant diversity (number of unique participants).

### **15.2. Methodology**

Users will be presented with a segment of an audio or video file. Files will vary in length, to determine if/how segment length impacts accuracy. Users will be asked to transcribe the audio or video files into English text. The mechanism for scoring the transcription is pending.

### **15.3. Platform**

We will conduct our experiment on an open source crowdsourcing platform, such as Mechanical Turk, Zooniverse, or CrowdFlower.

### **15.4. Data Collection and Analysis**

This experiment will cultivate (collect, or if needed, create) a body of publicly available audio and video files in the English language, and collect a minimum of three transcriptions into English-language text. The experiment will collect a single user identification, along with the user's transcription and timestamp for completion of each audio or visual file, in order to evaluate the measures (accuracy, timeliness, completion, and diversity) described above.

### **15.5. Participants**

#### **15.5.1. Recruitment**

Participants will be recruited via standard practices on the selected crowdsourcing platform. All potential participants will be provided with a short description of the task, and an estimate of the time required to complete the task. Due to the online and open (anyone can participate) nature of the platform, we will neither target nor discriminate against vulnerable populations.

### **15.5.2. Consent**

Users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the ease in discontinuing use, continued participation can presume continued consent, and users can stop at any time.

### **15.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with the labels they assigned to each image. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **16. EXPERIMENT 4 (LANL): ELICITING TOPICS: CONCEPTS AND KEY WORDS**

### **16.1. Research Question**

The objective that drives the three related LANL experiments is to explore the efficacy of crowdsourcing methods to elicit and structure specialized expert knowledge quickly and efficiently that is very difficult to derive from data-driven methods alone and very expensive to solicit from a small cadre of experts. If successful, such resulting knowledge structures may be used for knowledge management systems, the basis for more refined ontology or Bayesian network construction, or become a key part of heterogeneous human-machine learning architectures.

The first research question we want to test is: Can crowdsourcing elicit a comprehensive set of concepts and key words related to a topic of interest?

### **16.2. Methodology**

#### **16.2.1. *Gamified Version***

Participants will be presented with a topic area of interest and asked to provide words associated with that topic within a specified time limit with the goal of providing the most words (that result in a score) in the allotted time period. As words are repeated, they will be added to a “Taboo List” in accordance with a predefined threshold and no longer help the player to accrue points. The player with the most points is the high scorer on a list. Players can only play once for each topic.

#### **16.2.2. *Regular Version***

Participants will be presented with a topic area of interest and asked to provide words associated with that topic within a specified time limit with the goal of providing the most words in the allotted time period.

We may elect to take submitted words and present them as topics in a follow-on iteration of the experiment.

Whether we execute the gamified version or the regular version will depend on the cost and mechanism for distribution for implementing the gamified version.

### **16.3. Platform**

There are a number of platforms we can take advantage of: For a predefined expert crowd, the gamification version can be executed in a mini-workshop setting ideally on a LANL-approved website on the green network. The regular version can be executed over email or through quick interviews. For a broader audience, online resources include: Crowdcrafting, Crowdfunder, and Zooniverse but vary in the ability to put controls on participation based on the expertise of the crowd.

#### **16.4. Data Collection and Analysis**

In addition to collecting concepts and key words associated with topics, we are also interested in the frequency of terms, and the sequence over which they accumulate. Here we want to identify the most strongly associated terms with a topic as defined by both frequency and time across users.

#### **16.5. Participants**

##### **16.5.1. Recruitment**

We plan to explore expert crowd recruitment as well as online resources. If we use the online platforms, participants will be recruited via standard practices on the selected crowdsourcing platform. All potential participants will be provided with a short description of the task, and an estimate of the time required to complete the task. We will neither target nor discriminate against vulnerable populations.

##### **16.5.2. Consent**

We will secure a written consent for any expert participation. If we use the online platforms, users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the ease in discontinuing use, continued participation can presume continued consent, and users can stop at any time.

##### **16.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with their input. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **17. EXPERIMENT 5 (LANL): ESTABLISHING RELATIONSHIPS BETWEEN CONCEPTS AND KEY WORDS FOR KNOWLEDGE STRUCTURING**

### **17.1. Research Question**

The objective that drives the three related LANL experiments is to explore the efficacy of crowdsourcing methods to elicit and structure specialized expert knowledge quickly and efficiently that is very difficult to derive from data-driven methods alone and very expensive to solicit from a small cadre of experts. If successful, such resulting knowledge structures may be used for knowledge management systems, the basis for more refined ontology or Bayesian network construction, or become a key part of heterogeneous human-machine learning architectures.

Once we have the set of concepts and keywords associated with a topic, the second research question we want to test: Can crowdsourcing help to structure the knowledge obtained in Experiment 4 and establish relationships in the predefined set of concepts and key words such as hierarchical relationships, causal relationships, etc.?

### **17.2. Methodology**

Here we will present participants with either a complete list of the words we obtained in Experiment 4 or pairs ordered by priority based on their frequency. We ask participants to mark relationships between words with lines or arrows and the option to label the relationship between the two words. If the participant is provided the full set of words, the resulting structure does not need to be connected but rather a number of structures can be end result. There will be a “parking lot” for words that cannot be related to others with an option of adding them at any time in the process.

### **17.3. Platform**

There are a number of platforms we can take advantage of: For a predefined expert crowd, the version can be executed in a mini-workshop setting ideally on a LANL-approved website on the green network. For a broader audience, the same online resources as for Experiment 4 could be used: Crowdcrafting, Crowdfunder, and Zooniverse.

### **17.4. Data Collection and Analysis**

For this experiment, we are looking to build the knowledge structure that leverages the input of all of the participants. We will track structural consensus by frequencies on arcs with the simple line as the primitive. Where there is conflict over structure, we will develop alternative structures to demonstrate alternative morphologies with frequencies attached to the arcs.

### **17.5. Participants**

#### **17.5.1. Recruitment**

We plan to explore expert crowd recruitment as well as online resources. If we use the online platforms, participants will be recruited via standard practices on the selected crowdsourcing platform. All potential participants will be provided with a short description of the task, and an

estimate of the time required to complete the task. We will neither target nor discriminate against vulnerable populations.

### **17.5.2. Consent**

We will secure a written consent for any expert participation. If we use the online platforms, users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the ease in discontinuing use, continued participation can presume continued consent, and users can stop at any time.

### **17.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with their input. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **18. EXPERIMENT 6 (LANL): VALIDATION OF THE KNOWLEDGE STRUCTURE**

### **18.1. Research Question**

The objective that drives the three related LANL experiments is to explore the efficacy of crowdsourcing methods to elicit and structure specialized expert knowledge quickly and efficiently that is very difficult to derive from data-driven methods alone and very expensive to solicit from a small cadre of experts. If successful, such resulting knowledge structures may be used for knowledge management systems, the basis for more refined ontology or Bayesian network construction, or become a key part of heterogeneous human-machine learning architectures.

If time and resources permit, we would like to pursue a third research question. Once we have a rudimentary crowd-created “folksonomy”, the last research question we want to test: To what degree can crowdsourcing be used to validate the knowledge structure?

### **18.2. Methodology**

We will present participants with the knowledge structures built and allow for voting on the consensus structure.

### **18.3. Platform**

There are a number of platforms we can take advantage of: For a predefined expert crowd, the version can be executed in a mini-workshop setting ideally on a LANL-approved website on the green network. For a broader audience, the same online resources as for Experiment 4 could be used: Crowdcrafting, Crowdfunder, and Zooniverse.

### **18.4. Data Collection and Analysis**

For this experiment, we will track the votes in favor of and against particular arcs to validate the structure. We will use these votes to come up with a candidate structure (and possibly alternates if votes are evenly split) and present to a knowledge structure expert for a final decision.

### **18.5. Participants**

#### **18.5.1. Recruitment**

We plan to explore expert crowd recruitment as well as online resources. If we use the online platforms, participants will be recruited via standard practices on the selected crowdsourcing platform. All potential participants will be provided with a short description of the task, and an estimate of the time required to complete the task. We will neither target nor discriminate against vulnerable populations.

#### **18.5.2. Consent**

We will secure a written consent for any expert participation. If we use the online platforms, users of online crowdsourcing platforms are generally recognized as agreeing to willfully participate in the activity (given there is sufficient explanation of the task on the site). Due to the

ease in discontinuing use, continued participation can presume continued consent, and users can stop at any time.

### **18.5.3. Privacy**

Participant information will be collected in order to identify the number of unique participants, and to associate a participant with their input. The research team will make efforts to limit the connection between a user identification method and association with actual identifying information of the user.

## **19. NEXT STEPS**

Upon review by NA-241, each laboratory will submit their proposed experiments to their respective Institutional Review Board for approval, and will commence deploying their experiments as appropriate. A final report of experimental design and outcomes will be provided at the end of the experimental period, no later than September 2018.

## REFERENCES

M. Barletta, M. Fowler, J. Khaled (2016) “Integrating Multimedia Information in IAEA Safeguards” *Proceedings of the American Nuclear Society Advances in Nuclear Nonproliferation Technology and Policy Conference*. Santa Fe, NM, USA.

N. Bilton, (2013) “Knowing Where to Focus the Wisdom of Crowds” *New York Times*, April 22, 2013

Createc “Createc wins IAEA Technology Challenge” (2016) <https://www.createc.co.uk/createc-wins-iaea-technology-challenge/> retrieved May 26, 2107.

C. Fishwick (2014) “Tomnod – the online search party looking for Malaysian Airlines flight MH370” *The Guardian* <https://www.theguardian.com/world/2014/mar/14/tomnod-online-search-malaysian-airlines-flight-mh370>

M. Fowler, J. Khaled, T. Sköld (2016) “Using Social Media Information in Analysis for IAEA Safeguards” Presented at the American Nuclear Society Social Media for Nuclear Nonproliferation Workshop. Santa Fe, NM, USA.

Z.N. Gastelum (Forthcoming 2017) “Societal Verification for Nuclear Nonproliferation and Arms Control” *Nuclear Non-proliferation and Arms Control Verification: Innovative Systems Concepts*. Eds., M. Dreicer, I. Niehmeyer, G. Stein. Springer.

N. Gerami (2013) “Attracting a crowd and what societal verification means for arms control” *Bulletin of the Atomic Scientists* v.63, no. 3, pp.14-18.

D.A. Grier (2017) “Crowdsourcing for Dummies-Cheat Sheet” <http://www.dummies.com/business/start-a-business/crowdsourcing-for-dummies-cheat-sheet/> retrieved May 26, 2017.

K.L. Hartigan, C. Hinderstein (2013) “The Opportunities and Limits of Societal Verification” *Proceedings of the Institute of Nuclear Materials Management Annual Meeting*, Palm Desert, CA.

C. Hinderstein et al. (2014) *Innovating Verification: New Tools & and New Actors to Reduce Nuclear Risks: Redefining Societal Verification* Nuclear Threat Initiative.

International Atomic Energy Agency (1995) “Strengthening the Effectiveness and Improving the Efficiency of the Safeguards System: A Report by the Director General” GOV/2784, pp. 5.

International Atomic Energy Agency (2007) “IAEA: Staying Ahead of the Game” <https://www.iaea.org/sites/default/files/safeguards0707.pdf> retrieved January 30, 2017.

B. Lee, M. Zolotova (2013) “New Media Solutions in Nonproliferation and Arms Control—Opportunities and Challenges” James Martin Center for Proliferation Studies, Technical Report.

M. I. Pinsel (1981) “The Wind and Current Chart Series Produced by Matthew Fontaine Maury” *Navigation: Journal of the Institute of Navigation* v. 28, no. 2, p.123-138.

J M Leimeister, M Huber, U. Bretschneider, H. Krcmar (2009) “Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition” *Journal of Management Information Systems*. v.26 no.1.

T. Lorenz and Y. Feldman (2014) “The Efficacy of Social Media as a Research Tool and Information Source for Safeguards Verification” *Proceedings of the Institute of Nuclear Materials Management Information Analysis Technologies, Techniques and Methods for Safeguards, Nonproliferation and Arms Control Verification Workshop*. Portland, OR, USA

E. Loukis, Y. Charalabidis (2015) “Active and Passive Crowdsourcing in Government” in *Policy and Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research*. M. Janssen, M.A. Wimmer, A. Deljoo eds. New York: Springer.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978) “The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research” [https://videocast.nih.gov/pdf/ohrp\\_belmont\\_report.pdf](https://videocast.nih.gov/pdf/ohrp_belmont_report.pdf) retrieved August 9, 2017.

A. Oboler, K. Welsch, L. Cruz (2012) “The Danger of Big Data: Social Media as Computational Social Science” *First Monday*, Vol 17, Number 7, 2 July 2012. <http://firstmonday.org/ojs/index.php/fm/article/view/3993/3269> retrieved March 23, 2017.

R. Richard, M. Edward, L. Deci (2000) “Self-Determination Theory and the facilitation of Intrinsic Motivation, Social development, and Well-Being”. *American Psychologist*, Vol.55, No. 1, 68-78.

J. Roberts, I-H Hann, S. Slaughter (2006) “Understanding the Motivations, Participation and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects”. *Marshall School of Business Working Paper No. IOM 01-06*; Management Science July 2006,

## DISTRIBUTION

- 1 National Nuclear Security Administration  
Attn: C. Stanuch  
U.S. Department of Energy  
1000 Independence Ave., S.W.  
Washington, DC 20585 (electronic copy)
- 1 Los Alamos National Laboratory  
Attn: Kari Sentz  
P.O. Box 1663  
Los Alamos, NM 87545 (electronic copy)
- |   |        |                   |                         |
|---|--------|-------------------|-------------------------|
| 1 | MS1371 | Tina Hernandez    | 06832 (electronic copy) |
| 1 | MS1371 | Zoe Gastelum      | 06832 (electronic copy) |
| 1 | MS0899 | Technical Library | 9536 (electronic copy)  |

