Open and Crowd-Sourced Data for Treaty Verification

Contact: Dan McMorrow — dmcmorrow@mitre.org

October 2014

JSR-14-Task-015

JASON
The MITRE Corporation
7515 Colshire Drive
McLean, Virginia 22102-7508
(703) 983-6997

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
October 31, 2014	Technical	

Open and Crowd-Sourced Data for Treaty Verification

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
1314JA01

5e. TASK NUMBER
PS

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
JASON Program Office
7515 Colshire Drive
McLean, Virginia 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

JSR-14-Task -015

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Biological Weapons
Bureau of Verification, Compliance and Implementation Washington, DC
ACV/VTT, Suite 5871
2201 C. Street, NW
Washington, DC

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Department of State and DTRA jointly sponsored the present study, requesting that it "explore the validation of open source data so as to withstand the rigor of verifying treating compliance for U.S. policy makers and for sharing with the international community. It will examine tools for automated validation of open source information and assess the potential utility of open source data for treaty verification, transparency, and confidence building. The open source information under consideration includes traditional reporting of public discourse as well as newer forms of open source information, such as public domain social media, data from public domain sensor systems, and data introduced into the public domain through mobile devices."

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	b. ABSTRACT	c. THIS PAGE			Brian Nordmann
U	U	U	U		**19b. TELEPHONE NUMBER** *(include area code)* 202-647-2408

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Contents

EXECUTIVE SUMMARY

In response to a request from the Department of State and the Defense Threat Reduction Agency, JASON completed a study that provides evidence of open and crowd-sourced data being potentially useful and effective for treaty verification.

There has been a sea change in society due to the development of widely distributed sensors, and of the Internet: never before has so much information and analysis been so widely and openly available. The opportunities for addressing future monitoring challenges include the ability to track activity, materials and components in far more detail than previously, both for purposes of treaty verification and to counter WMD proliferation.

We organized our findings in response to the seven areas of potential research and development identified as study objectives by the sponsor. Our findings and recommendations are summarized as follows:

Findings

1. **Data sources**: The proliferation of inexpensive, networked sensors and the rise of social media provide significant new sources of information enabling pervasive monitoring and increasing societal transparency. The large number of potential observations can in many instances make up for relatively crude measurements made by the public. Users are motivated to communicate environmental information because they themselves benefit from early warning and hazard prevention.

2. **Data integrity**: Deliberate deception is a difficult challenge. Still, though potentially vulnerable to falsification, the newly available information can be validated through independent sources, and background clutter can be checked for consistency. Data corruption during transfer and storage can be addressed through known techniques.

3. **Signal clutter**: Extracting reliable knowledge from open source data is a discrimination challenge. However, suitably characterized clutter in technical data can be used for calibration and at least partial validation.

4. **Standards**: Protocols are de-facto standards that evolve rapidly, as driven by public and commercial interests. The government needs to be responsive to keep up, with flexibility and agility being key to exploiting data in an effective manner.

5. **Cooperative verification and monitoring**: Open and crowd-sourced data can enhance transparency and confidence through cooperative activities with other nations, and through exchanges among technical experts in other areas of mutual interest, such as the environment, climate, and public health.

6. **Public sharing of information and analysis**: Open sharing, as appropriate, improves the reliability and utility of both data and analysis. There are technical means of reliably transmitting data and preserving it from corruption or accidental deletion.

7. **Actionability**: Open source data and analysis have value even if not fully validated, by contributing to cueing, in conjunction with other data. Determining risks of action versus non-action is a political decision once the data have been analyzed, and there is no guarantee against false alarms.

Recommendations

1. Monitor and promote the development of ubiquitous, networked sensors and associated analysis.

2. Monitor trends in social media and open-source observations: i) judiciously distinguish those domains of specialized interest or capability in government from the much larger efforts in public and commercial sectors; ii) promote "grand challenge" projects to implement end-to-end

solutions, from conception and engineering to final analysis, in order to raise technical readiness levels (e.g., environmental gas monitoring with multi-agency partnering); and iii) establish web portals for assimilation and distribution of relevant open-source information.

3. Characterize the clutter in measurements of interest for purposes of discrimination, calibration, and validation. Recognize and support non-governmental efforts to calibrate and validate open-source information and analysis.

4. Charge a working group, including private-sector and other external members, with the responsibility of addressing cultural, ethical, legal and social implications of open-source exploitation.

5. Promote transparency and validation by: i) keeping open-source information and analysis open to the maximum degree possible and appropriate; ii) recognizing successful approaches and technologies (e.g., through awards); and iii) rewarding transparency in businesses worldwide.

6. State Department staff in foreign missions should be trained and tasked to identify open-source opportunities for application at home and abroad.

1 INTRODUCTION

This is a revolutionary time in the ability of private citizens as well as governments to both acquire and communicate massive amounts of information, including highly technical data. It is reminiscent of the late 1950's and early 1960's, when new technologies epitomized by the U2, the SR71 and CORONA, the first photo reconnaissance satellite, made it possible to penetrate the Soviet Union's Iron Curtain and dispel the myth of a missile gap.

The exponential increase in data volume and connectivity, and the relentless evolution toward inexpensive – therefore ubiquitous – sensors provide a rapidly changing landscape for monitoring and verifying international treaties. Of comparable importance is the potential for widely distributed miniature sensors to track activities of interest, whether or not in violation of formal treaties. In short, transparency is being significantly enhanced around the world due to the revolution in quantity, quality, ubiquity and availability of open source information.

New technologies can monitor potentially threatening activities, objects or people. And, because these technologies have many other beneficial applications – from enhancing home security to providing novel marketing tools for commerce – they are widely available and inexpensive. These open sources of information can offer persistent access, supplementing the capabilities of National Technical Means (NTM) to verify compliance with formal arms control treaties. They can also complement shared technical means, such as the International Monitoring System (IMS) to monitor nuclear explosive testing, and cooperative aerial monitoring under the Open Skies Treaty.

As technology advances, it presents challenges for government to develop an effective use of open-source information and analysis, in maintaining integrity of the data coming from numerous public sources (including countering deliberate efforts to introduce falsified information) while also addressing the balance between effectiveness and privacy. Moreover, government recog-

nizes that the total control of information possible in the past is no longer realistic. There is precedent for addressing some of these organizational and operational issues based on experience gained from such academic programs as Galaxy Zoo and the Quake Catcher Network.

Many new sensors are multi-use: they provide new means of monitoring treaty-related activities, while also collecting other data important to society, from air quality, soil moisture and localized sources of CO_2 emission to identifying and sending warnings of major catastrophes, including real-time tracking of health crises. Their multi-use character gains these sensors broad public acceptance.

Another aspect of the Information Technology (IT) revolution is the ability to process large numbers of data. Handling "Big Data" efficiently and effectively will be crucial to successful open-source analysis, and requires effective means of collecting and organizing the data.

Currently, the U.S. is addressing several critical arms-control problems. These include the verification of strict and comprehensive compliance requirements on Iran's nuclear program, as specified in the temporary agreement of the Six-Party talks, while provisions for a long-term treaty are being worked out. Detailed verification is also needed to follow up on the successful international campaign that removed and destroyed Syria's chemical weapons. These are two examples from which the U.S. can learn what it should be emphasizing in its R&D programs to develop effective systems for future treaty monitoring.

Another recent event addresses the use of social media, as reported by the *Wall Street Journal*'s Julian Barnes (August 6, 2014): "Minutes after Malaysia Airlines Flight 17 went down on July 17 in eastern Ukraine, killing all 298 on board, a Defense Intelligence Agency analyst sifting social media communications got 'a hit.' The Russian-speaking analyst saw a posting from pro-Russia separatists in Ukraine, on Russia's VK social media site, claiming to have shot down a Ukrainian military cargo plane." DIA chief

Lt. Gen. Michael Flynn is quoted in the article as having said "The first indication of who shot it, what shot it and when and where it was shot was all social media... It was literally within minutes."

1.1 Charge to the Panel

The Department of State and DTRA jointly sponsored the present study, requesting that it "explore the validation of open source data so as to withstand the rigor of verifying treating compliance for U.S. policy makers and for sharing with the international community. It will examine tools for automated validation of open source information and assess the potential utility of open source data for treaty verification, transparency, and confidence building. The open source information under consideration includes traditional reporting of public discourse as well as newer forms of open source information, such as public domain social media, data from public domain sensor systems, and data introduced into the public domain through mobile devices."

The statement of work lists the following seven study objectives:

- *Data sources – what open and crowd-sources exist and lend themselves to exploitation for treaty- verification-relevant analysis? How might these sources be exploited? What are the practical and technical limitations to use of these sources?*
- *Data Integrity – How do we detect spoofing and deception? The JASON study will examine possible methodologies to find and vet relevant open source data and metadata, and lay out the architecture for implementing an optimal mix of these methodologies in a future automated open source validation tool.*
- *Signal Clutter – In heterogeneous and distributed sensor networks, how does one find meaning in all of the noise. The JASON study will assess*

possible approaches for using existing commercial or government tools or creating more customized tools to strain and analyze information present on public-domain open source data.

- *Standards – What kind of uniform standards should one use for both data and metadata? The issue of data standards is integral to the design of the open source validation tool and aid to uniform and rapid signal processing.*

- *Cooperative verification and monitoring – What are the limitations on cooperative observation, validation and analysis? Are there models for cooperative sharing of verification and monitoring activities in this regime?*

- *Public sharing of information and analysis – to what extent can collection and analysis be based on publically-available technology and techniques? What are the limitations on public use of the results of this collection and analysis?*

- *Actionability – Once the integrity and utility of the data is established, how does one determine its actionability? Transformation of data into information is only a first step. The next step is establishing a process for determining the feasibility of the various approaches.*

2 PUBLIC DOMAIN/OPEN SOURCE DATA

This section explores the categories and potential verification utility of 1) data derived from sensors and transducers that generate non-imaging digital information from physical measurements, and 2) public-domain images at both optical and radar wavelengths. We identify the following classes of sensor-derived open source information that can potentially be useful for treaty verification purposes:

Non-Imaging

- Seismic
- Radioactivity
- Trace Gases
- Audio Information

Imaging

- Remote Sensing, optical and infrared
- Remote Sensing Synthetic Aperture Radar
- Still images
- Video

It is also helpful to consider the different kinds and configurations of sensors that might be generating these data, namely

1. Embedded smartphone sensors
2. Sensors attached as peripheral devices to smartphones or computers

3. Stand-alone sensors, with appropriate power and communication capabilities.

In many – perhaps most – instances public interest in sensor networks will be driven by considerations other than treaty verification. We attempt below to identify these potential drivers of public domain sensor deployment that might also provide open data of value to treaty verification.

A potential application of these sensing technologies is for site-specific monitoring. The traditional approach of using laboratory-grade instruments can be augmented by appropriate distributed sensors with the flavor of the devices described below.

In general, the concept of operation is one in which individuals or organizations (e.g., businesses, educational institutions, interest groups) widely share data from sensors that they deploy and manage, in return for a service that can range from warning of hazardous conditions (poor air quality, impending tsunami or shaking from a nearby earthquake, etc.) to improvement of standards of living (e.g., enhanced energy efficiency, improved environmental conditions or mitigating climate change). Internet-based communication can make a huge number of sensors available for monitoring the environment (in the broadest sense) and, in return, for providing useful services to those who contribute to this monitoring infrastructure. We come back to this issue with a more general discussion of incentives for open-source treaty verification in the next section.

It is important to recognize that there is still much to be learned from experiments and by experience in order to develop confidence in the applications of these public domain sensing technologies. This is an area in which rapid progress is being made, and one can expect that much will be learned on the full potential as well as practical limitations of these technologies in the near future. Clarifying and understanding the limitations will be important for the policy planners as well as the technical enthusiasts.

2.1 Seismic Sensing

Seismic information has a longstanding role in the verification of treaties that prohibit or limit underground nuclear-explosion testing. The International Monitoring System (IMS) that has been established for the Comprehensive Test Ban Treaty has 170 seismic monitoring stations across the globe (`www.ctbto.org`). This network is augmented by the international scientific network of seismometers, as well as by numerous national and local seismic stations. These seismometers have exquisite sensitivity, with performance that is limited by the seismic background noise of the Earth as opposed to sensor noise.

Seismologists are exploring a variety of crowd-sourced approaches to seismic monitoring. The use of a dense network of less capable sensors is seen as a valuable supplement to the existing but sparser network of high-performance seismometers. An example of a crowd-sourced seismic network that is under development is the Community Seismic Network (CSN) [1], which provides free accelerometers to residents of Pasadena in return for their connecting the seismic sensor to the network.

2.1.1 MEMS accelerometers for seismic sensing

MEMS accelerometers are used in smartphones. For example the iPhone 5 uses a LIS331DLH sensor. Somewhat higher performance MEMS devices are the Phidget sensors ($140, `http://www.phidgets.com`) that are used by the Community Seismic Network (`http://csn.caltech.edu`). MEMS devices that have been specifically engineered by HP for seismic applications and that achieve substantially lower noise are described by Milligan *et al.* [2]. These MEMS devices directly measure accelerations, and their performance parameters are summarized in Table 2.1.

The low-cost MEMS accelerometers exhibit many orders of magnitude higher noise than the scientific-grade seismic sensors, but we suggest that the

Table 2.1: MEMS Accelerometer Performance Comparison. The columns list the device type, acceleration noise spectral density, cost, and N, the number of sensors needed to achieve a network sensitivity adequate to detect a 1 kiloton nuclear explosion from a distance of 370 km, with 5 σ significance in a bandwidth of 10 Hz. (*The HP device is not currently commercially available.)

Device	Acceleration Noise Spectral Density (m $s^{-2}/\sqrt{Hz}$)	approximate cost	N
iPhone 5 accelerometer	2×10^{-3}	Free with your phone	400
Phidget 1044_0	10^{-4}	$140	10
Silicon Designs 1221-002	5×10^{-5}	few hundred $	1
HP Seismic-optimized MEMS sensor	10^{-7}	*	1
CTBT/IMS/scientific	Background-limited	$\sim \$10^4$	1

combination of higher density (where we may gain in the ratio of signal to sensor noise as $\sqrt{N}$ where N is the number of sensors in the network) and close range (where the peak ground acceleration signal scales with standoff distance R as roughly $1/R$) can provide adequate aggregate sensitivity to measure seismic signatures of interest.

2.1.2 Seismic signal levels from underground nuclear explosions

The relationship between nuclear yield and propagating seismic energy depends on numerous factors, including whether the detonation is fully coupled or not, and the local details of Earth's structure. For the purposes of this report we adopt the approximate relationship between seismic magnitude m_b and nuclear yield Y used by Kim and Richards [3] for a fully coupled nuclear detonation in hard rock, namely $m_b = 4.45 + 0.75 \log_{10}(Y/1 \text{ kt})$.

A 1 kiloton explosion corresponds to $m_b \sim 4.45$. The power spectrum of the surface accelerations depends on the nature of the event, and the low-pass filter imposed by propagation through the Earth. We are primarily interested in distances of a few hundred km, at which the power spectrum is expected to peak at frequencies of a few Hz.

To convert from m_b magnitudes into peak ground accelerations we used the empirical data from the North Korean 2006 nuclear test, as reported in [3], with a peak vertical ground velocity of 9.7 microns/second recorded at the MJD seismic station a distance of 371 km from the detonation. We downloaded the seismic data record of measured ground velocities vs. time (from `www.iris.edu`), and took a numerical derivative to determine the accelerations corresponding to these recorded ground motions.

This $m_b = 4.3$ nuclear event, corresponding to a fully coupled yield of 0.6 kt, produced typical peak ground accelerations of 2 $\times 10^{-4}$ m s^{-2}. A 1 kiloton event would be expected to produce an acceleration that is $10^{(4.45-4.3)}$=1.4 times larger, or 2.8 $\times 10^{-4}$ m s^{-2}. Given local propagation uncertainties and our goal of making simple estimates, we'll round this up to an acceleration signature of 3 $\times 10^{-4}$ m s^{-2} for a 1 kiloton explosion at a standoff distance of 370 km, with a scaling with yield Y and distance R of $\text{peak accel}(\text{Y}, \text{R}) \sim 3 \times 10^{-4}\text{m s}^{-2}(370\text{km/R})(10^{0.75\log_{10}(\text{Y(kt)})})$.

Assuming we require a 5 σ detection against the noise floor of a distributed sensor network with N seismic sensor nodes and that we can perform coherent waveform analysis, for a 10 Hz bandwidth we require that the system satisfy 3×10^{-4} m s^{-2} $> (5/\sqrt{3})\ \sqrt{10\,\text{Hz}}$ SANSD $N^{-0.5}$, where SANSD is the sensor acceleration noise spectral density in m s^{-2} per $\sqrt{\text{Hz}}$ listed in Table 2.1. The $\sqrt{3}$ factor in the denominator comes from presuming a coherent signal in three independent sensor axes, per node. This allows us to compute the requisite number of various MEMS accelerometer types needed in the network in order to detect a 1 kiloton explosion from a standoff distance of 370 km, and leads to the entries in the final column of Table 2.1.

A single iPhone should have adequate sensitivity to detect a 30 kt nuclear explosion from a standoff distance of 240 km, at 5 σ; for fully coupled explosions of 0.5 kt and 0.1 kt, similar (5 σ) detection is possible at distances of 200 km and 70 km, respectively.

The main point of Table 2.1, however, is that 400 of these relatively crude sensors match the capabilities of a high-end accelerometer having sensitivity near or at Earth's background noise level. As there are more than 1.5 billion smart phones worldwide, this amounts to a capability analogous to that of nearly 4 million high-quality sensors: to be compared with the few thousand research-grade seismometers deployed around the world (including the 170 IMS stations). To be sure, smart phones are not uniformly distributed across all countries, and they suffer from additional causes of signal degradation (e.g., the accelerometers are not tightly coupled to the ground), nor has it been demonstrated that their noise adds incoherently, and smoothly across the relevant spectral band, nor that additional sources of noise, such as motion of the sensor platforms are insignificant. Nevertheless, this example illustrates how a ubiquitous deployment of sensors can supplement the capabilities of research-grade arrays, because large numbers may make up for reduced performance.

Considerable research has been done to make accelerometers embedded in computers and cell phones useful for seismology. First, there are algorithms that filter out data collected from sensors showing ongoing motion (e.g., a telephone carried by someone walking). Second, signals from otherwise-still sensors are compared with data streams from nearby sensors to ensure that multiple sensors have picked up the same signalsthis prevents, for example, the signal from a phone knocked off a table from being considered valid.

There are other filters applied to validate the data, in addition to phones having valid (confirmed) location tracking and showing that they are otherwise stationary. One might think that few accelerometers would then be available for transmitting valid data, but in fact most phones and computers (including laptops) are at rest for long periods of time-for example, while the owner is at sleep. Given the large numbers of these accelerometers currently deployed, let alone the enormous increase expected over the coming years, there is a vast array of sensors that–even after filtering for false positives, bad calibration, etc.–can provide significant information. However, as al-

ready emphasized at the end of Section 2 on page 9, we are still in the early stages of exploring both the potential as well as practical limitations of data gained from public domain/open source sensors.

2.1.3 Prospects for exploiting expanded public seismic networks for treaty monitoring

There is substantial public interest in monitoring seismic activity in earthquake-prone regions, so we anticipate increasing the density of public-domain seismic sensor coverage in regions where high population density coincides with seismic hazards. More specifically, there is a benefit to individuals and organizations (businesses, government agencies and other institutions) to participate in early-warning systems, through which individuals contribute their computer and smart-phone accelerometer readings into a centralized system and in return obtain warning of impending hazards due to ground shaking or tsunamis [4, 5]. Communication is automatic, over the Internet, and the system is activated when a sufficient number of sensors are triggered in an appropriate manner (e.g., correcting for the possibility that a computer has been jostled, or a phone has been picked up). The data are processed automatically upon receipt, such that there can be seconds to minutes of warning time in regions tens to hundreds of kilometers away from the source. This is enough time to significantly reduce damage and casualties, whether by shutting down large equipment (e.g., trains, subways, power plants) or by individuals taking shelter or moving to high ground in regions subject to tsunamis.

Figure 2-1 suggests to us that a public seismic network in East Asia is likely; indeed, Japan is a leader in early-warning systems for seismic shaking and tsunamis. Public domain seismic sensors in South Korea would be located 300–400 km from the existing nuclear test site in North Korea, and a seismic network in Japan would be 700–800 km away from the NK test site. Most of the interior of earthquake-prone Iran is within 200 km of the three major cities of Tehran, Shiraz and Isfahan. A public seismic network

in that country would provide substantial coverage of potential underground test sites as a side benefit.

Even a few dozen installations per country with the sensors used by the Community Seismic Network (CSN) in Southern California would provide sensitivity to kiloton-scale underground explosions over distances of hundreds of kilometers, provided we can perform coherent waveform analysis on the resulting data archive. An additional verification benefit will come from calibration of local seismic energy propagation with a finer spatial sampling than typically available from scientific and IMS seismic monitoring stations.

Figure 2-1: Map of Seismic Hazards, from `http://www.esa.int/spaceinimages/Images/2004/07/Seismic_hazard_map`. We anticipate an expansion of public domain seismic sensors in populated areas with substantial earthquake hazard. This can also be considered a resource for open source monitoring for underground nuclear explosions.

2.2 Radioactivity

The Fukushima earthquake and tsunami prompted a number of grass-roots open source radiation monitoring programs, which included sensor development, data archiving and analysis, and associated social media fora. We

anticipate a segment of the public having a continuing interest in using sensors to monitor the radioactive character of their surroundings.

This interest can arise from natural or man made sources of radiation. Figures 2-2 and 2-3 show the worldwide distribution of radon, and proposed nuclear power reactors. Both of these distributions will likely drive the installation of public domain radiation monitoring systems. If we imagine a correlation between nuclear power installations and potential nuclear proliferation threats, we can benefit from public interest in monitoring local radioactive background levels.

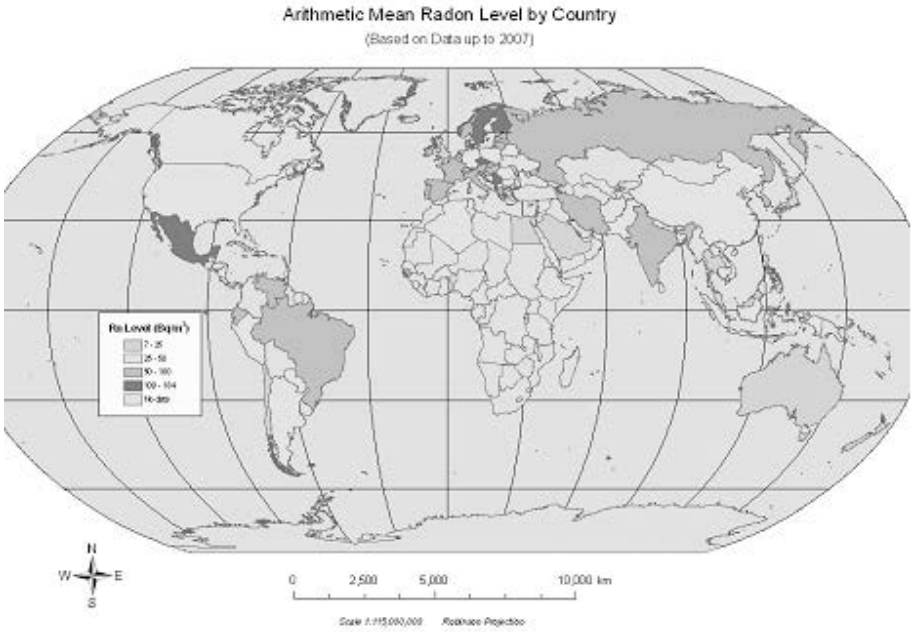

Figure 2-2: Map of Worldwide Radon, from `http://www.mclaughlincentre.ca/research/map_radon/Index.htm`. We anticipate an expansion of public domain radiation sensors in areas with public concern about radon concentration, or where the public has concerns about nuclear power accidents.

One of the notable developments since the March, 2011 Fukushima event is that the detectors can be highly networked, uploading the results onto the Internet for combining and mapping with others' measurements. A conflu-

Figure 2-3: Map of Worldwide Proposed Reactors, from `http://www.climatecentral.org/blogs/nuclear-maps/`.

ence of public interest and associated social structures (e.g., DIY and Makers movements; social networking), along with wide availability of smart phones and of cell-phone infrastructure has had at least as profound an impact on radiation detection as have improvements in the sensors themselves.

Broadly, four types of technology are available for nuclear radiation: CMOS, PIN diode and scintillator detectors are sensitive to γ radiation, whereas Geiger-Müller (G-M) detectors are sensitive to both β and γ radiation, and also to α radiation if a thin enough window is used (Table 2.2 gives an illustrative summary). Roughly speaking, sensitivity is proportional to price, with cell-phone cameras providing effective detection at essentially no extra cost, whereas more expensive technologies have either more sensitivity, or responsiveness to more types of radiation, or both (Figure 2-4).

To provide context, γ radiation has a range of hundreds of meters in air (i.e., is essentially absorbed within a kilometer or two), so can in principle be measured at considerable distance from a source. Under many circumstances, these measurements can sensibly be made from a moving vehicle, such as a car or helicopter. In contrast, α and β radiation are fully absorbed within

Table 2.2: Radiation Detectors

Technology	**Sensitivity**	**Price**	**Comments**
CMOS cell phone camera (γ)		Free App	uncalibrated (gives event hits)
CMOS cell phone camera (γ)	0.03–1.3 cpm/μSv/h	\$4.99	10-20 min initialization, 3-5 min for low level; mapping/communication function to come
PIN diode (γ)	10–16 cpm/μSv/h	\$72–99	mapping/communication function included
Geiger-Müller (β, γ)	130 cpm/μSv/h	\$160–300	Limited (beta) mapping/communication function included
Geiger-Müller (α, β, γ)	110–360 cpm/μSv/h	\$450–1000	mapping/communication function included; safecast.org DIY kits available
Geiger-Müller (α, β, γ)	110–360 cpm/μSv/h	\$470–595	data logging via radcast.org 0.01-2000 μSv/h range
Geiger-Müller (α, β, γ)	110–360 cpm/μSv/h	\$699	data logging via safecast.org
CsI (Tl) scintillator (γ)	$= 10^3$ cpm/μSv/h	\$1280–1860	PA-1100 (Japan) has mapping/communication function; link to Japan Electric Measuring Instruments Manufacturers' Association (JEMIMA)

Sources: Ishigaki, *et al.* [6]; Cogliati, *et al.* [7]; DeBarber and Yamamoto [8].

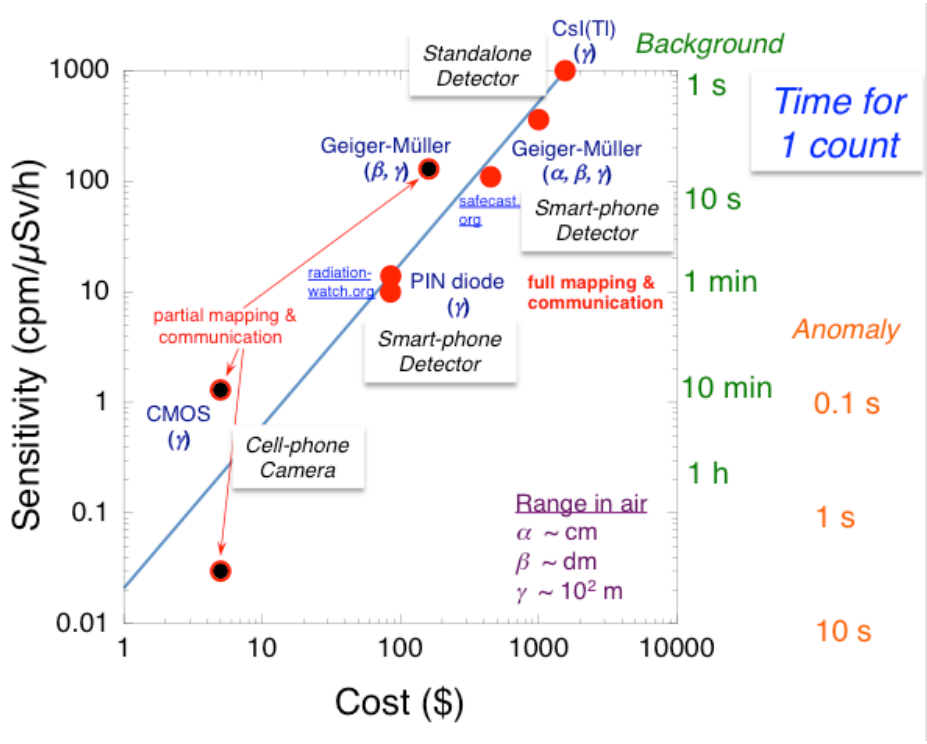

Figure 2-4: Sensitivity (counts per minute per μSv/h) as a function of cost (2014 USD) for a selection of technologies available for detecting nuclear radiation: red symbols indicate that software is available for mapping and communicating results on the Internet, whereas black symbols indicate that this capability is not (or only partly) in place. For a given sensitivity, the corresponding time to record 1 count from average background radiation (0.07 μSv/h) and from the lowest-level anomaly on the International Nuclear Event Scale (Level 1) is shown in green and orange, respectively, on the right (see Figure 2-5). Measuring background is often required for calibrating or validating the detector, and is useful for mapping the ambient field as a function of location and time ("clutter").

many centimeters to a few meters, respectively, so can only be measured in close proximity to a source. Still, such measurements may be useful in the case of radioactive dust dispersed in the environment (even if only α and β-emitting) because, in this case, the detector can be close to the source (dust or gas distributed throughout the air).

The ranges of readings to be expected from natural background as well as a variety of anomalous events. Figure 2-5 shows that even the lowest-cost sensors – cell phone cameras – can provide rapid indications of anomalies in radiation dose (e.g., Level 2 or even 1). The importance of more sensitive detectors is in better characterizing the background, as well as in being responsive to γ -radiation out to greater distances from a source.

A well-run nuclear facility – whether a nuclear power plant, or a site for enrichment, reprocessing, storage or waste isolation of nuclear fuel or components – is expected to maintain dose levels at or well below the natural background. Therefore, no radioactivity anomaly is in general expected from such facilities. However, should there be an accidental release, then history shows that there is a significant chance that modern detection systems may record the event. Mayak (Kyshtym), Chernobyl and Fukushima are well-known examples of nuclear-accident sites, but even Hanford experienced some releases above background [9].

Therefore, careful characterization of background radiation as a function of location and time can provide a baseline for identifying nuclear activities. This is an example of how useful it can be to characterize background clutter, in that detection of change even from a complex background is typically much easier than an absolute detection of an anomaly. Of course, there is still opportunity for false alarms, for example due to transport of medical isotopes or the nearby passage of an individual who has recently had certain nuclear medicine procedures. The importance of ubiquitous sensing – a large number of sensors widely distributed across a geographic area – networked to the Internet is that this background clutter can be richly documented by the public at large, and therefore in a manner that is likely to be far more

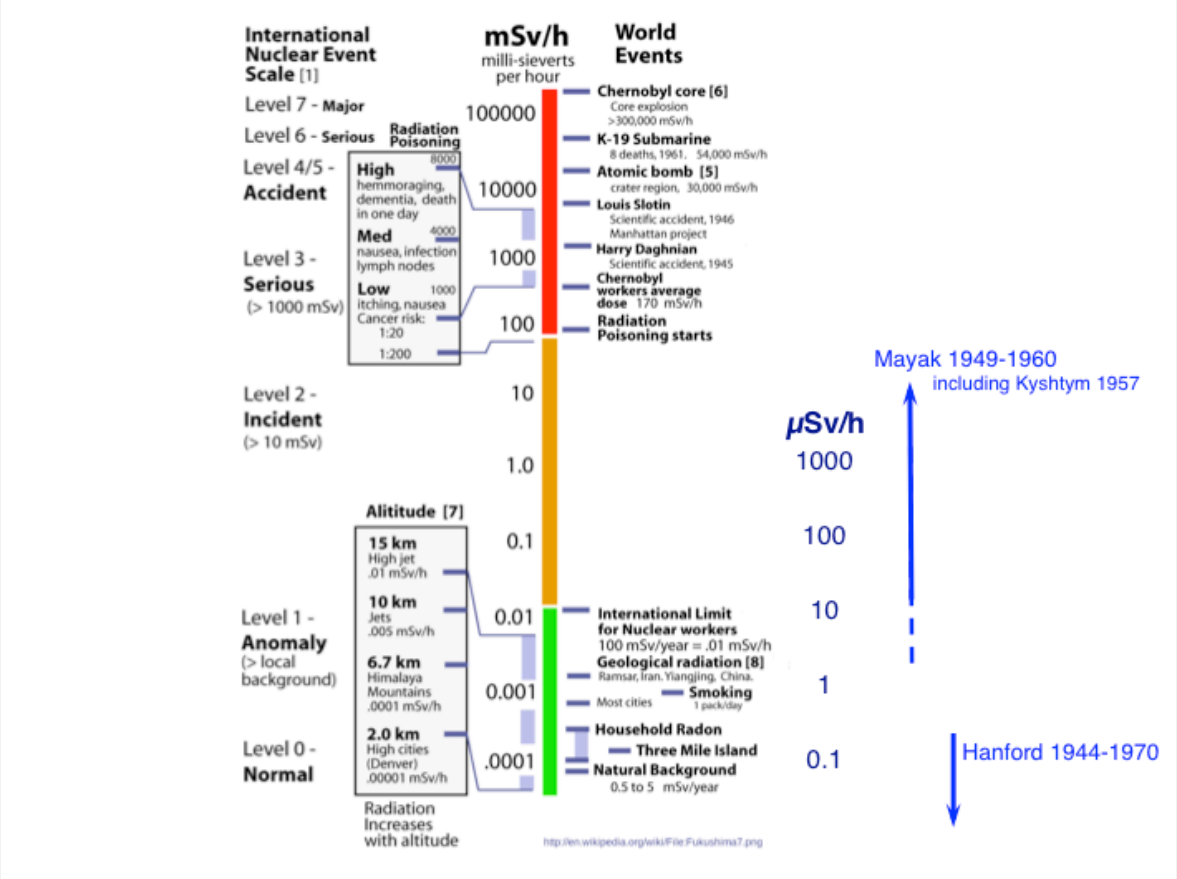

Figure 2-5: International Nuclear Event Scale and relevant past events (modified from `http://en.wikipedia.org/wiki/File:Fukushima7.png`). Emissions from Mayak between 1949 and 1960 (including the 1957 Kyshtym event) and from Hanford between 1944 and 1970 are shown on the right.

efficient than if the background were monitored by a single organization (e.g., a government agency). The likelihood of detection of a localized source is much increased.

An important but different use of background readings is the "tertiary" calibration described by Drukier, *et al.* [10], who show that it is possible to quantify the performance of γ-radiation readings from cell-phone cameras

by analysis of a large number of measurements. Because their software communicates the camera readings directly onto the Internet, these researchers were able to process large amounts of data from numerous types and numbers of sensors in many different locations. Remarkably, they find that they can calibrate the cell phones remotely, even though they have only limited information from each sensor (e.g., camera sensitivities and array sizes as well as exposure lengths are not known, and all of these span broad ranges). Their results are consistent, within necessarily broad uncertainties, with more direct measurements [7].

2.3 Trace Gases

The challenges in detecting trace gases of interest are both sensitivity and discrimination. Unintended releases of gases associated with nuclear proliferation or chemical weapons can produce distinct signatures, but the analysis methods are in general not yet accessible to individuals.

Certain other gases are amenable to public domain monitoring. Measuring CO_2 concentration is straightforward using infrared attenuation, with a per-sensor cost on the order of a hundred dollars. But the detection of radioactive isotopes of Krypton or Xenon or of HF (an indicator of UF_6 release) at low concentrations is not currently possible with simple, cheap USB-interfaced sensors.

The development of micro (1–10 cm) scale gas chromatographs and mass spectrometers is making them fit within the dimensions of smart phones [11, 12], although their cost, complexity and lack of consumer applications is likely to keep them from becoming everyday gadgetry at the present time.

Applying developments in high-discrimination compact sensing methods to treaty verification will likely require the targeted development of sensors that are optimized for each gas of interest. The verification community can perhaps leverage other government efforts in this arena, by the DoD and

DHS for example. In particular we advocate conforming to any interface and metadata standards that might emerge from these efforts.

We stress that there are areas of 100% overlap between certain treaty verification needs and battlefield sensing systems that are under development for the protection of the warfighter against chemical and biological agents. But deploying a wide array of open sources sensors with this narrow range of applicability is likely to be limited to areas of ongoing conflict and high tension. For example we can well imagine that the residents of Syrian cities would leap at the chance to deploy networked sensors for chemical weapons, were they available.

Another ongoing area of relevant development in the commercial sector is in medical peripheral devices. There is a strong overlap with some aspects of this technology and verification needs. In particular if device-to-smartphone interface and metadata standards emerge through the medical device market, it makes sense to adhere to these standards for any verification-optimized sensors.

Two specific technologies that might provide the combination of sensitivity and discrimination we seek are 1) high resolution optical and infrared spectroscopy, and 2) sensors that rely on highly selective chemical bonding. An example of a selective-bonding device currently under development is shown in Figure 2-6, in which an array of gas-specific sensors are monitored over a smartphone data link.

The infrared absorption spectrum of molecules frequently provides a distinctive fingerprint for detecting chemical species of interest. Figure 2-7 shows typical absorption spectra in the near infrared wavelength range between 700 nm and 3 μm.

A number of laboratories are pursuing small, compact sensors that use differential absorption for the detection of trace gases of interest. An example that is targeted at sensing chemical weapons is given by Holthoff *et al.* [14].

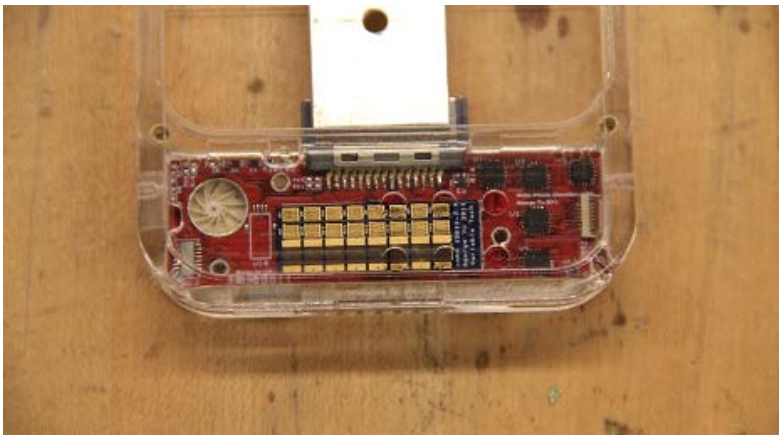

Figure 2-6: A gas sensing smartphone peripheral under development at NASA. From `http://gizmodo.com/5881097/this-is-nasas-cancer-sniffing-cellphone-sensor/`

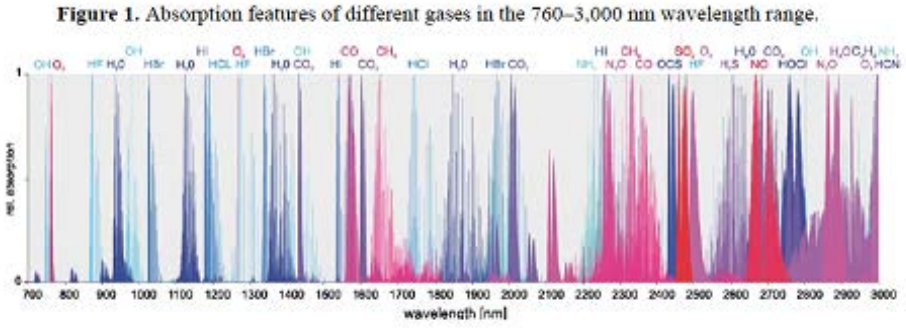

Figure 2-7: Illustrative infrared absorption spectra of molecules, indicating the potential for identifying specific species of interest. From [13].

Figure 2-8 shows a device that has been incorporated into stand-alone sensor packages with a wireless data interface. The principle is to modulate a diffractive element in the optical beam to rapidly switch between on-band and off-band wavelengths, thereby differentially searching for the spectral absorption signature of interest. This device is envisioned for use on oil rigs[1] to sniff for trace amounts of flammable gas, but in principle the same techniques could be brought to bear on detecting any trace gas other than noble gases.

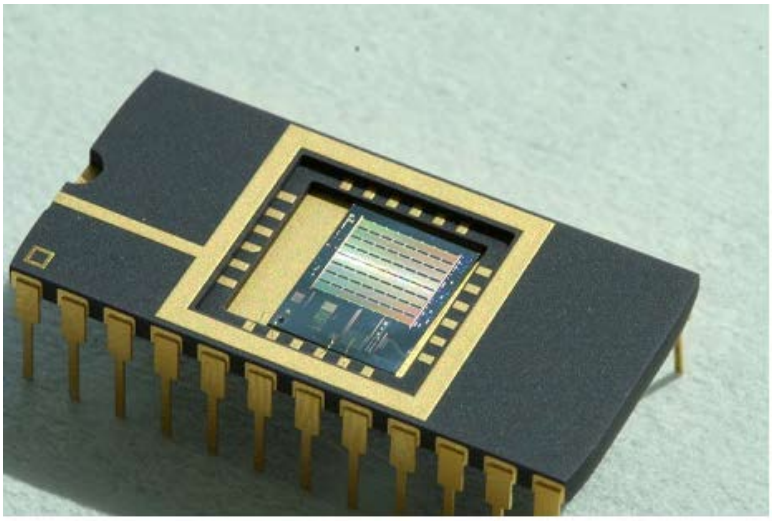

Figure 2-8: MEMS-based Infrared differential absorption spectrograph, for trace gas detection. From `http://www.esa.int/Our_Activities/Technology/MEMS_sensors_to_guard_oil_rigs_against_dangerous_gases` .

The trend toward environmental sensing of various gases of interest, ranging from pollutants to greenhouse gases, will likely drive this field forward. If infrared spectroscopic instruments are sufficiently generic, then extracting signatures of verification interest might be as simple as including appropriate template spectra in the public domain analysis software.

[1] `http://www.esa.int/Our_Activities/Technology/MEMS_sensors_to_guard_oil_rigs_against_dangerous_gases`

2.4 Audio

There are potential verification applications of acoustic information that is captured by audio recording devices, and then posted in a publicly accessible location. Here we have in mind such publicly available files as sound tracks from YouTube videos. Forensic analysis techniques can be brought to bear on establishing a precise time-tag for the audio track, and applying "voice recognition" techniques. In addition, transient acoustic signals of interest (such as explosions, sonic booms, and missile launches) can perhaps be extracted, and automated analysis of these transient signals (including as background noise) is worth pursuing as long-range (e.g., academic) research.

2.5 Commercial Satellites for Remote Imaging and Sensing

Satellite imagery of the Earth's surface has been publicly available from government and commercial sources since the early 1970's. Resolution (technically we refer to ground sample distance, the size of a single pixel on the Earth) began at 80 meters and has steadily improved, with the US government now allowing resolution as good as 0.25 meters. The 1990's saw a major change when the first private firms began to operate their own imaging satellites and sell the products. In the 2000's private radar satellites began to operate. Beginning about 2010 a further major change began, with new private firms developing satellites faster, with larger fleets of smaller satellites, with faster turnaround for imagery of specific sites, and with the eventual goal of daily imagery of the entire surface of the Earth. It is impossible to predict the outcome of this most recent period of change, or which of the new companies will survive. At present the main tradeoffs are among resolution delivered, frequency and pervasiveness of imagery, size of satellite, and number of satellites in use. The prospects appear excellent for such imagery to contribute importantly to public treaty monitoring (PTM), but different companies are exploring different parts of the trade space, and it is not yet

clear which will be most useful for PTM; the US government should track the trends closely and engage with some of the companies now.

The table below displays some of these parameters for five representative selected companies: Two (DigitalGlobe and Airbus) provide imagery services to governments as well as commercial imagery; two (Skybox and Planet Labs) are recent startups with exploratory business models; and one (Radarsat) provides radar synthetic aperture imagery rather than optical imagery. (Sources: web sites of the respective companies).

	DigitalGlobe	**Airbus Pleiades**	**Skybox**	**Planet Labs**	**RADARSAT**
Ground resolution	0.41 m (later 0.25 m)	0.5	0.9 m	3-5 m	1-100 m (radar)
NIIRS level	6-7	6	5	3	n/a
Swath width	16 km	20 km	8 km	10 km	500 km at 100 m res
Number of Satellites (current)	4	2	2	43	1
Number of Satellites (planned)	similar	2	24	>100	1-2
Revisit time	~ 1 day	~ 1 day	< 1 day	< 1 day	
Time to cover Earth	1 million km^2 per day, per satellite	1 million km^2 per day, per satellite	?	1 week	?
Sample Image	Figure 2-9	(similar to previous)	Figure 2-10	Figure 2-11	Figure 2-12

The sample images below illustrate the capabilities of different satellites, though the specified resolutions are actually several times better than shown

Figure 2-9: Digital Globe image of airport in Madrid, Spain, August 21, 2014. Picture width is $\sim$ 0.4 km. (Source: www.digitalglobe.com)

here. Other things being equal, one would prefer high resolution images such as Figure 2-9. But high resolution images necessarily cover a small area, so cadence of image collection will be faster for lower resolution images, especially for monitoring of and search in large areas such as entire countries. In the end one may prefer to have both frequent low resolution images for survey, search and change detection, and high resolution images of individual sites gathered on command as desired.

The radar product (see Figure 2-12) is clearly in a different category. It is not a direct image, but rather processed data overlaid on a map, to show flooding and wetlands in this example. Radar will usually require expert analysis. Its great advantage is that radar is always available: it works both day and night, and can "see" through clouds. For instance, it could be used to monitor activity at a site at night, or under cloud cover, when used at higher resolution than shown.

2.6 Natanz

The public unveiling and confirmation of the Natanz nuclear site in Iran was an important early example of PTM. In 2002, the National Council of Resistance of Iran (NCRI), an émigré dissident group which aspires to be a government in exile, publicly announced that the government of Iran

Figure 2-10: Skybox image of airport in Nice, France, December 7, 2013. Picture width is $\sim$ 3 km (Source: `http://www.firstimagery.skybox.com`)

Figure 2-11: Planet Labs picture of airport in Beijing, China, Jul 25, 2014. Picture width is $\sim$ 10 km (Source: www.planet.com)

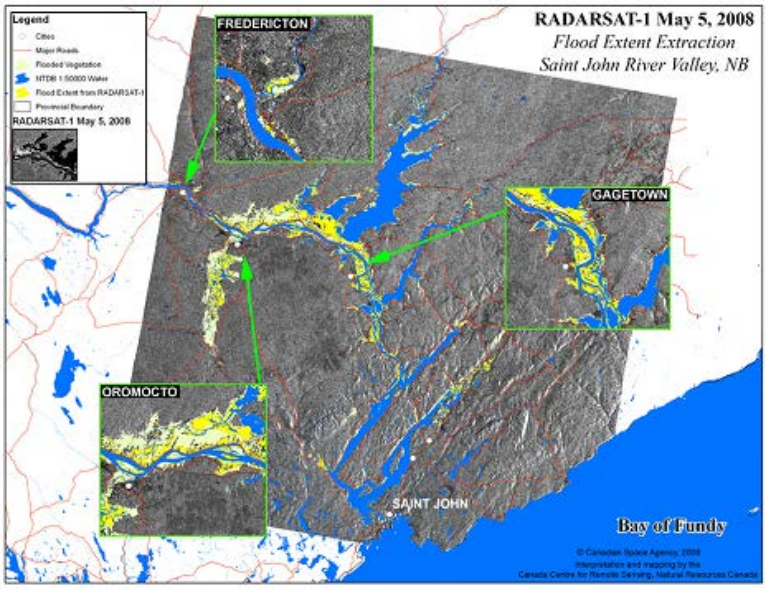

Figure 2-12: Radarsat-2 image product (Source: `http://gs.mdacorporation.com/`)

was carrying out clandestine nuclear activities at Natanz and Arak, Natanz being called a fabrication site for nuclear fuel. Later in December 2002, the Institute for Science and International Security (ISIS), an independent policy institute in Washington, DC, released commercial satellite images of Natanz, and based on these images assessed it to be a gas-centrifuge plant for uranium isotope separation, which turned out to be correct. (Sources: `http://www.isisnucleariran.org/www.ncr-iran.org/`)

More recently and more generally, the ISIS NuclearIran web site provides an impressive public interface for information about many nuclear sites around the world at `http://www.isisnucleariran.org/from-the-sky/` as an overlay on Google maps. The user can zoom into the individually marked sites on the standard Google satellite imagery and read attached information. The Ages of the images were not provided but they are likely to be a year or two old. The map is already a creative platform for public communication. In the future, when whole-world satellite imagery may be updated on a time

Figure 2-13: Natanz nuclear complex in Iran, imaged by GeoEye on Aug 12, 2006 (Source: `http://www.isisnucleariran.org/sites/detail/natanz/`)

scale of weeks, days or hours, such a platform could become a basis for open and public discovery and analysis of such sites, or changes and levels of activity therein. By cloning such sites, US government agencies could provide their own confidential overlays, whenever necessary, to benefit from the public analysis without revealing US interest in it, or revealing other non-public US information that is correlated with it.

For the specific verification challenge of identifying, counting and geolocating specific delivery systems or other large items of interest, this imminent flood of rapid cadence imaging opens up a new set of options. Whether the government elects to purchase images from these vendors, encourages NGO's to obtain and analyze images, or both, is a matter of verification implementation policy. We also note that the Skybox business model is to sell derived knowledge – that is, professional analysis – not just images. It

might be useful to compare cost and performance of the photo interpretation results from the commercial sector to having government analysts look at the same images.

Another impact of this revolution is the opportunity to share all this information widely, without the classification restrictions placed on NTM images.

2.7 Amateur still images and video, from the ground and from the air

The final category of public domain information we explore is images, both still and video, obtained by the general public and posted for unrestricted and open access. An interesting example that illustrates this was the flood of images from dashboard cameras that were posted on the Internet after the Chelyabinsk meteor event in February 2013. Ionov [15] used the archived video data from dashboard cameras to determine the trajectory of the meteor and its fragments through the atmosphere. This illustrates the potential of extracting quantitative information from redundant video data.

There are numerous non-trivial challenges to using the barrage of image and video data that are uploaded daily for the purpose of treaty verification. However, one advantage is that it is highly structured information, especially when compared to much of the information on the web, and is frequently accompanied by geo-tagged and time-stamped metadata. Moreover, there is considerable academic research underway in the area of image analysis, (e.g., Zhu *et al.* [16], is but a single recent example of a large and active area of commercial as well as university research).

One potentially interesting regime we foresee is public domain image data obtained from UAVs. There is a growing hobbyist community (perhaps soon to be augmented by an activist community) that obtains images from UAVs.

2.8 Finding and Recommendation

Finding: *The proliferation of inexpensive, networked sensors provides significant new sources of information enabling pervasive monitoring and societal transparency. The potentially large number of observations made by the public can in many instances make up for crude instrumentation. Users are motivated to communicate environmental information because they benefit from early warning and hazard prevention provided by other subscribers.*

Recommendation: *Government should track public-sector activities in sensor development and data sharing in support of increased transparency, facilitating these – as appropriate – through targeted calls for research proposals and through challenges rewarded by recognition and prizes.*

3 INCENTIVES FOR OPEN SOURCE TREATY VERIFICATION

This section discusses potential incentive issues that might arise in trying to develop mechanisms for open source treaty verification.

Crowd-sourcing might prove useful for treaty verification in several different ways. One possibility is to rely on the efforts of citizens to collect and transmit information that is relevant for verification. Individual participation could mean taking measurements or images, or noticing irregular activities, and then sharing the data either publicly or with governments or organizations involved in treaty verification. A second avenue is to encourage people to take public data (satellite imagery, social media posts, etc.) and look for patterns or signals that might be informative about treaty violations or compliance. In this case, citizens worldwide might participate.

These types of open source engagement involve rather different incentive challenges. We start with crowd-sourced data gathering, which is both an easier and a harder problem, and then consider crowd-sourced data processing, where we can draw on parallels with existing open source efforts.

3.1 Incentives for Open Source Data Gathering and Reporting

The reason that open source data gathering is an easier problem is that, increasingly, at least some of it will happen as a matter of course. More and more people carry mobile phones, are connected to the Internet, and actively use social media. Even with no specific effort to create incentives or crowd-sourcing mechanisms there is likely to be a wealth of images and sensor data publicly and freely shared from practically every country and region in the world.

Incentive issues arise if one wants to go beyond this and ask citizens

to consciously gather information and report on the activities of their own governments. This could involve substantial effort or personal risk, and even if people are motivated, they might need appropriate equipment and safeguards (e.g. technology to measure factory emissions or effluents, direction about what activities to report, the ability to send reports anonymously, and so forth).

In principle, one can imagine trying to create direct incentives, by providing financial rewards or recognition for certain images or measurements. For example, the US Rewards for Justice Program, established in 1984, provides rewards for information on terrorist-related activities.[2] The program has paid more than $125 million to informants. Of course even large rewards do not always work in motivating informants. After 9/11, the US offered a $25 million award for information leading to the capture of Osama Bin Laden. But the reward was not paid, and when Bin Laden was killed in May 2011, the Administration stated that the information that led to the raid came from electronic signals intelligence.

Providing information on illegal nuclear or biological or nuclear weapons development arguably is similar to reporting on terrorist activity, with the difference that in the former case, citizens may be reporting on their own government (they might, alternatively, be reporting on a terrorist or other criminal organization). As a result, promising rewards could run the risk of making data collection more difficult. A publicly posted award for pictures of military activities or weapons systems could make it more risky to take pictures, or even to carry a camera or phone in sensitive areas.

A less direct but potentially effective approach might be to facilitate data collection by providing people with information on what to look for or how to gather data that would be relevant for treaty verification. For instance, in a situation where a government uses chemical or biological weapons on its own citizens, or has alienated a significant portion of the citizenry, motivation may not be the problem. The issue may be providing an appropriate

[2] http://www.state.gov/m/ds/terrorism/c8651.htm

set of tools so that people can gather evidence of wrongdoing, and report it in a credible and fully verifiable way without betraying their activities.

Chemical and biological weapons research is a case in point, because the relevant professional organizations have well-established ethical guidelines precluding such activities. International law, professional ethics and cultural mores can provide strong incentives to monitor activities, materials and people considered beyond the pale, implying that there can be significant value in providing tools to facilitate such monitoring. Examples of such tools include the sensors described in the previous section; additional examples are given below.

3.2 Incentives for Open Source Data Processing

The second type of open source engagement, crowd-sourced data mining, has parallels in a whole range of open source efforts over the past few decades. There are plenty of successful examples, and these efforts have involved a range of different participation models and incentive mechanisms.

Some famous crowd-sourcing efforts involve financial prizes. A celebrated case is the Canadian mining firm Goldcorp, Inc.[3] In March 2000, Goldcorp's Red Lake mine was struggling and the company was unsure how to improve the mine's performance. The company started the Goldcorp Challenge, which offered prize money of $575,000 to geologists who could help identify promising locations using the company's geological data. The data were posted online, and more than 1400 geologists from 50 countries entered the challenge. Two Australian firms partnered to win the competition and the company successfully drilled their top targets, increasing the mine's yield by an order of magnitude.

A similar approach was used later by Netflix,[4] which posted a large data set on user movie ratings online and promised $1 million if someone could

[3] http://www.fastcompany.com/44917/he-struck-gold-net-really
[4] http://en.wikipedia.org/wiki/Netflix_Prize

improve by 10% on Netflix's movie recommendation engine. The contest started in October 2006, and within a week, a team had improved on Netflix's system. The winning effort came two and half years later after several leading teams merged their efforts.

Similar competitions – both with financial prizes and without – can be found today on the web site Kaggle.[5] Current examples range from predicting which shoppers will become repeat buyers (financial reward) to estimating fire insurance losses (financial reward), and to separating Higgs Boson decay signals from background data using data from the ATLAS experiment conducted at CERN (no financial reward).

All of these competitive mechanisms share some key ingredients. They provide participants with data and specify a clear method for evaluation (for instance, the mean-squared error of a predictor). And they provide some motivating reward – a financial prize, the prospect of participating in something of social value, public recognition, or a combination.

One feature of these examples is that they have the goal of attracting a relatively small number of highly qualified people to work on a particular problem. The sense in which there is "crowd" participation is not that that there are millions of entrants to a competition. It's that the identity of the participants is not known in advance – entry is open and efforts are parallel, so that in comparison to hiring a team or issuing an RFP and selecting a single contractor to work on a problem, the sponsor motivates a range of people, and gets to see who makes the most progress.

A similar participatory model is used in open source software development. The number of developers working intensively on a piece of software at a given time may be small, but if someone shows up with a great innovation to include in the software, it is likely to be adopted. One difference is that there is not a clearly specified evaluation process for open source software. Instead there is a governance process to decide which contributions to include

[5] `https://www.kaggle.com/competitions/` Accessed July 12, 2014.

in any new release of software, and not every contribution is included.

There are also examples of successful open source efforts that attempt to encourage more widespread, and perhaps less expert, public involvement. A good example is the case of user-generated Internet content, such as Wikipedia, or Amazon reviews. Another example is the Zooniverse citizen science project,[6] which makes large data sets available and asks volunteers to engage in transcription of texts (e.g. the Ancient Lives project to transcribe papyri), or pattern recognition (e.g. the Planet Hunters project to identify planets using NASA's Kepler dataset), or other activities that still cannot be automated effectively using computer data-mining methods.

These efforts differ from the preceding examples in several ways. The participation model is different – the projects try to engage many non-experts, even if for limited amounts of time, rather than a handful of experts devoting large amounts of time – and so is the incentive model. There is no financial reward for updating Wikipedia or writing online reviews on Amazon (or Trip Advisor, Yelp, etc.) or for participating in a Zooniverse project. Instead, there may be some intrinsic reward or social recognition from becoming a top reviewer on Amazon, or a co-author on a Planet Hunters paper reporting on a planet discovery. And there is the opportunity to participate in a collaborative project of social value.

One point to emphasize about the "widespread" crowd-sourcing examples is that the incentives of people to participate may depend on how big a contribution they feel they are making. Someone who writes a Wikipedia article gets to see his other writing published immediately in one of the most prominent places on the Internet for everyone in the world to read. If each contribution had to go through a review process with a 5% acceptance rate, there would probably be a lot fewer editors.

To take another example, the Asteroid Zoo project asks volunteers to help in finding unknown asteroids, by looking at Catalina Sky Survey images

[6] `https://www.zooniverse.org/` Accessed July 12, 2014.

and classifying the contents.[7] Presumably only a small fraction of people who look at the web site will actually discover an asteroid, but a large number of negative reports on an image may still be valuable. Whether the project will be a success probably depends on whether people find it sufficiently fun and rewarding to contribute even though their prospects of new asteroid discovery are pretty low.

How could these examples be applied to crowd sourced data mining to improve treaty verification? It depends on what task is actually needed. If what is needed is better data processing models to take Twitter or Instagram feeds, or geologic or atmospheric or satellite imagery data and look for relevant signals, then the former set of examples seem more relevant. The lesson seems to be to offer talented people a well-defined task and the tools or data to work on it, a benchmark against which to measure progress, and some prospect of a reward, whether financial or not. The output might not be actual detection of treaty violations, but the development of better data mining methods.

On the other hand, if what is needed is not computer models, but human effort and the gradual accumulation of small amounts of data processing, then the latter examples are more relevant. The lesson of these models seems to be to make participation easy, fun and collaborative, and at least somewhat rewarding. Again, the output might not be the detection of violations, but the identification of places to look or not look with more expert methods.

[7] `http://www.asteroidzoo.org`

4 OPEN SOURCE INFORMATION AND ANALYSIS

This section contains the study's response to the seven objectives listed in our Statement of Work that defines "areas of potential research and development pertaining to the use of public domain data to support treaty verification."

4.1 Data Sources

"What open and crowd-sources exist and lend themselves to exploitation for treaty-verification-relevant analysis? How might these sources be exploited? What are the practical and technical limitations to use of these sources?"

The Bureau of Arms Control, Verification and Compliance (AVC) in the Department of State[8] has prepared an unclassified document describing its priority needs for R&D programs, addressing future as well as continuing needs in treaty verification and transparency, and in identifying covert activities threatening to U.S. national security. Our focus is on the emerging and rapidly expanding interaction between technology and citizens equipped with powerful and widely distributed capability for collecting and sharing information, and also for crowd-sourced analysis. Important advances will include better tools – sensors, communications infrastructure, analytical methods – to acquire, communicate and analyze information free from spoofing and deception, and to preserve participants' anonymity to protect them from retribution.

Listed below is the variety of common sources that are contributing copiously, from sensors to social media.

- Posting systems: These are systems that allow users to share information in a very public way. They include Facebook, Twitter, and their

[8] http://www.state.gov/t/avc/

foreign equivalents; photo sharing; there are many other sites that allow "sharing" of various interests: music, videos, group chats, restaurants, activities, friend location; video sharing (YouTube is the most popular, but there are many).

- Personal sensors: The development of inexpensive sensors has led to an explosion in the sensor space. While mobile phones are the most obvious, they are not the only personal sensing platforms (location, accelerometers, camera, microphone, and potentially others); locators (devices the size of a coin) that provide location of keys, phone, shoes, ... There is also the notion of the "instrumented self" that provides information about the person wearing the sensors: accelerometer, blood sugar, heart rate, blood oxygen, posture and other data that might be useful for detecting the release of harmful substances.

- Vehicles: The explosive development of sensors has also affected vehicles, and these are valuable for detecting movement of material. More generally, vehicles have a large number of sensors, including location, cameras, microphones, mm-wave radars (for parking and collision avoidance). A Ford executive recently said[9] that they knew the location of every Ford car built since a certain date in real time. Insurance companies are now putting location and recording systems into cars, and offering discounts for people willing to install them. Oregon is experimenting with using GPS to charge per mile to calculate the gasoline taxes, and this will provide a track of the vehicle location.

- Pervasive sensing: Sensors are being installed in every imaginable location, including "smart meters," alarms and cameras in and around houses; video monitoring of many and eventually most public places; chemical sensors, all of which are getting cheaper and with improved sensitivity and accuracy. Wireless communication makes it easy to communicate with these sensors, and leads to the "Internet of things."

[9] `http://www.businessinsider.com/ford-exec-gps-2014-1`

- Public record data: In the past, government records have been difficult to access, but with the shift to electronic records, all of these are now available. These include municipal, county, state, national and regulatory data, in the United States and abroad. Environmental regulations require extensive monitoring by companies, leading to many additional signals.

- Imagery: Traditionally, we have relied on NTM for important tasks such as treaty verification. But recently there has been a revolution in publicly available commercial imagery, Skybox, Planetlab, etc. are of growing importance and provide a complementary capability to NTM. In addition to the imagery, these companies will provide analysis that can then be used to trigger further investigation. We also expect to see imagery from drones.

- Commercial data: We hear constantly about "Big Data" and some of the biggest data is that collected by commercial entities, including but not exclusively retail and distribution channels. There is an enormous amount of information behind corporate firewalls, including all the data regarding communications (mobile telephony, regular telephony), subscriptions, advertisements clicked, web sites visited, prescriptions, groceries, and other purchases.

We need to keep in mind that it is often not a single data source but the fusion of several data sources that allow us to draw inferences.

Social media provide an immense and ever-expanding collection of data, much of it available to the public. It is not only the vast amount of data, but its variety that makes it interesting for our purposes. What we find in social media tomorrow will be beyond what we see today as new ideas for "sharing" occur, and the willingness of the public to "share" nearly every aspect of their lives continues to increase.

Social media represents a moving target: their rapidly evolving nature means that it is impossible to provide a complete discussion of the scope

or even an accurate inventory of their current state. But a representative discussion of a few of the key social media sites provides an indication of the opportunities and challenges that are presently available.

Consider, for example, digital images. Current estimates are that 10% of the total number of photographs in the history of the world were taken in the past year, and the rate is increasing. Nearly every mobile device has a multi-megapixel camera. There can be significant detail in the background of seemingly uninteresting pictures, so the large numbers of images is a real factor when considering the value of open-source data. Instagram, one of the popular photograph sharing sites, held more than 20 billion images as of June 2014, and more than 300 billion images have been loaded onto Facebook, with more than 350 million new images added each day.

Social media amount to more than photographs and videos, but also contain written content and audio. While much of the written content is repeated or may be of low information content, important trends can often be spotted. In addition, many organizations – from NGOs to self-identified terrorist organizations to individuals who may be unaware of the importance of their information – post, tweet, share, like and maintain all manner of information. As of September 2013, Facebook had 4.75 billion status updates, wall posts, photos, videos, and comments posted each day; in addition each day there were 4.5 billion "likes" (indications of interest or assent), and 10 billion messages sent. For Twitter, in 2013, there were more than 500 million "tweets" per week and a peak of 143,199 tweets/second, and that continues to increase. The challenges will be to extract meaning from this flood of information and to separate the signal from the noise. Interestingly, what might be considered clutter may actually provide additional information that will be useful in validating data, as discussed later.

The fact that today's data are but a small subset of what will be available in the future emphasizes the need for agility in designing an effective system of searching, collecting, archiving and analyzing large quantities of data.

Finding: The advent of powerful technologies (e.g., smartphones) and associated societal connectivity through the Internet is proving to be a game changer in the national security arena because of the enormously increased quality and quantity of openly available information. We anticipate that this will extend to the domain of arms control and treaty verification as well.

To be effective, however, this emerging interaction between citizens and the new technology requires careful analysis for applications to treaty monitoring, and more generally for enhancing transparency in order to track activities that may threaten national or international security outside the domain of treaty regimes (e.g., proliferation, terrorism, organized crime, etc.). Incentives may or may not be needed for public contributions to open data (e.g., environmental or other societally relevant information, perhaps coupled to an "early warning" system, may provide reason enough for individual contributions), but appropriate approaches to incentives need to be defined, as summarized in the previous section.

In this regard, specific treaties in general have specific requirements, some of which may be more or less amenable to open-source information and analysis than others. It is also useful to identify technologies and operational procedures that improve prospects for success, for example by improving sensitivities and avoiding (to the degree possible) the need to incentivize broad participation.

Recommendation: Analyze and prioritize arms control agreements and diplomatic initiatives to which open source information or analysis could contribute most effectively.

4.2 Data Integrity

"How do we detect spoofing and deception? The JASON study will examine possible methodologies to find and vet relevant open

source data and metadata, and lay out the architecture for implementing an optimal mix of these methodologies in a future automated open source validation tool."

The briefings we received from commercial and academic institutions to interpret and validate the integrity of data flowing in from open sources made clear the enormous difficulties still to be faced and work to be done before claiming confidence in the reliability of data obtained from open sources.

There are at least three aspects to data integrity that must be considered in the present context: i) ensuring the validity of raw input from sensors or people; ii) ensuring integrity of data in storage and transmission; and iii) making data reliably available through proper archiving and broader access or dissemination. We specifically do not consider the integrity of analysis performed on these data: good data may or may not be subject to faulty analysis. Good analysis also means ensuring that the input data are suitable for the problem of interest. In short, we consider data to be valid if they are not misrepresented, whether or not they addresses a topic of interest.

The most difficult step is the first, assuring validity of raw input, as the other steps are well addressed by current technical capabilities in cryptography and data management. Regarding the other two topics, we note that the IMS is an example of successfully ensuring data integrity in transmission and storage, all in the context of an international treaty. Perhaps more challenging at the present time is assurance of proper archiving and dissemination, which requires effective communication with the ever-changing public at large as well as with evolving technology (e.g., for storage). These are among the challenges already being addressed by government agencies, however, so do not call for approaches unique to open-source information or analysis.

The validity of the original data is especially problematic with social media, in that eyewitnesses are notoriously unreliable (especially when reporting on unanticipated or extreme events [17]) and there can be strong

reinforcement of messages – whether true or not – when a topic "goes viral" on the Internet. Also, much distributed (crowd sourced) sensing may be performed with readily available and inexpensive detectors that are far from the technical state of the art and may or may not be calibrated.

In addition to such errors, taken here to be honest mistakes and limitations of sensors, there is the possibility of willful deception and spoofing of the raw data, including through conscious insertion or duplication of false information. This process can be automated; multiple re-postings of similar (let alone identical) information is simply no guarantee of data validity.

Finding: The only reliable means of validating raw data is through confirmation from independent means.

Even one independent replication of a datum may be sufficient for validation or, alternatively, proof of corruption. Multiple independent sources reduce the possibility of error, so it is preferable to have multiple sources than just one. The critical issue comes down to validating the independence of the confirming information or source.

Independence is of primary importance, as distinct from raw numbers of observations. This conclusion brings into focus the significance of crowdsourcing and big data, both in sensing and analysis of information. For purposes of vetting, it is not the large (reported) number of observations or participants that is in itself important, but the degree to which large numbers increase the chances of there being independent input or analysis that matters.

However, it is also true that large numbers of independent observations typically allow random and uncorrelated errors to be reduced. This is accomplished by way of well-established statistics. Independent observations of a given type provide improved accuracy by reducing "random" errors. If the errors are uncorrelated and have zero mean the uncertainty of their mean is proportional to $1/\sqrt{N}$ for N observations (100 times more independent observations reduce random errors by 10). Error or uncertainty decreases

less rapidly with increasing number of observations when correlated (non-random) or systematic errors are present, and may not decrease at all in the case of highly correlated errors (e.g., systematic biases). Still, observations of sufficiently different types can uncover systematic errors, and thereby improve absolute accuracy. Validation amounts to removing biases (improving accuracy) by comparing independent types of observations.

Just as we distinguish between random and systematic errors, it is important to distinguish between errors and clutter. Although related, these are distinct concepts and clutter can in fact be beneficial in calibrating and even validating observations. This is especially important in mitigating deliberate spoofing or deception.

Recommendation: Identify independent validation of information from all-source capabilities, and including checks for internal consistency from clutter analysis.

Recommendation: Recognize, and if possible support or reward, efforts by others (NGOs, academics, etc.) to calibrate and validate treaty-relevant sensors and other potentially useful sources of open information and analysis.

4.3 Signal Clutter

"In heterogeneous and distributed sensor networks, how does one find meaning in all of the noise? The JASON study will assess possible approaches for using existing commercial or government tools or creating more customized tools to strain and analyze information present on public-domain open source data."

Clutter refers to the actual heterogeneity (variability) in space, time or other sample characteristics of the system being observed. In contrast, error refers to a combination of unreliability and lack of resolution of a measurement, observation or analysis. For example, the sea surface is rough if

measured at scales of cm, and is at least as rough if the measurements are made at scales of mm: roughness is a form of clutter that is intrinsic to the sea surface. Multiple independent measurements can decrease the error (increase the reliability and resolution) of such observations, but cannot reduce the scatter in observed values due to clutter.

As such, clutter can in many circumstances define the initial limit of resolution of a set of measurements or observations. Increasing sensitivity does not help, as this variability is intrinsic to the system, and if the desired signal is small relative to the clutter one is confronted with the classic problem of searching for a needle in a haystack. In this sense, clutter provides a baseline for the maximum sensitivity or resolution that can be achieved.

This last conclusion is true if one knows nothing about the clutter. Clutter, however, can be a form of information, such that if it is well characterized one may be able to: i) remove much of it (for example, by averaging), so as to successfully find the needle in the haystack; and ii) use the clutter as at least partial validation of the data.

An image of an object of interest including confusing or obscuring elements in the background and foreground, respectively, offers an analogy. If enough is known about the foreground and background, these elements can be removed so as to highlight the object of interest. There may still be loss of information, if part of the object has been obscured, for example. However, a good model of the clutter can serve to enhance the signal in an otherwise overwhelming background of "noise" due to clutter.

Better yet, the clutter may reveal other information about the image. The background of a photo image may provide indicators of location, time of day or other items (including people) of potential interest. This associated information is not necessarily reliable, but itself can be used to test the validity of the image. Unfortunately, it is often the case that such information can only be used to check for consistency (e.g., that the image was taken where and when purported), rather than to provide truly independent validation.

Still, multiple consistency checks of clutter may provide enough confidence to effectively validate the original data.

A case in point is the background noise in seismic records (i.e., the small but non-zero readings of seismometers between earthquakes or other seismic events). The ambient seismic field, sometimes referred to as seismic "daylight" (or "background noise"), is now being extensively studied, so that it is difficult to spoof a seismic record without losing consistency with records obtained from other stations. The background of small or distant earthquakes, as well as signals of storms, volcanic eruptions and other natural phenomena, adds to the clutter of this ambient field in such a manner that seismic records can in many respects be considered self-validating, at least in principle. Now that it is understood, there are automated means in effect for analyzing the ambient seismic field, and of utilizing this "clutter" to map the subsurface and even monitor it with time (e.g., monitoring extraction from oil and gas reservoirs).

There are at least two additional aspects of background clutter that are important for the topic at hand. First, the background can be used to calibrate sensors, as illustrated by Druiker, *et al.*'s [18] self-calibration of cell-phone cameras for sensing gamma radiation (note that it would be an indication of suspicious data – flawed or possibly spoofed – were such "tertiary" calibration to produce highly anomalous results when compared to laboratory measurements). Second, background clutter can be used to cross-calibrate between different types of sensors or sensor networks. In this sense, the capabilities of a sensitive but more focused system (e.g., NTM) can potentially be greatly enhanced through judicious combination with a large, low-sensitivity, poorly calibrated network of ubiquitous sensors, the "clutter" (as well as individual events, if available) providing necessary tie-points between the two sources of data.

Background radioactivity, which is notoriously variable in space and time, is a good example of a clutter field that, if well characterized by a dense distribution of sensors, could turn from being a problem to being an attribute

of future radiation-sensing regimes. This is an example of how large numbers of crude sensors can potentially beat the performance of smaller numbers of higher-quality sensors.

None of these approaches is absolutely reliable, in the face of noise, equipment failure and other realities of collecting real observations from the field. Still, they make the point that clutter can provide a useful signal if it is well characterized. The key is to invest in thorough studies of the clutter associated with any data (or analysis) one is pursuing, both to enhance the signal relative to background noise (find the needle in the haystack) and to help validate the integrity of the raw data. There is no known means of automatically nulling or using background clutter; however, once properly characterized, it is possible to apply automated procedures to reduce and even utilize this clutter.

Finding: Clutter can, if properly characterized and understood, prove useful for calibrating and validating open source information.

Recommendation: Characterize the clutter in measurements of interest for purposes of calibration and validation of open-source information and implementation of automated processing.

4.4 Standards:

"What kind of uniform standards should one use for both data and metadata? The issue of data standards is integral to the design of the open source validation tool and aid to uniform and rapid signal processing."

Because we are in the midst of a rapidly increasing volume of data generated by social media, and the relentless evolution towards globally available inexpensive, miniaturized sensors, modern technologies can now greatly

increase the ability to verify treaties and to monitor activities potentially harmful to U.S. and international security.

This time of rapid development would be perhaps the "worst of times" for government to attempt setting uniform standards for both data and metadata. Most of the expertise that is driving the revolution in technology is in the commercial and academic sector, and is driven by scientific and technical curiosity, as well as by commercial interests. Standards are ultimately necessary, but protocols – de-facto standards – are rapidly evolving in an era characterized by agility and diversity (of interests and motivations, as well as means and media) rather than uniformity.

Government can best help keep pace with progress by sponsoring and/or participating in the open conferences on new developments, supplemented by its own specialized meetings for those interested in addressing the nation's specific challenges. This approach also applies to other aspects of making effective use of the new technology, including integrity of data and signal processing. In short, any system developed for the government must be flexible and adaptable, not following the existing procurement model but striving for the agility that one observes in technology companies.

An example of success in this area was the development of the TCP/IP protocol stack. It was an open process, focused on ease of implementation and not over-specified. It is essential that all such standards have interoperability as their primary metric for success. This is embodied in what is known as Postel's law, first articulated in RFC 760, the standard for the IP protocol: "an implementation should be conservative in its sending behavior, and liberal in its receiving behavior". This approach has served the Internet well, and is restated as the Robustness Principle in RFC 1122, which defines the lower layers of the Internet: "Be liberal in what you accept, and conservative in what you send" (NRC [2014] [19] offers a good primer on Internet standards in the context of security).

The commercial and public standards processes operate at several different levels. Organizations such as the Institute of Electrical and Electronics Engineers (IEEE) are good at developing standards for relatively low-level devices that are manufactured and must interoperate, for example IEEE 754 for floating point arithmetic or IEEE 802.11 for wireless networking. Similar standards exist in the medical community, promulgated by various organizations, including IEEE. Each has a formal process that works more or less well for that standards body, but if it is functioning properly then the issuance of the standard enhances functionality, reduces cost, and has only minor impact on innovation. New sensors are developed and make use of these standards, dozens of them, as a way to more rapidly produce innovation. The combination of many components, results in a new device and perhaps a novel data source (that one day may result in a standard). Standards do not come before the innovation, but an innovation may become a standard.

JASON believes that the process of developing standards, and by this we mean specifications that allow for the interoperability and exchange of data, necessarily must be open and collaborative. There are several examples of success in this area, with the best-known being the Internet. The Internet is perhaps the greatest interoperability success story in history, and this can be credited to the process by which Internet standards are created through the Internet Engineering Task Force (IETF). The World-Wide Web Consortium (W3C) has a similar process, and issues some of its standards jointly with the IETF.

Both the IETF and the W3C function through the open process of creating and commenting on and revising draft proposals. The IETF drafts are called "Request for Comments" (RFC), and interestingly remain as such even when they are an adopted standard. The W3C has a more formalized system with Working Draft (WD), Candidate Recommendation (CR), Proposed Recommendation (PR), and W3C Recommendation (REC). In all cases the process is open to all interested parties and so the standard, or recommendation, is fully vetted. As with cryptography, the best approach is

to subject the proposal to open public scrutiny.

What this means for government is that the primary focus should be on identifying and adopting the most common and effective standards, and finding ways of translating databases or analyses from one set of standards to another. Although there may be reason to translate into a set of specialized governmental formats for specific forms of analysis, we advocate working to the maximum degree possible with the protocols developed in commercial and public arenas. For example, there may have been a time when geospatial information would be formatted in a government-specified manner, but such standards as are in Google's Keyhole Markup Language (KML) system are now so prevalent that government analysts have to know how to use it and may as well do so for their own data, to the degree feasible.

This is no more than advocating for software what is already happening with hardware, whereby commercial-off-the-shelf (COTS) is systematically replacing government-specified standards (e.g., MILSPEC). There are specific instances when this is not possible, but in general the drive toward increasing use of COTS is not only motivated by enhanced efficiencies and cost savings, but simply by the fact that government-specified is no longer available. Just as reliable integration of COTS is a requirement for modern hardware, there will be a premium on developing the capability to translate between different (non-governmental) standards and protocols when it comes to handling databases and software.

Finding: Standards are evolving rapidly, driven by public and commercial interests, such that government needs to be responsive to keep up: flexibility and agility are key to exploiting data in an effective manner.

Recommendation: Implement a strategic plan for keeping up with evolving protocols developed and applied in the public (especially commercial) sector, and for translating among these formats.

4.5 Cooperative Verification and Monitoring:

"What are the limitations on cooperative observation, validation and analysis? Are there models for cooperative sharing of verification and monitoring activities in this regime?"

Key to validating the findings of open sources will be confirming the independence of two or more of the sources. Multi-reporting – "retweeting" – or posting of the same information is no substitute for establishing credibility, though repeating the same information can provide an indication of importance or urgency (e.g., the occurrence of the 2008 Wenchuan earthquake was first noted in the United States as an anomalous increase in text-messaging: information moving at nearly the speed of light, so arriving well before the seismic waves reached US seismic stations). In cases where the open sources are used as a trigger for further investigation, perhaps by the intelligence community using its traditional means, validation is still important in setting priorities.

This observation points to the importance of developing a strategy for effective partnership with the private open source community, both for the sake of agility, and because much of the expertise and most of the infrastructure are outside government, in the commercial and academic world. In this sense one can think of a parallel with the aerospace industry that serves and supports U.S. government capabilities.

Open source cooperation between governments is also widely practiced currently in observation, validation, and analysis. Two examples are in countering crime and narcotics trafficking. With the broad proliferation of sensitive, inexpensive sensors there will be more areas that can benefit from open-source cooperation: e.g., limitations on environmental contamination by societies, natural disasters caused by earthquakes, tsunamis, or escaping nuclear radiation, and other effluents from industrial accidents.

In the area of arms control, we anticipate that arms control treaties will offer one means of addressing our national security needs, expanding beyond bilateral U.S.-Russian to multilateral agreements. More effort will need to be devoted to confidence building measures, with multi-national participation that requires monitoring activities, not just counting physical objects. These activities include, for example, efforts to covertly acquire or produce nuclear bomb fuel, the most difficult challenge en route to fission weapons; and military build-ups in regions of tension. Current international cooperation in the cyber domain points in the direction that needs to be pursued, but also illustrates the broad challenges that remain (e.g., successes in certain domains, such as constraining money laundering or child pornography, have yet to overcome fundamental disagreements in other aspects of cyber security and stability).

Since the demise of the Soviet Union in 1991, we have made significant advances with the Russian Federation and many other nations in cooperation and transparency in verifying treaty agreements. Both the United States and the Russian Federation have successfully conducted 18 annual on-site inspections of each other's strategic nuclear forces that are provided for under the New START treaty since it entered into in 2011 (as of August 21, 2014, 11 of the 18 have been completed this year by each country). The Open Skies Treaty, signed by 35 nations starting in 2002 and entering into force in 2012, has provided for more than 1,080 overflights as of September 2013, and is ongoing according to the treaty provisions.

Open sources can enhance confidence in assuring compliance through such cooperative activities. This is also true for verifying compliance with existing multilateral treaties banning biological and chemical weapons (BW, CW) by improving the likelihood of detecting and identifying suspicious activities and effluents. Commercial satellites have also provided valuable data for analysis of (non-) compliance with the Nuclear Non-Proliferation Treaty, and open sources proved valuable in detecting the use of chemical weapons in Aleppo, and in subsequent steps to remove such weapons from Syria. They

also informed the world about the Russian troop movements and threats to Ukraine.

Careful case studies of these activities can be of considerable value in designing a well-focused R&D program on technical needs at all levels of sophistication to plan for future systems and strategies.

Open sources are expected to be more effective if they include experts with an understanding of both the scientific and technical issues, as well as the culture of their counterparts. Many such personal contacts that have been made between academicians, scientists at the weapons laboratories, and participants in cooperative agreements such as OST, BW and CW discussions, have been very valuable, going back some 60 years to the early Pugwash meetings.[10]

The Department of State has been successfully working to increase its core strength in science and technology, because of their importance in shaping world events [19]. In addition to enhancing S&T capability within its US facilities, State may strengthen the capabilities of its foreign diplomatic missions in developing open-source information and analysis. We suggest this as part of the electronic outreach (e.g., Facebook and Twitter presence) already in place.

We emphasize that it is not only local capability for use of open sources, but also local sensibilities - political, cultural and otherwise - that should be carefully taken into account for such an effort to be constructive, and therefore both sustainable and ultimately successful. In particular, it is crucial that citizens or groups not be put at risk by encouraging open-source activities that might be interpreted as espionage. The line between open source sensors and "spy gear" is thin.

There is increasing recognition that businesses dealing with sensitive technologies have both an opportunity and a motivation for limiting the pro-

[10] `http://pugwash.org/`

liferation of dangerous technologies, including those constrained by trade and export laws because of potential military applications e.g., Kurzrok and Hund [20]; Maurer and von Engehardt, [21]]. Ralf Wirtz at Oerlikon (manufacturer of components having potential application in uranium enrichment), for example, has emphasized the importance of corporate responsibility in creating a sustainable market [22, 23]. Companies can have considerable situational awareness, both because of their business dealings (proposed as well as actual) and because of their presence in foreign locales. For these reasons, there is an opportunity for the U.S. to develop open-source capabilities by encouraging openness and transparency in the business community.

Finding: Open and crowd-sourced data and analysis can enhance transparency and confidence through cooperative activities with other nations, and through exchanges among technical experts in other areas of mutual interest, such as the environment, climate, and public health.

Recommendation: Identify specific partners and treaty regimes for cooperative monitoring, starting with bilateral agreements but with the idea of building these into multilateral efforts.

Recommendation: Train and task staff in foreign missions to identify sources of open information compatible with their locale.

Recommendation: Recognize, and if possible reward the business community, for transparency practices, including through sharing of open-source information and analysis.

4.6 Public Sharing of Information and Analysis:

"To what extent can collection and analysis be based on publically-available technology and techniques? What are the limitations on public use of the results of this collection and analysis?"

We have illustrated the value of open sources and social media for cooperative verification and monitoring. These have enormous and still rapidly growing capacity to store and process information gained form pervasive monitoring of diverse sensors and sources, and can respond with agility to rapid changes in circumstances (world events) as well as technology. This information can provide valuable triggers to help guide governmental means of identifying potentially harmful activities being pursued, but it is the openness of the data that make them most effective. That is, because numerous and disparate sources of information are being openly shared, mainly via the Internet, the aggregate of open sources and analyses can in many instances be far more capable than efforts based on a few highly sensitive sensors or specialized analysis.

An important reason for making information and, to the degree appropriate, analysis openly available is in order to maximize the chances that these are vetted by independent parties. Open discourse is the essence of the scientific method for this reason and, though the analogy is far from perfect, it is similarly useful to maintain as open a policy as possible for data, analysis, inferences and opinions.

There are two additional benefits to be gained from public sharing of open source information:

1) It can help to build political support for policy decisions or military actions by making more readily available the information on which these decisions are made.

2) It can encourage well-meaning individuals and non-governmental organizations (NGOs), many of whom possess numerous sensors, to pitch in with their observations, reporting, and processing of open source information.

Incentivizing issues arise if one wants to go beyond this and ask citizens to consciously gather information and report on the activities of their own governments, for example in monitoring environmental quality. This could involve effort or personal risk, and even if people are motivated, they might

need appropriate equipment and safeguards (e.g. technology to measure factory emissions or effluents, direction about what activities to report, the ability to send reports anonymously, and so forth).

It may also be useful to facilitate and incentivize data collection by broadcasting information about signs of what to look for in the way of treaty violations or suspicious behavior. For example, there are professional standards in the biological and chemical research communities – and associated industries – regarding work that could contribute toward development or deployment of biological or chemical weapons. There is widespread consensus that such work violates professional standards and is unacceptable, and it is part of advanced training to become aware of these potential dangers. Enhancing this capability for public support of BW and CW treaty regimes is well in line with applications of open source data and analysis.

Along these lines, the ethical, legal and social implications (ELSI) of open-source data have attracted considerable attention, and rightly so [24, 25, 26, 27, 28, 19]. To the degree that information is passively collected from the open Internet, there would seem to be no problem in governmental or nongovernmental organizations aggregating and analyzing the information. However, there is the possibility of retribution, should information considered sensitive to a country's (or its leadership's) security be revealed, even if inadvertently. This possibility is all the more problematic because of the difficulty in maintaining anonymity, even when participants are nominally anonymous [25, 29]. In this regard, Shakil Afridi provides an example of the potential consequences of errors in judgment.[11]

We agree with others' findings and associated recommendations that ELSI is an important matter, and any organized program to collect or analyze open-source data should include a significant effort at addressing these issues. In addition, we emphasize the importance of taking cultural sensitivities into account, both to avoid backlash and also to make most effective use of the information. The recent Facebook scandal is a case in point, with publication

[11] `http://en.wikipedia.org/wiki/Shakil_Afridi`

of Kramer *et al.*'s paper [30] being accompanied by an Editorial Expression of Concern. Regardless of the ethics of that study or its publication, there was considerable displeasure expressed by Facebook clients; notably, this backlash came from many sharing the same cultural norms as the authors.

It is therefore all the more important for US analysts (and scholars) to recognize the potential for unexpected backlash from those living in other cultures. One should recognize up front the potential for tension between US values and the values of others. This is not to say that others' values should be accepted at the expense of US values (e.g., regarding "transparency"), but that the consequences of potential clashes in values be considered as part of a decision to proceed with a given study or not and, if so, how to proceed.

In summary, specific attention has to be paid to cultural norms and context for open source studies. This means that country-specific sensitivities, social and political, as well as real or perceived security risks, be recognized as important for sustaining an open-source data mining process. Staff at US diplomatic missions can play a special role in alerting a program to local sensitivities, as well as to the special opportunities there may be in using social media and open information in a particular cultural and political context.

Finding: Open sharing, as appropriate, improves the reliability and utility of both data and analysis. Though there are technical means of reliably transmitting data and preserving it from corruption or accidental deletion, meaning is more difficult to preserve.

Throughout this report we refer to "appropriate" sharing of information to acknowledge that such sharing must be consistent with legal and ethical standards. We also suggest that, to be effective, sharing should be done in a manner cognizant of social and cultural norms.

Recommendation: Keep open source information and analysis open to the maximum degree possible and appropriate, so as to encourage vetting as well as increased transparency.

Recommendation: Establish a cultural, ethical, legal and societal advisory process that includes participation from business and academia in order to ensure appropriateness and effectiveness of open-source information gathering and analysis.

4.7 Actionability

"Once the integrity and utility of the data are established, how does one determine its actionability? Transformation of data into information is only a first step. The next step is establishing a process for determining the risk associated with acting or failing to act and the feasibility of the various approaches."

Doubts are often expressed that open sources, while indicative, are not actionable. However we have past experience as to how to increase the value of open sources for verifying compliance with international treaties. The situation has much in common with existing international treaty obligations, monitoring and verification under the Comprehensive nuclear Test Ban Treaty (CTBT) and the Chemical Weapons Convention (CWC). Not only are there formal declarations and provisions for inspections, but the very fact of international agreement emboldens citizens to report what they perceive as potential violations, not only to the responsible organizations such as the CTBTO and the Organization for the Prohibition of Chemical Weapons (OPCW), but often to interested Parties such as the United Nations or even the United States. The U.S. invites such disclosures through private communications, and could no doubt do more to encourage such messages in the era of widespread web mail and social media.

Rarely is any source of information free from ambiguity, the possibility of denial and deception or the potential of false alarm. A single, possibly false alarm can have a lasting and disproportionate effect, decreasing the credibility of treaty regimes and undermining confidence between adversaries or even allies.

Determining risks of action verses non-action, in the end, is a policy decision. The same applies to open sources and social media that, while offering the promise of information important for making informed decisions, also bring new risks of deception, ambiguity and false alarm. Therefore as with more traditional approaches for gathering information, all-source analysis of remains the best way to evaluate the actionability of open-source information.

Still, because it can provide unique new insights, potentially from a large number of individuals or sensors in regions of interest, open-source information and analysis have value, even if not fully validated, by contributing to cueing and to interpreting other data. Because it can be widely shared, open-source information generally makes other evidence more actionable.

Finding: Open source data and analysis have value even if not fully validated, by contributing to cueing and interpreting other data. Determining risks of action versus non-action is a political decision once the data have been analyzed, and there is no guarantee against false alarms.

Recommendation: Carry out an end-to-end project on a topic that is not politically charged (even if not treaty-relevant), in order to develop experience. Example: environmental monitoring in cooperation with one or more US agencies and perhaps foreign partners.

5 CONCLUSIONS

Rapid advances in technology have led to the global proliferation of inexpensive, networked sensors that are now providing significant new levels of societal transparency. As a result of the increase in quality, quantity, connectivity and availability of open information and crowd-sourced analysis, the landscape for verifying compliance with international treaties has been greatly broadened. Of comparable importance is the impact more generally on tracking activities potentially threatening to US and international security.

These technologies present both challenges and opportunities for the government to make effective use of the available information and analysis. Agility will be required in adopting the new technologies and exploiting the data. To this end, government should give high priority to 1) tracking public sector activities involving sensor development and data sharing in support of increased transparency; and 2) developing a strategic plan for keeping up with the evolving protocols for collecting, transmitting and analyzing the resulting data. The data are significant in arms control treaty verification and monitoring.

Raw data obtained by these public means may be reliably validated by independent confirmation, supported by checks for internal consistency from quantitative analysis of the clutter (as distinct from errors) in the data. There is an opportunity to enhance transparency and confidence through cooperative activities with other nations, and through exchanges among technical experts in areas of mutual interest such as the environment, climate and public health. We also recommend training and tasking staff in foreign missions to identify appropriate sources of open information, and engaging the business community to share open source information and analysis. Finally, we advocate keeping open source information and analysis open to the greatest degree possible and appropriate so as to encourage vetting as well as increased transparency.

References

[1] R. W. Clayton, T. Heaton, M. Chandy, A. Krause, M. Kohler, J. Bunn, R. Guy, M. Olson, M. Faulkner, M. Cheng, *et al.*, "Community seismic network," *Annals of Geophysics*, vol. 54, no. 6, 2012.

[2] D. J. Milligan, B. Homeijer, and R. Walmsley, "An ultra-low noise mems accelerometer for seismic imaging," in *Sensors, 2011 IEEE*, pp. 1281–1284, IEEE, 2011.

[3] W.-Y. Kim and P. G. Richards, "North Korean nuclear test: Seismic discrimination low yield," *Eos, Transactions American Geophysical Union*, vol. 88, no. 14, pp. 158–161, 2007.

[4] R. Allen, "Seismic hazards: Seconds count.," *Nature*, vol. 502, no. 7469, p. 29, 2013.

[5] S. C. Allen and D. J. Greenslade, "Indices for the objective assessment of tsunami forecast models," *Pure and Applied Geophysics*, vol. 170, no. 9-10, pp. 1601–1620, 2013.

[6] Y. Ishigaki, Y. Matsumoto, R. Ichimiya, and K. Tanaka, "Development of mobile radiation monitoring system utilizing smartphone and its field tests in fukushima," *IEEE Sensors*, no. 13, pp. 3520–3526, 2013.

[7] J. C. Cogliati, K. W. Derr, and J. Wharton, "Using CMOS sensors in a cellphone for gamma detection and classification," *arXiv*, no. 1401.0766v1, 2014.

[8] P. DeBarber and A. Yamamoto, "Smart phone enabled radiation monitor," in *Air Sensors 2014 Conference*, (Research Park, NC), June 2014.

[9] R. E. Gephart, "A short history of Hanford waste generation, storage, and release," Tech. Rep. PNNL-13605, Pacific Northwest National Laboratory, 2003.

[10] G. A. Drukier, J. C. Kessler, Y. B. Rubenstein, and E. P. Rubenstein, "Crowd-sourced calibration of uncontrolled radiation detectors," in *2012 IEEE Conference on Technologies for Homeland Security (HST)*, pp. 373–377, IEEE, 2012.

[11] K. Cheung, L. F. Velasquez-Garcia, and A. I. Akinwande, "Chip-scale quadrupole mass filters for portable mass spectrometry," *Journal of Microelectromechanical Systems*, vol. 19, no. 3, pp. 469–483, 2010.

[12] A. Doerr, "Mass spectrometers on a chip," *Nature Methods*, vol. 6, no. 8, p. 555, 2009.

[13] W. Zeller, L. Naehle, P. Fuchs, F. Gerschuetz, L. Hildebrandt, and J. Koeth, "Dfb lasers between 760 nm and 16 μm for sensing applications," *Sensors*, vol. 10, no. 4, pp. 2492–2510, 2010.

[14] E. L. Holthoff, D. A. Heaps, and P. M. Pellegrino, "Development of a mems-scale photoacoustic chemical sensor using a quantum cascade laser," *Sensors Journal, IEEE*, vol. 10, no. 3, pp. 572–577, 2010.

[15] G. V. Ionov, "The determination of the trajectory of Chelyabinsk bolide according to the records of the drive cams and the simulation of the fragments motion in the atmosphere," *ArXiv e-prints*, Mar. 2014.

[16] J.-Y. Zhu, Y. Lee, and A. A. Efros, "Averageexplorer: Interactive exploration and alignment of visual data collections," *ACM Transactions on Graphics (SIGGRAPH 2014)*, Aug. 2014.

[17] Committee on Applied and Theoretical Statistics, "Strengthening forensic science in the united states: A path forward committee on identifying the needs of the forensic sciences community," 2009.

[18] G. A. Drukier, J. C. Kessler, Y. B. Rubenstein, and E. P. Rubenstein, "Crowd-sourced calibration of uncontrolled radiation detectors," in *2012 IEEE Conference on Technologies for Homeland Security (HST)*, pp. 373–377, IEEE, 2012.

[19] National Research Council, "Science and technology capabilities of the department of state: Letter report." `http://www.nap.edu/catalog.php?record_id=18761`, 2014.

[20] A. Kurzrok and G. Hund, "Beyond compliance: Integrating nonproliferation into corporate sustainability," *Bulletin of the Atomic Scientists*, vol. 69, pp. 31–42, 2013.

[21] S. M. Maurer and S. von Engelhardt, "Industry self-governance: A new way to manage dangerous technologies," *Bulletin of the Atomic Scientists*, vol. 69, pp. 53–62, 2013.

[22] R. Wirtz, "Role and responsibility of the civil sector in managing trade in specialized materials," in *Cultivating Confidence: Verification, Monitoring, and Enforcement for a World Free of Nuclear Weapons* (C. Hinderstein, ed.), pp. 253–281, Stanford, CA: Hoover Institution Press, 2010.

[23] R. Wirtz, "Industry contribution to thwart illicit nuclear trade," Tech. Rep. IAEA-CN-184/198, International Atomic Energy Agency (IAEA), 2011.

[24] B. Lee and M. Zolotova, "New medis solutions in nonproliferation and arms control: Opportunities and challenges," 2013.

[25] C. Stubbs and S. Drell, "Public domain treaty compliance verification in the digital age," *Technology and Society Magazine, IEEE*, vol. 32, no. 4, pp. 57–64, 2013.

[26] B. Lee, J. Lewis, and M. Hanham, "Assessing the potential of societal verification by means of new media," 2014.

[27] Nuclear Threat Initiative, "Redefining societal verification," 2014.

[28] National Research Council, "Emerging and readily available technologies and national security – a frameowrk for addressing ethical, legal, and societal issues," 2014.

[29] A. Narayanan and V. Shmatikov, "De-anonymizing social networks," in *Thirtieth IEEE Symposium on Security and Privacy*, pp. 173–187, IEEE, 2009.

[30] A. D. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Science*, vol. 111, pp. 8788–8790, 2014.