



European
Commission

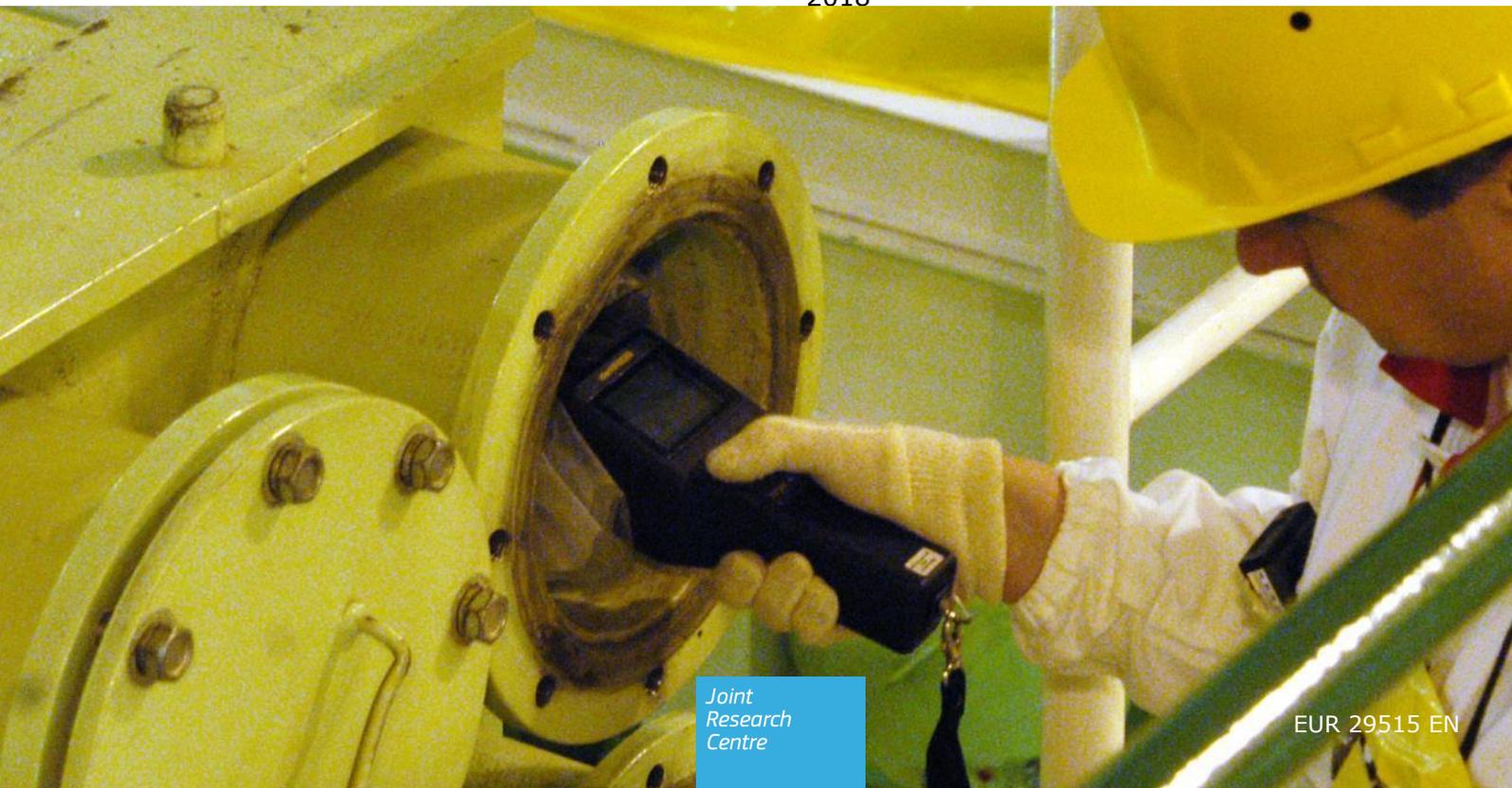
JRC TECHNICAL REPORTS

Open Source Information Analysis in Support to Non-Proliferation

*A Systems Thinking
Approach*

Renda, G.
Cojazzi, G.G.M.

2018



This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact Information

Name: Guido RENDA

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 421, 21027 Ispra (VA), Italy

E-mail: guido.renda@ec.europa.eu

Tel.: +39 033278 3964

JRC Science Hub

<https://ec.europa.eu/jrc>

JRC102286

EUR 29515 EN

PDF ISBN 978-92-79-98286-6 ISSN 1831-9424 doi: 10.2760/87306

Print

Luxembourg: Publications Office of the European Union, 2018

© European Atomic Energy Community, 2018

The reuse policy of the European Commission is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Reuse is authorised, provided the source of the document is acknowledged and its original meaning or message is not distorted. The European Commission shall not be liable for any consequence stemming from the reuse. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

How to cite this report: Renda, G., and Cojazzi, G.G.M., 'Open Source Information Analysis in Support to Non-Proliferation - A Systems Thinking Approach', EUR 29515 EN, Publication Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-98286-6, doi:10.2760/87306, JRC102286

All images © European Atomic Energy Community 2018, except: see images and/or related captions

Table of contents

Acknowledgements	3
Abstract	4
1 Introduction	5
2 Some Preliminary Definitions.....	6
2.1. Defining Open Source Information.....	6
2.2. Defining the System to be Analyzed: a Systems Thinking Approach.....	6
2.3. Defining Open Source Information Analysis	7
3 An Integrated Approach to Monitoring an Engineering Programme.....	8
3.1. Hard Layers' Observables.....	8
3.2. Soft Layers' Observables.....	8
3.3. Taking the Context into Account.....	11
3.4. Uncertainty Identification and Management.....	13
3.5. Putting everything together: a non-nuclear Example	15
4 Types of OS Analysis Scenarios for Nuclear Non-Proliferation	17
4.1. The OS Analysis for Nonproliferation Research Dimensions	17
4.2. Puzzle Solving	17
4.3. Data Foraging.....	18
4.4. Model Building.....	19
4.5. Mystery Framing.....	21
5 Conclusions	23
References	25
List of figures.....	30
List of tables.....	31

Acknowledgements

The reflections here presented stem from the activities performed within the activities *Innovative Concepts and Methodologies for Nuclear Safeguards (INSAF)* and *Open Source Information For Nuclear Security (OSINS)*, later incorporated in the project *Methods, data analysis and knowledge management for Nuclear Non Proliferation, Safeguards & Security (MEDAKNOW)* funded within the European Commission (EC) Euratom Horizon 2020 Research and Training Programme. The authors are grateful to all the researchers that contributed to these projects, In particular Dr. Lance K. Kim, Dr. Rainer Jungwirth, Dr. Simone Cagno, Mr. Frank Pabian and Mr. Erik Wolfart. The discussions with them fertilized the ideas that are here expressed, even though any incorrectness or mistake is the authors' sole responsibility. The views expressed in this report are those of the authors and do not necessarily represent the official views of the European Commission.

Abstract

Open source information, here defined as “publicly available information that anyone can lawfully obtain by request, purchase, or observation” is playing an increasing role in treaty monitoring, compliance verification and control. The increasing availability of data from a growing number of sources on a vast range of topics has the potential to provide cues about complex programmes subject to international treaties such as the *Treaty on the Non-Proliferation of Nuclear Weapons* (NPT).

This report suggests a system’s thinking view of open source analysis in support to non-proliferation analysis, identifying the possible dimensions (hard/soft/context) involved and discussing different types of scenarios an open source analyst might face.

Modelling a nuclear engineering programme by explicitly acknowledging the peculiarities of its hard and soft layers allows the analyst to consider which are the types of insights that each layer can provide and which is the the best tool/technique to investigate them.

Once a particular analysis is set, the analyst might face different types of analysis scenarios, according to the type of the problem to be tackled and the type of data at his disposal. An open source analyst in support of non-proliferation will also have to handle many different forms of uncertainties, whose proper understanding is critical for the analyst to perform dependable assessments and for the decision maker to take properly informed decisions.

A system’s thinking approach to open source analysis has the potential to integrate synergically with the other tools available in the international treaty monitoring toolkit, helping in increasing the international community’s confidence in its ability to detect an undeclared proliferation programme.

1 Introduction

Nuclear energy programmes are complex, long-term endeavours, involving multiple sites and a substantial number of personnel with a wide range of expertise. States willing to embark in such an enterprise would have to either sign the *Treaty on the Non-Proliferation of Nuclear Weapons* (NPT) and benefit from the transfer of technology enabled by its membership or develop the necessary science, technology and industrial capability autonomously. States signatories of the NPT renounce the possibility to develop a nuclear military programme and are subject to international nuclear safeguards implemented by the *International Atomic Energy Agency* (IAEA).

Within its State Level Concept [1], the IAEA envisions an objective-based and information-driven approach for designing and implementing State Level Approaches (SLAs). To achieve its objectives the Agency foresees the adoption of a holistic approach, in which all the information gathered is analyzed and assessed as a whole “. . . to draw and maintain a conclusion of the absence of undeclared nuclear material and activities in that State.” ([2], p. 20)

While the Agency, in its assessment, performs an *all-source analysis* making use of all the information and knowledge available to it¹ [3], either closed-source (official State’s declarations, inspection data, third parties information) or open source, other treaty monitoring regimes often rely on fewer potential sources. No matter the domain in which open source analysis is employed to inform international treaties monitoring, the IAEA intuition about the need for holism to be effective needs to be seriously taken into account.

This report investigates the potential of open source analysis in support to non-proliferation adopting a systems thinking approach. The starting point of the investigation is the acknowledgement that every engineering project implies the complex interaction between “hard” technical systems and infrastructures and “soft”, social systems (the people and organizations that operate and interact with the technical systems), within a given geopolitical and social context. Each of these three layers potentially generates observables that might be captured by open source monitoring. Although these observables—taken by themselves—are usually too weakly correlated to a nuclear programme to be relevant, when linked to other week observables they might form an overall picture compatible with the existence of such a programme.

The reader will be guided through an overview of *what* is open source analysis (section 2) when taking a holistic perspective and *how* it could be used to detect and monitor a complex engineering programme as a nuclear energy programme, here described adopting a *systems thinking, holistic* [4] point of view (section 3). In the process, some examples on how this could be applied to nuclear non proliferation are provided. Finally, section 4 proposes some reflections on possible types of open source analysis scenarios that can be encountered when analysing complex programmes for treaty monitoring purposes.

¹ In this respect, although they do collect and make use of open source information, the IAEA analysts are here considered not to be open source analysts.

2 Some Preliminary Definitions

Objectives for this section:

- To provide a definition of Open Source Information Analysis.
- To define the system to be analyzed according to the *Doing It Differently*(DID) Systems Thinking approach.

2.1 Defining Open Source Information

There is no normative definition of open source information that is universally accepted, as there is no universal definition for open source analysis. For this exposition, open source information will be defined as “publicly available material that anyone can lawfully obtain by request, purchase, or observation.” ([5], p. 5) This definition is very broad, and covers a very vast range of potential sources of information. For instance, the IAEA sets it to include “information generally available from external sources, such as scientific literature, official information, information issued by public organizations, commercial companies and the news media, and commercial satellite imagery ” ([6], p.10) and trade analysis [7,8].

While classified information is certainly outside the open source information envelope, the exact boundaries of this envelope are not well defined: in some cases e.g. grey literature is considered to be open source [9], even if technically not reachable by everybody.

2.2 Defining the System to be Analyzed: a Systems Thinking Approach

There are many ways of formalizing a complex engineering programme, investigated by the Systems Engineering discipline.² Depending on the final goal one might work better than another. This discussion will propose a formalization of a complex system adapted from the “Doing it Differently” (DID) approach [14], originally conceived to bring a holistic approach to the construction industry and then explored in several other fields, including that of nuclear proliferation resistance of nuclear energy systems³ [15] and nuclear safeguards [16].

A central idea in the DID approach is the fact that complex issues involve many different players with diverse needs, wants and objectives which must be taken into due account. In the DID approach everything is modelled as a *process*.

A process is defined by setting out the need for its existence and how it develops through time towards its objectives. The first process is that of *being*. The very fact of existing through time involves changes that might have considerable impact on the eventual success of the project.

Every process involves an interaction between “*hard*”, technical systems (which might be very basic equipment/products to be used or very complex technical systems to be operated) and “*soft*”, social systems (the behaviour of people who play various roles in the

² As Dori and Sillito pointed out [10], the concept *system* is used in science since the times of Aristotle, who affirmed that “the whole is something over and above its parts, and not just the sum of them all.” ([11], quoted in [10], p. 209). Despite this, the definition of *system* in modern science is still not universally codified. [10] Similarly, also modern Systems Engineering doesn’t have a universally codified definition. Here Systems Engineering will be defined as “. . . a discipline that concentrates on the design and application of the whole (system) as distinct from the parts. It involves looking at a problem in its entirety, taking into account all the facets and all the variables and relating the social to the technical aspect.” ([12], quoted in [13])

³ The presentation of the DID approach outlined here is reproduced and adapted from [15], Appendix I.

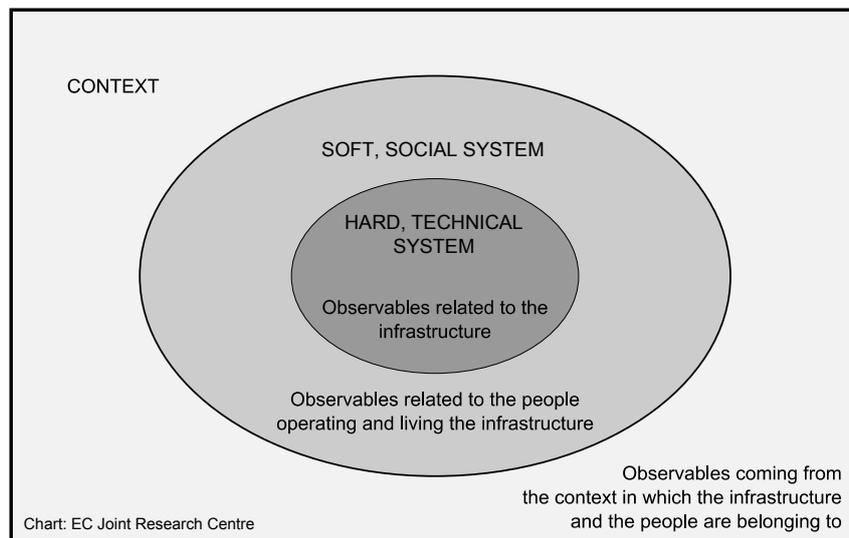


Figure 1: System's hard, soft and context layers adapted from [14].

process and who will interact with the hard system). All of this is set within a given *context* [17] (Fig. 1).

Each process is considered to be a *holon*, i.e. both a whole and a part [4]. Every process is a whole made up of the *hard* and *soft* layers that interact to produce change, but it is also a part of other processes to which it contributes. For example a nuclear reactor is a whole made up of all the primary and subsidiary technical systems operated by the personnel of the site and according to the regulatory framework (process "operating a nuclear power plant"). The reactor is also a part of the overall nuclear fuel cycle in which it is embedded (process "operating a civilian nuclear fuel cycle") [15].

The complexity of a programme is described as a process view where the traditional separation of the *hard* technical layer from the *soft* social one and the *context* in which they take place is replaced by a systemic comprehensive view of all aspects, where the processes of a complex system are organised into a hierarchical structure with increasing levels of definition.

2.3 Defining Open Source Information Analysis

Within a systems thinking approach, open source analysis could be seen as a process of "getting the right information (what) to the right people (who) at the right time (when) for the right purpose (why) in the right forum (where) and in the right way (how)" ([17], p. 190) by merging openly available data and information coming from a wide variety of accessible sources into an overall comprehensive and cohesive picture. Usually the process involves the gathering and analysis of a large amount of data and information, a very small percentage of it being relevant. A common scenario involves filtering an enormous amount of data to end up with a sparse and incomplete set of information not all of which contributes to knowledge [18]. When investigating a covert engineering programme, the analyst will have to deal with low quality data and is always exposed to deliberate deception [19,20]. Nonetheless, the analyst may be able to obtain valuable insights about what a State might be pursuing.

3 An Integrated Approach to Monitoring an Engineering Programme

Objectives for this section:

- To introduce Hard and Soft Layers' observables and the insights they might provide to the analyst.
- To reflect on the importance of the context in which a system operates.
- To reflect on the uncertainties an analyst needs to cope with when dealing with open source information in support to non-proliferation.
- To illustrate, via a non-nuclear example, how the observables from all the existing layers can be merged into a consistent and informative picture.

3.1 Hard Layers' Observables

The *hard* layer of an engineering programme refers to its technical infrastructure. The observables typically relate to the existence and functioning of the technical infrastructure of the system. They provide information about the topology, layout, shape and sizes of the infrastructure. When peculiar shapes are recognizable (e.g. the reactor dome of a nuclear power plant), they give the analyst the opportunity to gain insights about the technical layer's function.

Usually these observables are continuously available (e.g. the visual signature of a building is always available for collection), but—at least in the field of open source analysis—signals collection is discontinuous to varying degrees (e.g. the visual signature of a building will be captured only during an overhead passage of a imagery commercial satellite).

Since the technical infrastructure is generally geo-localized, the data gathered is mostly *structured information*, and can be processed and analyzed with geospatial tools.

Examples of open source information related to hard layers' observables are obtained through satellite/overhead imagery and remote sensing in general [21]. These techniques already demonstrated their effectiveness in analyzing and monitoring nuclear-related sites to gain knowledge of their status and activities, both in the field of nuclear safeguards and non-proliferation. For instance, the interest in monitoring proliferation sensitive fuel cycle steps such as irradiation in reactors and enrichment via remote sensing is not new [22,23] and e.g. the nuclear-related sites in Iran have been extensively monitored in the last years by NGOs active in the field of nuclear non-proliferation (see e.g. [24]). The DPRK nuclear test site has been object of several geospatial analyses, published both in scientific journals (see e.g. [25,26]) and non-proliferation blogs (see e.g. [27,28]). Open source analysis has also investigated sites that underwent nuclear accidents [29,30]. Fig. 2 provides an example of open source analysis of technical infrastructure through the use of remote sensors and satellite imagery as presented in [21].

3.2 Soft Layers' Observables

The *soft* layer of an engineering programme relates to the people that conceive, design, build and operate the *hard* layer under investigation. This layer gives the possibility to look into the social aspect of the programme under analysis, and gives the analyst the possibility to gain insights about the function of the various parts of the infrastructure, the people working in and operating it, the working security rules and policies of the system, together with hints about the context to which the process belongs.

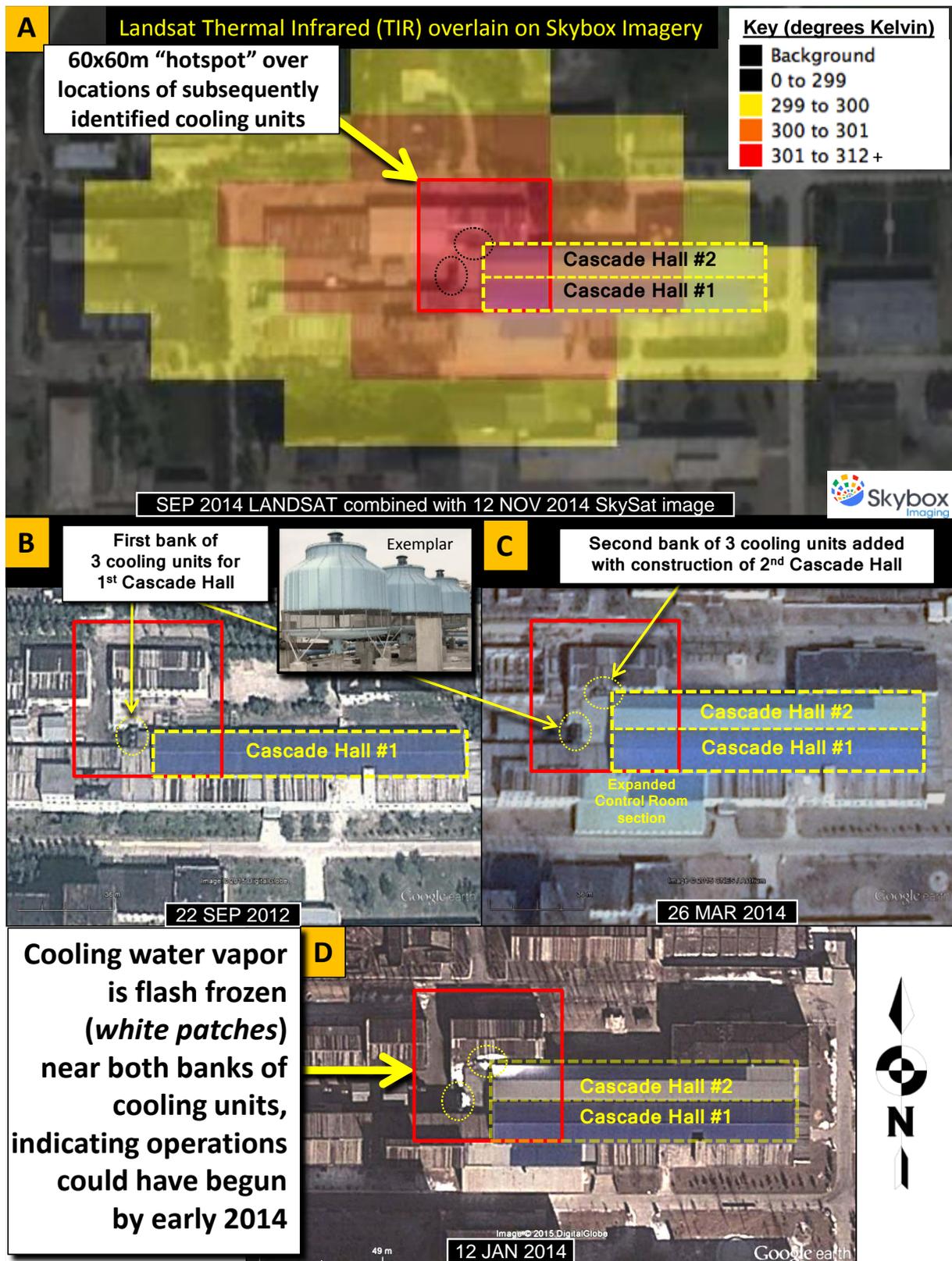


Figure 2: "Multi-satellite, multi-sensor [analysis] for verification applications. The figure shows the North Korean gas centrifuge uranium enrichment plant at Yongbyon. The original cascade hall was visited by Dr. Siegfried Hecker in 2010 [31]. In 2013, what appeared to be a duplicate cascade hall and probable expanded control room were constructed. The low resolution LANDSAT thermal imagery directed [the] analytical focus to an area where cooling towers were identified with higher (finer) resolution optical imagery. While the lower (coarser) resolution thermal imagery could not distinguish whether one or both towers were operational in September 2014, the January 2014 DigitalGlobe finer resolution imagery (as is viewable with Google Earth) suggests that both were possibly operating as early as January 2014." ([21])

Signal emitters of a soft layer include the people themselves. In addition to the scientific literature and “traditional” media sources, the advent of the internet and social networks [32, 33], coupled with the technological advancements in the field of mobile connectivity, led to the fact that nowadays many people own and regularly use internet-enabled camera phones. The penetration of smartphones worldwide grew “globally from 15% in 2012 to 24% in 2014” ([34], p. 1) and in 2016 31.2% of the world population is foreseen to “own at least one smartphone and use the smartphone(s) at least once per month.” ([35]) Evidence suggests that in remote rural areas, if mobile connectivity is available, their adoption rate is not far from the one experienced in urban areas [34].

Where people have access to reliable connectivity as in the European Union and in the United States of America, a high percentage of them regularly uses one or more social networks. In the US in 2014 “52% of online adults [used] two or more social media sites.” ([36], p. 2) Having a cameraphone in one’s pocket increases the likelihood of having people sharing pictures of the places and events they frequent, to the extent that in 2014 the photo sharing social network flickr.com saw an average number of photos shared per day of 1 million⁴ and in 2015 the average daily images upload on Instagram was in the range of 70 million.

Compared to the observables for the *hard* layer described in section 3.1, *soft* layer observables are discontinuous in nature (the single person publishes or posts only once in a while), but the enormous number of potential emitters still provides a remarkable signal—albeit most of it is, from a Treaty monitoring perspective, just noise. Contrary to the *hard* layer’s observables, some of the *soft* ones enable potential *continuous collection* to the open source analyst: posts on social networks such as Twitter or Instagram can be monitored near-real time, and the same is technically possible with (micro-)blogs and online media sources [37].

The information collected can either be *geo-referenced* (as in the case of some social network posts) or *unstructured* (as in the case of scientific and official literature, online blogs, online and traditional media).

When social networks data are analysed, the analyst is potentially able to derive information from much more than just the actual content shared by the user. In particular, the following aspects of social networks’ posts can be of potential interest:

- *Layout* patterns: when geo-referenced, the pattern of the posts (i.e. from where the information was being posted) can reveal useful information about e.g. the security rules of a nuclear site. When analysed over a long enough period of time, the pattern would reveal e.g. whether the site allows smartphones use or not, and in case restrictions are limited to some areas and buildings, which are the buildings interested to this limitation. Usually civilian nuclear research sites are very active, but military nuclear sites and research centres are virtually silent. Fig. 3 shows two different nuclear research centres and their respective integral social network signals layout over a given period of time. It is evident how the two patterns are very different.
- *Temporal* patterns: temporal lines of posts from one source can provide insights of the system’s evolution (people and related activities being moved from one building to another, dismissal of activities); concentration of posts from a given building in a given period of the day (e.g. 12H00–14H00) can provide insights of the function of a given building (e.g. a building active between 12H00–14H00 is likely the canteen/cafeteria), providing hints of its function and level of sensitivity. Sharp increases or decreases at certain periods of the day from the entire site might reveal the working habits, such as usual working hours. Fig. 4 shows the use of social networks posts’ temporal line to acknowledge a given subject’s relocation in a given date.
- *Content*, although most of the times irrelevant, can occasionally provide insights on the type of activities and equipments involved e.g. in laboratories, or whether renovation and refurbishment of the equipment is ongoing. Fig. 5 shows selected Social Network posts uploaded from a nuclear research centre’s experimental facility.

⁴ <http://tinyurl.com/techcrunch-flickr>, accessed on September, 5, 2018.



Figure 3: Social Network integral signal layout on two different nuclear research centres. The one on the left is a civilian nuclear research centre, the one on the right is a nuclear research centre related to the design of nuclear warheads.

Monitoring news and social media allows the open source analyst to follow the development and the evolution of the system under analysis in an almost dynamic way, resulting in being potentially very useful in filling the gap between the availability of signals collected from the *hard* layer. In particular, the *structured* information coming from social networks posts and online collaboration efforts can allow the follow-up and analysis of events such as radiological releases [38, 39], accidents, riots and protests [40].

While it might be an issue for any type of open source [41], the *soft* layer's observables, especially those related to social network posts, face the risks of being unreliable, inaccurate, or even subject to manipulation to misinform and deceive [20]. There have been cases of States having allegedly set up "troll farms" as a part of a disinformation operation [42]. The analyst should therefore be aware of this possibility and look for possible corroboration deriving from other independent sources.

3.3 Taking the Context into Account

Any human activity is shaped by and carried out within a particular *context*, the knowledge of which might be extremely important for the understanding of the observed process.

The information about the *context* relates to the bigger picture of which the programme under investigation is part. It can inform on both the *hard* and *soft* layers of the programme, and is usually colored by biases and points of view.

Typically, the information related to the context is unstructured in nature, and comes from a wide variety of sources, including news and media reports, international relationships and obligations, political and government statements and acts.

The start of complex programmes such as a nuclear proliferation programme is often influenced by the context perceived by the State initiating it, and influences how the programme could be structured and carried out. For instance "two programs, those of the United States and Soviet Union, were unquestionably initiated in response to what were perceived as di-



Figure 4: Example of the analysis of Social Network posts temporal line. In this example before a given date a given target was an active Social Network poster from a particular house. After the given date the posts of the same target resulted to originate from another house. In that period the target likely relocated from the first house to the second.

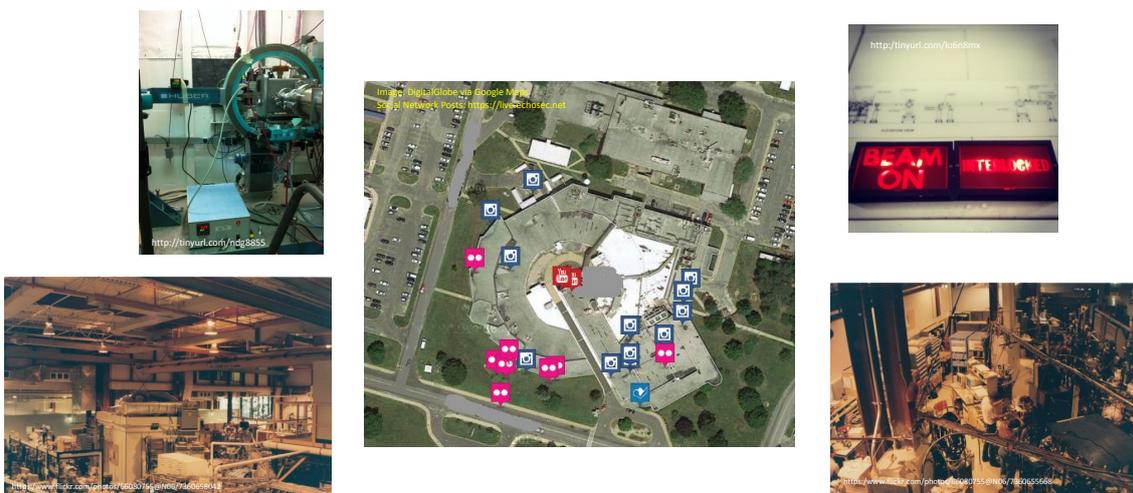


Figure 5: Example of Social Network posts content analysis. The figure shows selected Social Network posts uploaded from a nuclear research centre's experimental facility. As it is possible to see, the material posted might help in gaining a flavour of the inside of the experimental facility.

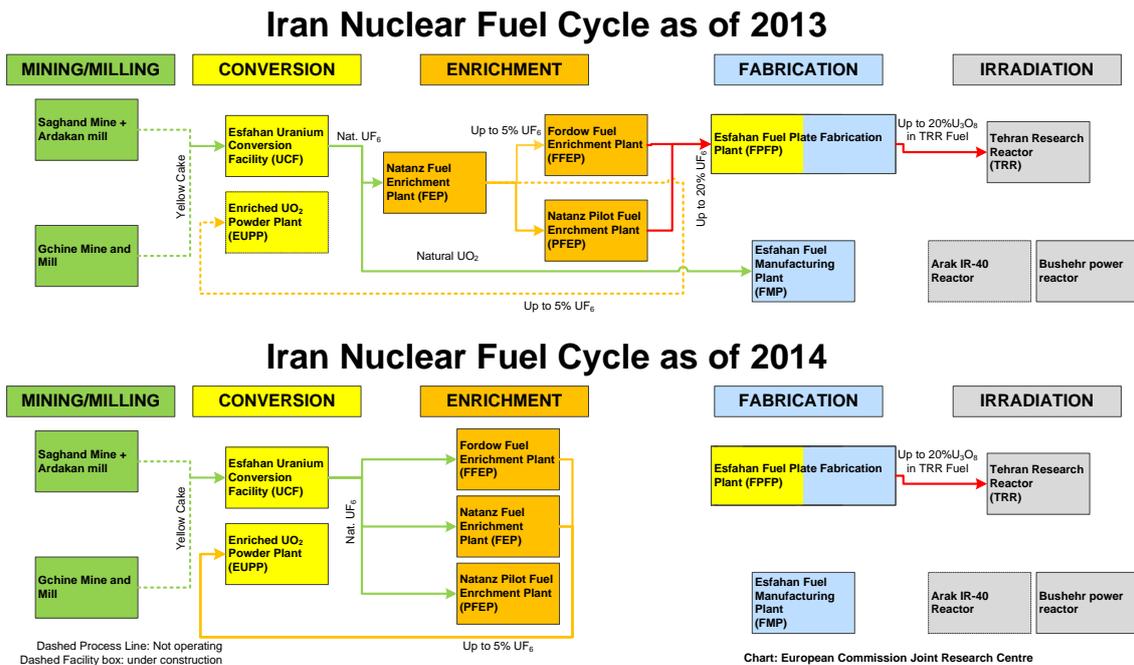


Figure 6: Snapshot of the Iranian nuclear fuel cycle in 2013 and in 2014 as emerging from the IAEA Director General reports to the IAEA Board of Governors [46, 47]. The evolution of the programme would not be understandable without the knowledge of the geo-political context (signature of the EU/E3+3 and Iran Joint Plan of Action related voluntary measures [48] foreseeing changes in the scope and operation of some Iranian nuclear fuel cycle's facilities).

rect outside threats and, in the rush to develop weapons, both the enriched uranium and plutonium routes were pursued simultaneously." ([43], p. 11)

Although there is no definitive answer to why states decide to start a nuclear weapons programme and how influential the context may be [44, 45], sometimes knowledge of the external influencing factors is the only way to explain the observed evolution of a nuclear programme. An example of how the context might be critical to understand the evolution of a nuclear programme is provided in Fig. 6.

The figure shows two snapshots of the Iranian nuclear fuel cycle: one in 2013 [46] and another in 2014 [47]. While in 2013 the fuel cycle covered all the steps of an open fuel cycle and presented a coherent picture from a process point of view, one year later the same fuel cycle apparently evolved into a much less understandable picture. The explanation for this evolution is the agreement between Iran and the EU/E3+3 (France, Germany, United Kingdom, China, the Russian Federation, the United States of America, the European Union) on a Joint Plan of Action [48] that shaped its nuclear fuel cycle as depicted.

3.4 Uncertainty Identification and Management

Each time an assessment that needs some kind of decision making is carried out, the aspect of uncertainty management plays an important role in ensuring that the analyst could convey to the decision maker the most complete and dependable message. An Open Source analysis in support of non-proliferation will have to handle many different forms of uncertainties, whose proper understanding is critical for the analyst to perform dependable assessments and for the decision maker to take properly informed decisions.

In general terms uncertainty could be categorized in two broad aspects: *aleatory* and *epis-*

*temic*⁵ [50]. Aleatory uncertainty relates to the intrinsic lack of specific pattern in the data under investigation, and is commonly referred to as randomness. It is the domain of classical probability theory and is intrinsic of the phenomenon under observation, i.e. this source of uncertainty cannot be eliminated. Epistemic uncertainty relates to the imperfect knowledge of the subject/data by the analyst, and could in theory—but seldom in practice—be eliminated. In any kind of analysis these two aspects are usually present and distinguishable, but intertwined [51].

Gathering evidence for the presence of indicators of a nuclear engineering programme in a state, especially when related to military aspects and therefore with active efforts to keep it concealed, means having to deal with several epistemological issues:

- The analyst is out to detect something whose existence is uncertain.⁶ Making a parallelism with any classic measurement performed by an inspector on declared nuclear material in a declared facility, [54, 55] while the three postulates of the theory of measurement [56] inform the latter that he will not ever know the absolute true value of the characteristics he is measuring, they also inform him that the true value does exist and it is—within the limit of the typical safeguards measurement campaigns—de facto constant. When searching for indicators of a possible clandestine nuclear programme, the analyst does not know whether such programme really exists.
- Assuming that the clandestine programme exists, the analyst does not know its characteristics,⁷ and therefore would not be in the position to choose the best detection method for finding indicators of its presence. This not only impacts effectiveness and efficiency, but also adds considerable epistemic uncertainty about the outcome of the verification activities.
- In addition to the above-mentioned issue, even in the case in which the analyst chooses the appropriate method for detecting the existence of a given indicator, the actual detection probability of the presence of an indicator given its existence when using a given detection/measurement technique is often not known as no dependable attempts to investigate this aspect are available.

In view of the above aspects, it is here necessary to introduce a further distinction related to the epistemic uncertainty, that will be here further characterized in *fuzziness* and *incompleteness*. [14, 57, 58] The first one (fuzziness, or imprecision of definition) pertains to the inability to perfectly describe and characterize an aspect that the analyst knows being relevant to the analysis. The second one (incompleteness) acknowledges the fact that the analyst deals with incomplete information, including both “known unknowns” and “unknown unknowns”. This calls for an *open world* assumption⁸ while dealing with uncertainty management and its formal treatment.⁹ This requires the relaxation of the additivity of probabilities requirement¹⁰. [15] provides an overview of a possible way to account for all the above types of uncertainties via the use of an adapted version of Interval Probability Theory [57].

⁵ For additional possible distinctions and their interpretation see e.g. [49].

⁶ Actually the history of non-proliferation tells us that it is safe to assume that most of the times there is *no* clandestine nuclear activity going on in a State. For an overview of past known proliferation efforts, see e.g. [43, 52, 53]

⁷ Among all the possible variables, there are two main routes to nuclear proliferation: one foresees the acquisition of weapons-usable plutonium as the main nuclear material for a nuclear weapons-arsenal and the other one foresees the acquisition of highly-enriched uranium. The two routes imply different technologies and therefore would produce different indicators. For an overview of the choices made by proliferators regarding these two options, see e.g. [43].

⁸ “The Open World Assumption (OWA) is the assumption that that what is not known to be true or false might be true, or absence of information is interpreted as unknown information, not as negative information. It assumes incomplete information about a given state of affairs, i.e., there may be more relevant information than what is provided.” ([59])

⁹ Examples of uncertainty management approaches allowing an open world assumptions are e.g. [57, 60, 61].

¹⁰ The non-additivity of probabilities is not a recent concept: Bernoulli already accepted it in the XVII Century when he tried to understand how could the force of different pieces of evidence be combined [62].

3.5 Putting everything together: a non-nuclear Example

As an example of synergy of *soft* and *hard* layers' observables, Fig. 7 shows an effort to locate the Ebola cemetery opened by the Liberian authorities out of Monrovia at the beginning of 2015, based on the open source information available at that time.

In 2014 and 2015 Liberia experienced a dramatic outbreak of the Ebola virus, requiring a coordinated international intervention in support to the Liberian authorities in coping with this regional emergency. At the beginning of 2015, under the need to provide an alternative to cremation for the Ebola virus victims, a new safe burial site was opened outside Monrovia.

Following the description of a blog post on openstreetmap and Google Earth, in March 2015 it was possible to identify a candidate area, but not to corroborate it as the last freely available high resolution satellite image dated back to 2013, well before the establishment of the cemetery. Low resolution images from DigitalGlobe ImageFinder confirmed the presence of new activities after the date of establishment of the cemetery but were not sufficient to link the activities to the cemetery.

Investigating social network posts, some images on the UNMEER official Flickr account were found , dated January 2015, related to the burial site (see Fig. 7). Despite the credibility of the source, the geographical coordinates attached to the images did not seem to be consistent neither with the very accurate driving directions given by the blog post nor with the information that could be derived by low-resolution images via DigitalGlobe ImageFinder¹¹. Images uploaded on social networks taken around Decoration Day (a Liberian National Holiday since 1916 in which the people remember their dead relatives i.a. by decorating their graves) in the area under investigation corroborated the inaccuracy of the UNMEER images' geotags and supported the candidate site identification. Another image uploaded in July 2015 further increased the confidence in the correct identification. Finally, in early 2016 a new high resolution satellite image of the area was made available in Google Earth, which confirmed the candidate site as the actual burial site.

This example shows how the use of social networks, collaborative tools and online blogs can work synergistic with hard layers' observables such as high resolution satellite images to inform the open source analyst about a particular site/system.

Soft layers' observables are also potentially able to fill the gap between the last available satellite image and a given point in time, allowing the analyst to monitor the evolution of a system through time in a more dependable way.

Knowledge of the context (existence and meaning of Decoration Day in the Liberian tradition) helped the analyst in focusing his attention on selected sources of information around a precise point in time, allowing him to retrieve important information that proved valuable to reach the investigation's goal.

The Example also highlighted how the use of open source information, even when coming from reputable and credible sources, requires a careful vetting of its correctness before being able to depend on it.

¹¹ The images taken after the opening of the burial site did not show any significant activity in the area corresponding to the images geographical coordinates.

ReliefWeb report <http://reliefweb.int/report/liberia/safe-burial-ebola-victim-liberia>

Liberia: A safe burial for an Ebola victim in Liberia

Original published date: 10 Feb 2015
 Country: Liberia
 Theme: Health, Water Sanitation Hygiene
 Content format: News and Press Release
 Language: English
 Source: UN Children's Fund
 Disaster type: Epidemic
 Origin notes: http://www.unicef.org/infoycountry/liberia_79760.html?utm_source=unicef_news&utm_medium=rss&utm_campaign=rss_link

By Helene Sandbu Ryeng

MONROVIA, Liberia, 9 February 2015 – “Drive along the Robertsfield highway past Tower Hill, make a sharp right turn after you pass Ambush Curve, and proceed a few hundred metres until you reach Disco Hill. It’s close to smell no taste.”

These are the directions to Liberia’s newest cemetery, one in which more than 350 people have been buried in a dignified – and safe – way.

Sources: maps: OpenStreetMap
 Non-matching geo-tagged pics: UNMEER flickr account
 Matching geo-tagged pics: Instagram: @iamjimuttile (worker), @cariyell (graves)
 Jan 2013 satellite images: Google Earth
 Mar 2015 satellite image: DigitalGlobe ImageFinder

Arm Bush Curve = Ambush Curve
 (i.e. typo in OSM?)
 Arm Bush Curve

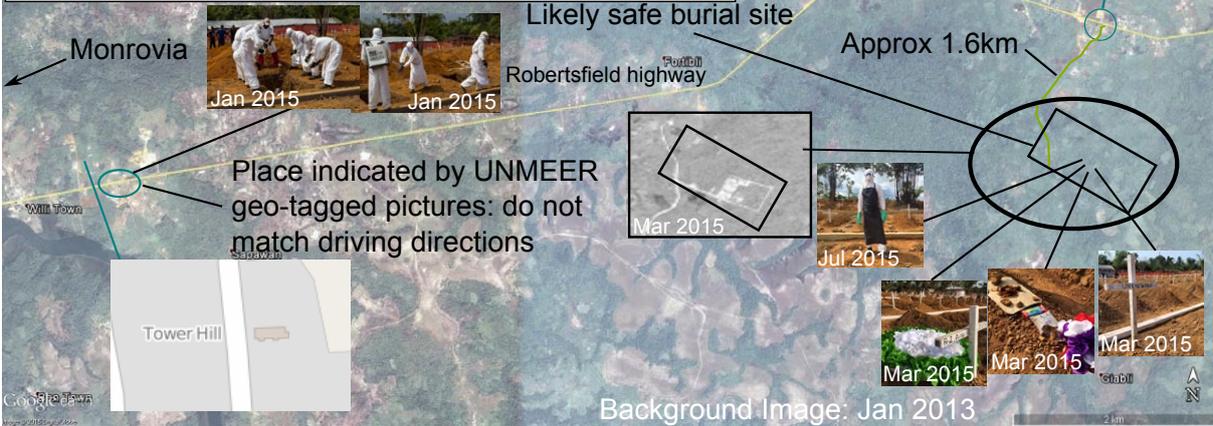


Figure 7: Effort to use social media information to locate the ebola cemetery opened by the Liberian authorities out of Monrovia at the beginning of 2015, being the last high-resolution satellite image of the area available for free dating back to 2013. The bottom picture shows an August 2015 Satellite image of the identified area, confirming the identification. Credits as in the figure, annotations: EC JRC.

4 Types of OS Analysis Scenarios for Nuclear Non-Proliferation

Objectives for this section:

- To introduce the possible types of OS analyses that could be performed to support non-proliferation regimes.
- To provide examples of such analyses.

4.1 The OS Analysis for Nonproliferation Research Dimensions

As with any type of analysis, an Open Source analysis in support to nuclear non-proliferation comprises a set of data to be analysed and a paradigm¹² to make sense of those data. Depending on the quality and quantity of the data that are potentially available and the quality of the models¹³ at the analyst's disposal, four different analysis research dimensions can be identified [65], illustrated in Table 1. To be able to effectively perform OS analysis in support to nuclear non-proliferation, these scenarios are to be understood and assessed.

Table 1: Research dimensions for OS analysis applied to nuclear non-proliferation. General Frame adapted from [65].

		Paradigm Strength	
		Low	High
Data Quality/Quantity	High	Model Building What could work for NP?	Puzzle Solving How likely in NP?
	Low	Mystery Framing How to find "unknown unknowns"?	Data Foraging What and how in NP?

The following paragraphs will briefly discuss the four categories of Table 1. The concept is illustrated in the following paragraphs by means of examples.

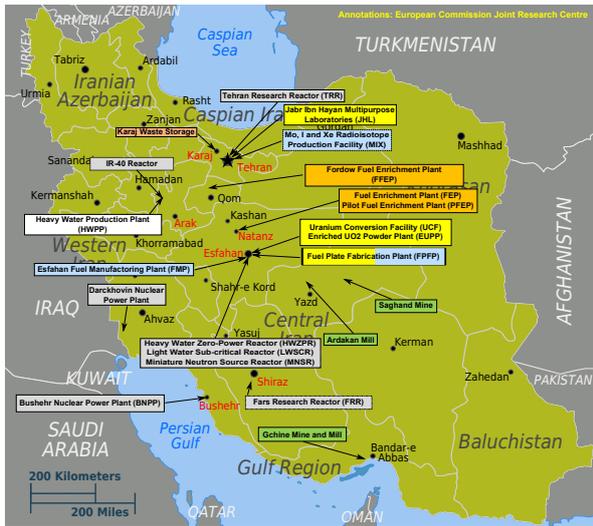
It has to be noted that the classification of the analysis as being part of one of the four scenario types should not be considered to be rigid and depends on various factors, including the analysis and the analyst: an analysis could fall into e.g. the *data foraging* category for one analyst and into e.g. the *model building* one for another. In addition, an analysis might start out as a mystery framing problem, which after new inputs can evolve into a data foraging or a model building one, and (rarely) ending up in a puzzle solving exercise. Section 4.5 provides an example of such migration from mystery framing through model building and data foraging.

4.2 Puzzle Solving

When high quality data and a well established paradigm with dependable models are available to the analyst, the outcome of the analysis is usually dependable and the analyst's role is that of gathering the data, processing them according to the need and visualizing the outcome in the most suitable way. An example of puzzle solving analysis, in addition to the description of the Iranian production nuclear fuel cycle depicted in Fig. 6, is shown in Fig. 8. The information used comes mainly from the IAEA Director General reports to the IAEA Board of Governors. The nuclear fuel cycle production processes are very well known in their general terms and make for a high quality model for the purpose of the analysis. The outcome can be considered as a dependable picture of the status of the Iranian nuclear fuel cycle's *declared* processes.

¹² Following Kuhn, a *paradigm* is here defined as "universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners." ([63], p. viii)

¹³ A model is here defined as "a physical, mathematical or logical representation of a system entity, phenomenon, or process." ([64], p. 105)



- Tehran:**
1. Tehran Research Reactor (TRR)
 2. Molybdenum, Iodine and Xenon Radioisotope Production (MIX) Facility
 3. Jabr Ibn Hayan Multipurpose Laboratories (JHL)
- Esfahan:**
4. Miniature Neutron Source Reactor (MNSR)
 5. Light Water Sub-Critical Reactor (LWSCR)
 6. Heavy Water Zero Power Reactor (HWZPR)
 7. Uranium Conversion Facility (UCF)
 8. Fuel Manufacturing Plant (FMP)
 9. Fuel Plate Fabrication Plant (FPFP)
 10. Enriched UO₂ Powder Plant (EUPP)
- Natanz:**
11. Fuel Enrichment Plant (FEP)
 12. Pilot Fuel Enrichment Plant (PFEP)
- Fordow:**
13. Fordow Fuel Enrichment Plant (FFEP)
- Arak:**
14. Iran Nuclear Research Reactor (IR-40 Reactor)
- Karaj:**
15. Karaj Waste Storage
- Bushehr:**
16. Bushehr Nuclear Power Plant (BNPP)
- Darkhovin:**
17. 360 MW Nuclear Power Plant
- Shiraz:**
18. 10 MW Fars Research Reactor (FRR)

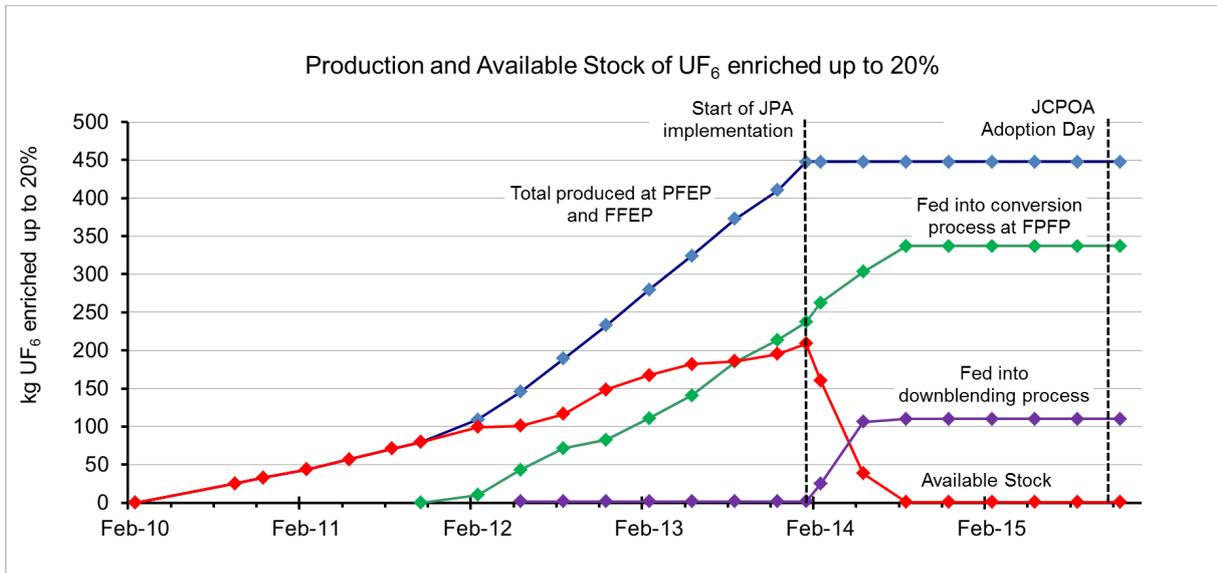


Figure 8: Example of puzzle-solving OS analysis. The figure shows the name and location of the declared Iranian nuclear facilities. The chart shows the overall evolution of the stock of UF₆ enriched up to 20 % produced in Iran between February 2010 and November 2015. The period includes the implementation of the EU/E3+3 and Iran Joint Plan of Action related voluntary measures [48] until the adoption of the EU/E3+3 and Iran Joint Comprehensive Plan of Action (JCPOA) [66] in October 2015. Data taken from the IAEA Director General reports to the IAEA Board of Governors.

While the Iran nuclear fuel cycle is a good example of puzzle solving, it is also a very particular one. Usually the amount of information available on a nuclear fuel cycle is much more limited and attributable to sources that are much less reliable than the IAEA. Thus, it is not expected to see many puzzle solving problems in performing OS analysis in support to nuclear non-proliferation.

4.3 Data Foraging

A more likely analysis scenario is one in which the paradigm is well established but dependable information is lacking. In this case, the main task and challenge for the OS analyst is to find and collect data as dependably as possible, and the dependability of the analysis will be mainly driven by the quality of the gathered data. This scenario also opens a research dimension, i.e. to look into innovative ways of finding and collecting data that may not have been generated/intended for non-proliferation analyses but can adequately cope with the task.

Name	Separative Power (kg SWU / centrifuge / y)	Overall Height (cm)	Length (casing) (cm)	Diameter casing (cm)	Rotor material	Rotor diameter (cm)	Rotor length (cm)	Number rotors	Bellow material	Number bellows	Type
IR-1	0.9-2.1	185	170	16	Aluminum	10	35	4	Maraging steel	3	Supercritical
IR-2		100	80	22	Carbon Fibre	13	47	1		0	Subcritical
IR-2m	2.7-6.3	130	123	22	Carbon Fibre	14	54	2	Maraging steel	1	Supercritical
IR-3		102	73	22	Carbon Fibre	13	47	1		0	Subcritical
IR-4	2.7-6.3	141	123	22	Carbon Fibre	14	54	2	Carbon Fibre	1	Supercritical
IR-5											
IR-6	5.4-12.6	128	115	29	Carbon Fibre			2		1	Supercritical
IR-6s											
IR-8	13.5-31.5, 24										

Data based on IAEA reports, measurements or Iranian statements

Assessment

Derived numbers

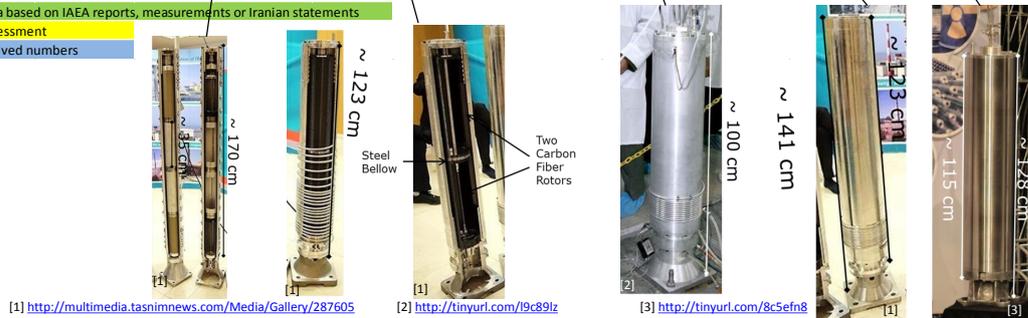


Figure 9: Example of data foraging OS analysis: Iranian centrifuges. Despite the exceptional amount of open information available about the Iranian nuclear programme, in 2014 little was known about the technical details of the various types of centrifuges developed in Iran. The figure shows an attempt to assemble and derive this information from publicly available sources in 2014, making use i.a. of image analysis to derive the centrifuges dimensions. Example taken from [67]. Image credits as in the figure. Analysis: R. Jungwirth.

An example of data foraging for a nuclear open source analysis is the one depicted in Fig. 9. While the knowledge of gas centrifuge enriching processes is well established in open literature, despite the exceptional amount of information released on the Iranian nuclear programme, at the time of this experiment (2014) little was available about the technical details of the Iranian centrifuges. By merging different sources of open information, including measurements derived from public photographs, the figure shows a 2014 attempt to derive the main technical characteristics and parameters of some centrifuges types developed by Iran. Another example of data foraging is the characterization of the DPRK Punggye-ri nuclear test site, with an effort to identify the exact location of the first four nuclear tests (Fig. 10).

In some areas data foraging can be supported by *ad hoc* tools. For example, in the domain of online news media monitoring tools to monitor, screen and collect information in multiple languages [69, 70] can inform the open source analyst and contribute to e.g. gaining early awareness of nuclear-related activities [37].

As with any data gathering campaign, a critical point is being able to assess “how much information is enough”. Having more information does not necessarily increase the accuracy of an analysis: studies show that “once an experienced analyst has the minimum information necessary to make an informed judgment, obtaining additional information generally does not improve the accuracy of his or her estimates. Additional information does, however, lead the analyst to become more confident in the judgment, to the point of overconfidence.” ([71], p. 52)

4.4 Model Building

Sometimes reliable data could be gathered, but the available paradigms are not adequate for making sense of the information there hidden. Where models are known, most of the times they are classified and cannot be considered being available in the public domain. There are therefore areas of investigation where a model building activity is necessary to be able to conduct an open source analysis. As mentioned, sometimes this activity is aimed at “rediscovering” in the public domain what is already known in classified environments,

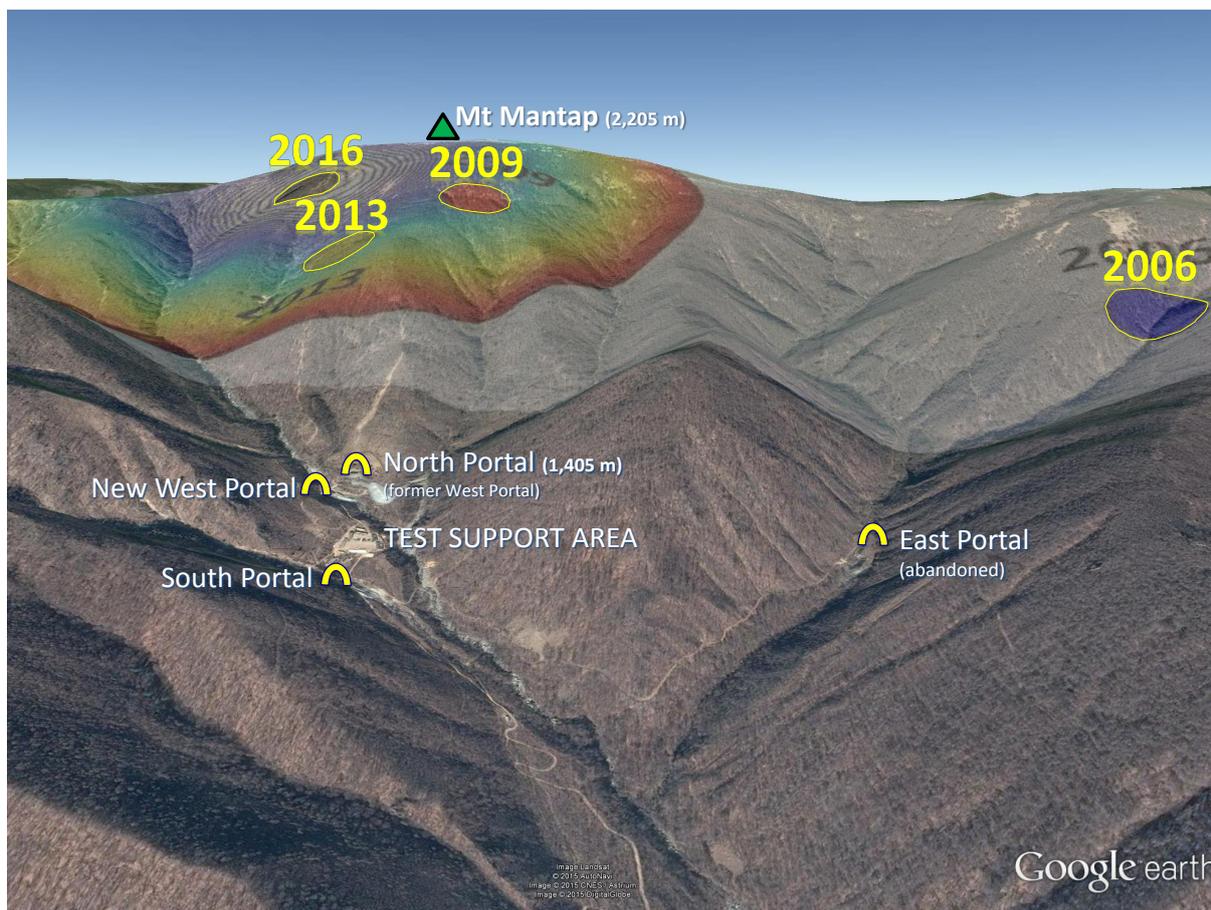


Figure 10: Example of data foraging OS analysis: geo-localization of the first four DPRK nuclear tests. Overview of the North Korea's Nuclear Test at Punggye-ri with Google Earth visualization of the CTBTO relative test locations plot released on January, 7, 2016. The georeferencing has been performed in line with the findings of Koblenz and Pabian about the characterization of the Punggye-ri test site. [68] The locations of the four existing tunnel portals and the central test support area are also shown. Image and analysis credit: F. Pabian.

as in the case of the suitability of the nuclear material available in a commercial nuclear fuel cycle for a military nuclear programme [72–75].

Sometimes there is a need to investigate the signature of nuclear facilities to compare with data and signals gathered from the open source domain:¹⁴ in early 2015, a German magazine published a reportage speculating on the possibility that a given site in Syria might have hosted an underground nuclear reactor. [76] Many commentators and nuclear experts dismissed the possibility based on the fact that the typical signatures of nuclear reactor cooling systems were missing.¹⁵ [77] The episode triggered the question “would it be possible to cool a small research reactor using ground-water only? And if it were possible, which would be the related signature available to a non-proliferation analyst?”. JRC carried out a “model building” exercise to try answering this question, highlighting the theoretical feasibility of this cooling approach, which would result in a signature substantially different from the one of typical research reactors cooling systems (Fig.11). [78]

Even when a well established and dependable paradigm is available, identifying the right model for a given analysis is a complicated task. For example, in an analysis of the technical difficulty for a given State to set up and operate e.g. a gas centrifuge enrichment facility—typical of an acquisition pathways analysis [79], it might be very difficult to select

¹⁴ for other examples, see also [22, 23]

¹⁵ The dismissal was mainly based on the issue of water management and the use of a near-by lake or river as a water source. The article reports that experts “have observed that the absence of large pipes that would be essential to supply a reactor with large amounts of water undermines Follath’s conclusion. After all, there is a reason why such reactors are built near rivers.” [77]

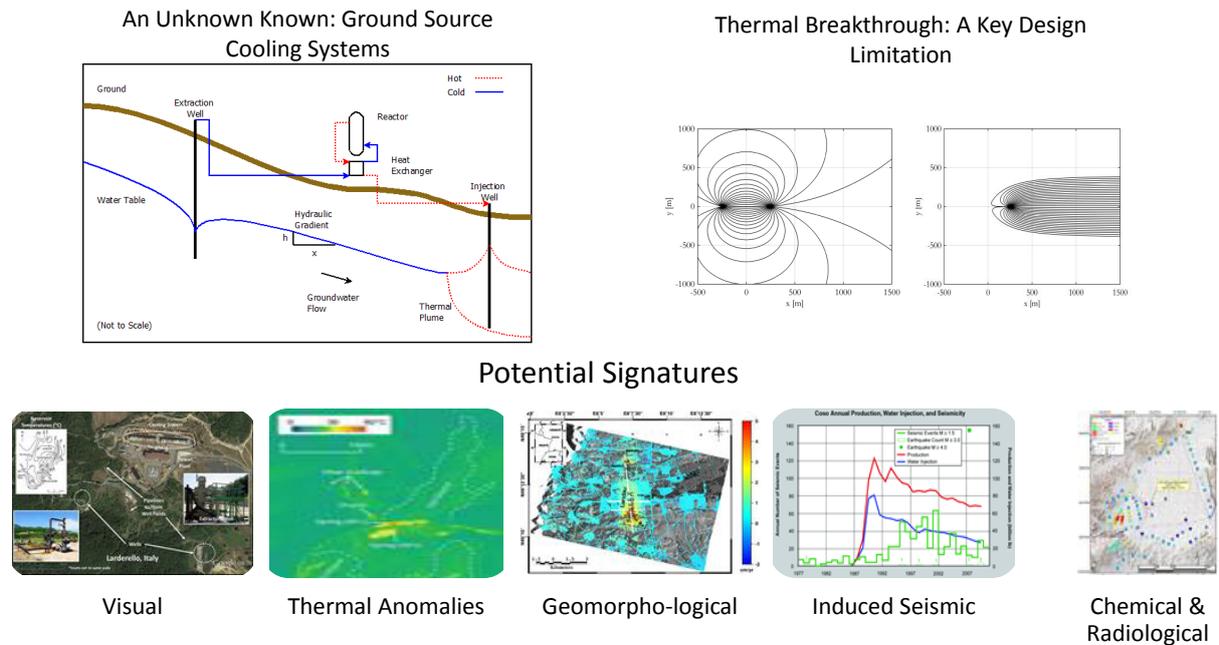


Figure 11: Example of model building OS analysis. Understanding operating principles of ground source cooling and potential signatures aids detection and deters construction. Challenges include ensuring sufficient recharge flow to prevent production well dry-out and overcoming structural engineering challenges on permeable foundations. If successfully implemented, the resulting signature is substantially different from that of typical research reactors cooling systems. [78]

the most appropriate model. Indeed, "... some analysts have claimed that '... all enrichment techniques demand sophisticated technology in large and expensive facilities' ([80], p. 31), suggesting that enrichment is out of reach of all but the most capable States. In contrast, others suggest that a small centrifuge plant is 'feasible for countries with no prior experience, that possess relatively little technical skills and which have relatively little industrial activity.' ([81], p. 47)." ([82], p. 6)

In this scenario, the dependability of the analysis will mainly depend on the quality of the model and on the final goal of the analysis.

4.5 Mystery Framing

In a nuclear proliferation effort the information related to the programme is usually a closely guarded secret, and both data and possible related models are unknown or highly uncertain. When both the quality of the data and the model are poor, the analyst faces a mystery framing scenario. This is typical of looking for clandestine nuclear weaponization activities, and is usually addressed by national intelligence services. Open source analysis could theoretically play a role but has to face formidable challenges in both data gathering and modelling. Moreover, failing in detecting a signal cannot lead to conclude the absence of a nuclear military activity.

Occasionally, the analyst stumbles into ambiguous data that do not fit neatly into any known signature but still cannot be dismissed as innocuous (or when there are multiple competing hypothesis that seem plausible). These are sometimes referred to as enigmas, and are another form of mystery framing. Usually the analysts facing enigmas will try to shift the scenario from a mystery framing one to either a model building or a data foraging one—often both. The closest non-classified example of an enigma in the nuclear non-proliferation community known to the authors was the Dair Alzour site in Syria, which was evidently subject to an air strike in September 2007.¹⁶ In the aftermath of the air strike, that caused surprise in the open source community, the claims that the destroyed site was

¹⁶ For a reconstruction of the events, see e.g. [83]

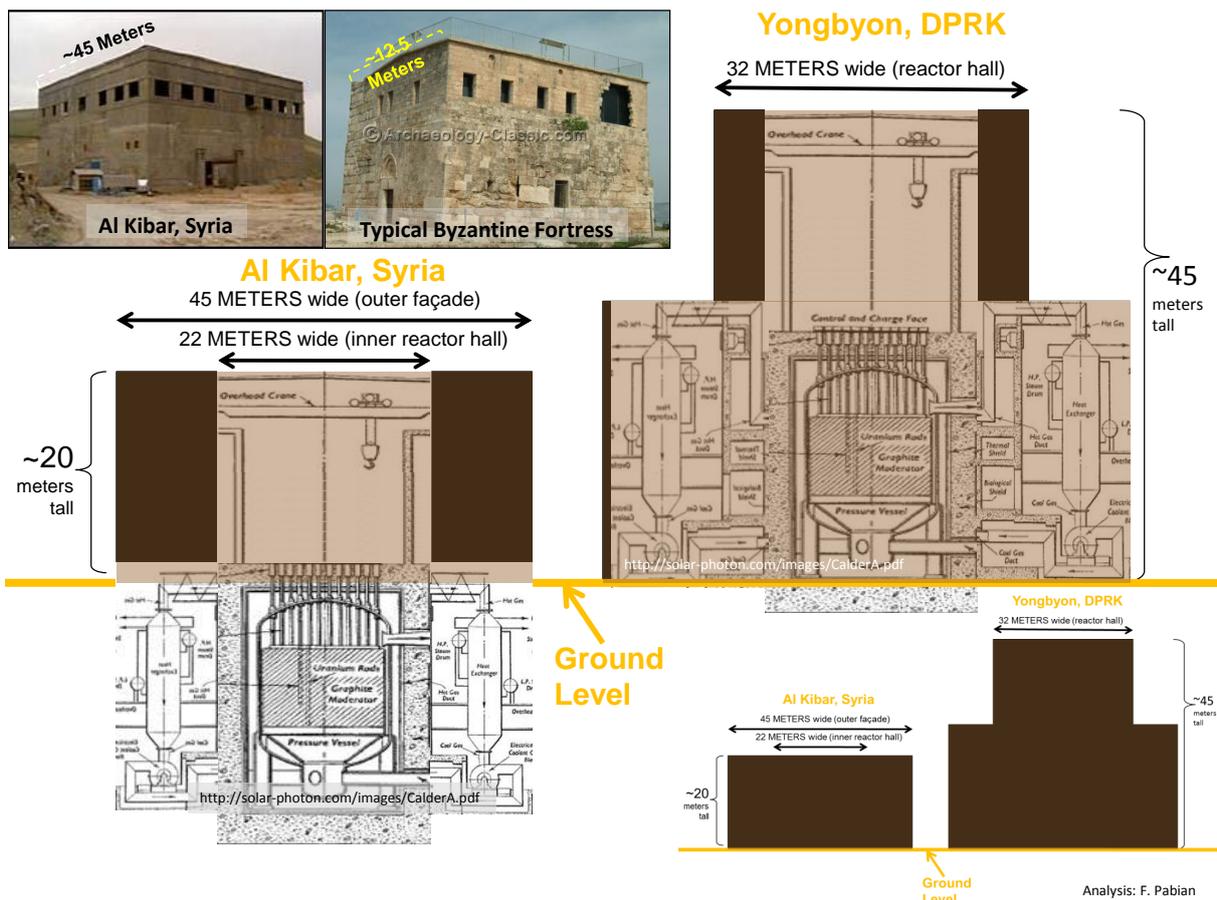


Figure 12: Example of model building applied to an “enigma”. The Figure shows an example of model building applied to the Al Kibar site to verify the possibility that the available data could be compatible with the presence of a Yong-byon-like research reactor. The top left images show the camouflage efforts to make the Al Kibar building looking like a typical byzantine fortress—just bigger. The center and bottom-right images show how the different size and shape of the Al Kibar building and the Yongbyon research reactor could be reconciled. The Figure is adapted from [89]. Analysis performed by F. Pabian and adapted with permission.

hosting a nuclear reactor [84] gave the analysts the possibility to switch the scenario to a data foraging one, as they started to gather as much evidence as possible to support or dismiss the claim. Some analysts disputed the possibility that the destroyed building could have been a nuclear reactor on the grounds that many of the usual signatures associated to this type of facility, such as “barbed wire or air defenses that would normally ring a sensitive military facility” ([85]) were not evident. Others investigated whether this discrepancy between the observed signature and the expected one could have been the outcome of a complex concealment effort [86,87], thus shifting again to a new analysis scenario: the model building one to see if a credible signature suppression effort might have been compatible with the observed data (see Fig. 12). In 2011 the IAEA assessed that “. . . the destroyed building was very likely a nuclear reactor. . .” ([88], para. 24).

5 Conclusions

Proliferation efforts are complex engineering programmes lasting several years, involving multiple sites and a lot of human resources with a wide range of expertise. As such, they are prone to leave behind several traces and potential cues that can be identified and monitored via open source analysis with varying degrees of success.

Within a systems thinking approach, open source analysis can be used to gain an insight of a given programme. Usually the process involves the gathering and analysis of a large amount of data and information, a very small percentage of it being relevant. The most common scenario implies filtering an enormous amount of data to end up with a sparse and heavily incomplete set of information to work with. When looking for a covert military engineering programme, the analyst will have to deal with low quality data and is always exposed to deliberate deception. Nonetheless, the insights that this overall picture could provide to the analyst is able to give him valuable insights about what a State might be pursuing.

This report suggested a system's thinking view of open source analysis in support to non-proliferation, identifying the possible dimensions involved and discussing different types of scenarios an open source analyst might face.

Modelling a nuclear engineering programme by explicitly acknowledging the peculiarities of its hard and soft layers allows the analyst to consider which are the types of insights that each layer can provide and which is the the best tool/technique to investigate them.

The hard layer can provide primary evidence of an infrastructure's topology, layout, shape and sizes. In many cases, the infrastructure's functions and operational status can be observed or deduced. The intrinsic geo-localization of this type of information allows and simplifies its representation in a visual form by overlaying it to satellite, aerial or ground imagery. The persistence of the signal emission can theoretically enable continuous signal collection, although in practice open source analyst would seldom be able to achieve it.

The soft layer can effectively corroborate and complement the information obtained by the hard layer signatures by providing insights on the people and the social organization operating the infrastructure. In particular, the possibility to monitor the social network activity of a given site or company can provide insights about the security practices (and therefore the potential sensitivity of the activities performed) and the evolution of a system (e.g. people and related activities being moved from one building to another, laboratory renovations and refurbishments, dismissal of activities). Some soft layer signals can in principle be subject to 24/7 monitoring by an open source analyst, and might provide a dynamic way of following the development and the evolution of a system potentially able to fill the temporal knowledge gap between the the availability of signals collected from the hard layer.

While the importance of the context in which a programme is conceived and deployed is often overlooked by technical analysts, it is essential to understand the evolution of a given programme and can provide valuable insights to the analyst on what, where and when to search for a particular signal/piece of information. Although the geo-political context is often a factor that cannot play a role in the estimation of a given metric, its knowledge often provides invaluable information for the understanding of the collected data.

Once a particular analysis is set, the analyst might face different types of analysis scenarios, according to the type of the problem to be tackled and the type of data at his disposal.

When high quality data and a well established paradigm with dependable models are avail-

able to the analyst, he/she deals with a "puzzle solving" type of analysis. The outcome can be considered as a dependable picture of the programme aspect under analysis.

When the analyst can rely on a well established paradigm but lacks dependable information, he/she is confronted with a "data foraging" type of analysis. The main task and challenge is to find and collect data as dependably as possible. The analysis dependability will be driven by the quality of the gathered data.

In cases where high quality data can be collected, but no dependable paradigm to make sense of those data is available, the analyst faces a "model building" type of analysis. The main challenge lies with the ability to build a dependable model to make sense of the available data, and the dependability of the analysis will mirror that of the model.

Sometimes the analyst is confronted with ambiguous data that do not fit into any known signature but still cannot be dismissed as innocuous. Given the fact that nuclear proliferation programmes are usually a closely guarded secret, analysts might be faced with both low quality data and models. To evolve, this "mystery framing" type of analysis will have to either improve the quality and quantity of data or ("data foraging") or find a suitable paradigm for making sense of the available data ("model building"), or more probably a combination of both.

An open source analyst in support of non-proliferation will have to handle many different forms of uncertainties, whose proper understanding is critical for the analyst to perform dependable assessments and for the decision maker to take properly informed decisions. Gathering evidence for the presence of indicators of a nuclear engineering programme in a state implies dealing with several epistemic issues, requiring a specific categorization in terms of fuzziness (impreciseness of definition) and incompleteness (existence of both "known unknowns" and "unknown unknowns"). An "open world" assumption might be beneficial to capture uncertainties related to these types of analysis.

A system's thinking approach to open source analysis has the potential to integrate synergically with the other tools available in the international treaty monitoring toolkit, helping in increasing the international community's confidence in its ability to detect an undeclared proliferation programme.

References

- [1] International Atomic Energy Agency. The Conceptualization and Development of Safeguards Implementation at the State Level - Report by the Director General (GOV/2013/38). Accessed via ISIS - Nuclear Iran, 2013. URL <http://www.isisnucleariran.org/assets/pdf/GOV201338.pdf>.
- [2] International Atomic Energy Agency. *IAEA Safeguards Glossary*. Number 3 in International Nuclear Verification Series. IAEA, 2002.
- [3] Rockwood, L. How the IAEA verifies if a country's nuclear program is peaceful or not: The legal basis. *Bulletin of the Atomic Scientists*, volume 74, no. 5:pp. 317–325, 2018. URL <https://doi.org/10.1080/00963402.2018.1507789>.
- [4] Smuts, J.C. *Holism and Evolution*. MacMillan and Co., Limited, 3rd edition, 1936.
- [5] Best, Jr., R.A. and Cumming, A. *Open Source Intelligence (OSINT): Issues for Congress*. Technical Report RL34270, Congressional Research Service, 2007.
- [6] International Atomic Energy Agency. The Safeguards System of the International Atomic Energy Agency. 2009.
- [7] Versino, C., Cojazzi, G.G.M., Contini, F., Tsois, A., Chatelus, R., and Tarvainen, M. Global trade data to support IAEA safeguards. *ESARDA Bulletin*, volume 45:pp. 14–21, 2010.
- [8] Marinova, E., Ardhammar, M., El Gebaly, A., Schot, P.M., and Ferguson, M. Analysis of nuclear relevant information on international procurement and industrial activities for safeguards purposes. In *Symposium on International Safeguards: Linking Strategy, Implementation and People, Vienna (Austria), 20-24 October 2014 . Book of Abstracts, Presentations and Papers*, CN-220-085. 2015.
- [9] Holland, B.R. *Enabling Open Source Intelligence (OSINT) in private social networks*. Technical report, Iowa State University, 2012.
- [10] Dori, D. and Sillitto, H. What is a System? An Ontological Framework. *Systems Engineering*, volume 20:pp. 207–219, 2017.
- [11] Barker, E. *The Politics of Aristotle*. Oxford University Press, 1946.
- [12] US Federal Aviation Administration. *NAS Systems Engineering Manual*, 3.1 edition, 2006.
- [13] INCOSE. *Systems Engineering Handbook: A Guide for System Lifecycle Processes and Activities*. International Council on Systems Engineering, 3.2 edition, 2006.
- [14] Blockley, D.I. and Godfrey, P. *Doing it differently: systems for rethinking construction*. Thomas Telford, 2000.
- [15] Renda, G. *Resisting Nuclear Proliferation through design — A systems Approach to Nuclear Proliferation Resistance Assessment*. Ph.D. thesis, University of Bristol, 2008.
- [16] van Wijk, L.G.A. *A Process Model for Nuclear Safeguarding*. Ph.D. thesis, University of Bristol, 2005.
- [17] Blockley, D.I. The Importance of Being Process. *Civil Engineering and Environmental Systems*, volume 27, no. 3:pp. 189–199, 2010.
- [18] Few, S. *Signal: Understanding What Matters in a World of Noise*. Analytics Press, 2015.
- [19] Katz, J.I. *Deception and Denial in Iraq: The Limits of Secret Intelligence and the Intelligent Adversary Corollary*. Technical report, Washington University, 2006.
- [20] Tsikerdekis, M. and Zeadally, S. Online deception in social media. *Communications of the ACM*, volume 57, no. 9:pp. 72–80, 2014.
- [21] Pabian, F., Renda, G., Jungwirth, R., Kim, L.K., Wolfart, E., and Cojazzi, G.G.M. Recent Developments Promoting Open-Source Geospatial Synergy: Emerging Trends

- and Their Impact for Nuclear Nonproliferation Analysis. In *Proceedings of the 56th INMM Annual Meeting, Indian Wells, USA*. 2015.
- [22] Zhang, H. and von Hippel, F.N. Using commercial imaging satellites to detect the operation of plutonium-production reactors and gaseous-diffusion plants. *Science & Global Security*, volume 8, no. 3:pp. 261–313, 2000.
- [23] Bernstein, A. Monitoring large enrichment plants using thermal imagery from commercial satellites: A case study. *Science & Global Security*, volume 9, no. 2:pp. 143–163, 2001.
- [24] Albright, D. and Hinderstein, C. The Iranian Gas Centrifuge Uranium Enrichment Plant at Natanz: Drawing from Commercial Satellite Images. *ISIS-Online*, 2003. URL http://www.isis-online.org/publications/iran/natanz03_02.html.
- [25] Schlittenhardt, J., Canty, M., and Grünberg, I. Satellite Earth observations support CTBT monitoring: A case study of the nuclear test in North Korea of Oct. 9, 2006 and comparison with seismic results. *Pure and applied geophysics*, volume 167, no. 4-5:pp. 601–618, 2010.
- [26] Pabian, F.V. and Hecker, S.S. Contemplating a third nuclear test in North Korea. *Bulletin of the Atomic Scientists*, volume 6, 2012. URL <http://thebulletin.org/contemplating-third-nuclear-test-north-korea>.
- [27] Lewis, J. The Tunnels at Punggye-ri: An Alternative View. 38 North, 2014. URL <http://38north.org/2014/03/jlewis032014/>.
- [28] Liu, J. North Korea's Punggye-ri Nuclear Test Site: No Indication of Nuclear Test Preparations. 38 North, 2013. URL <http://38north.org/2013/12/punggye122013/figure6-17/>.
- [29] Sadowski, F.G. and Covington, S.J. *Processing and analysis of commercial satellite image data of the nuclear accident near Chernobyl, USSR*. Technical report, US Geological Survey, 1987.
- [30] Albright, D., Brannan, P., and Walrond, C. Fukushima Crisis: Unmonitored Releases - Preliminary Assessment of Accident Sequences and Potential Atmospheric Radiation Releases. *ISIS-Online*, 2011. URL http://isis-online.org/uploads/isis-reports/documents/Accident_Sequence_Fukushima_31March2011.pdf.
- [31] Hecker, S.S. A return trip to North Korea's Yongbyon nuclear complex. *CISAC*, 2010. URL <http://iis-db.stanford.edu/pubs/23035/HeckerYongbyon.pdf>.
- [32] NTI. *Innovating Verification: New Tools & New Actors to Reduce Nuclear Risks - Redefining Societal Verification*. Technical report, Nuclear Threat Initiative, 2014. URL <http://tinyurl.com/pjk9eb8>.
- [33] Crosbie, V. What is New Media? *Sociology Central*, 2002. URL www.sociology.org.uk/as4mm3a.doc.
- [34] Heimerl, K., Menon, A., Hasan, S., Ali, K., Brewer, E., and Parikh, T. Analysis of smartphone adoption and usage in a rural community cellular network. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ACM, 2015.
- [35] eMarketer. Smartphone Users Worldwide Will Total 1.75 Billion in 2014. *eMarketer.com*, 2014. URL <http://tinyurl.com/qcjkj8>.
- [36] Duggan, M., Ellison, N.B., Lampe, C., Lenhart, A., and Madden, M. *Social Media Update 2014*. Technical report, Pew Research Centre, 2014. URL <http://www.pewinternet.org/2015/01/09/social-media-update-2014/>.
- [37] Cojazzi, G.G.M., van Der Goot, E., Verile, M., Wolfart, E., Rutan Fowler, M., Feldman, Y., Hammond, W., Schweighardt, J., and Ferguson, M. Collection and Analysis of Open Source News for Information Awareness and Early Warning in Nuclear Safeguards. *ESARDA Bulletin*, volume 50:pp. 94–105, 2013.

- [38] Plantin, J.C. *The Map is the Debate: Radiation Webmapping and Public Involvement During the Fukushima Issue*. Technical report, Université de Technologie de Compiègne, 2011. URL <http://dx.doi.org/10.2139/ssrn.1926276>.
- [39] Glassman, M. and Kang, M.J. Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior*, volume 28, no. 2:pp. 673–682, 2012.
- [40] Stefanidis, A., Crooks, A., Radzikowski, J., Croitoru, A., and Rice, M. Social media and the emergence of open-source geospatial intelligence. *Human Geography: Socio-Cultural Dynamics and Global Security*, US Geospatial Intelligence Foundation, Herndon, VA, volume 1:pp. 109–123, 2014.
- [41] Renda, G., Kim, L.K., Jungwirth, R., Pabian, F.V., and Cojazzi, G.G.M. The potential of open source information in supporting Acquisition Pathway Analysis to design IAEA State Level Approaches. In *Proceedings of the IAEA International Safeguards Symposium: Linking Strategy, Implementation and People*, Vienna. 2014.
- [42] Joseph, R. Open Source Indicators and Asymmetric Advantage in Security Planning. The international Relations and Security network, 2015.
- [43] Ullom, J. Enriched uranium versus plutonium: Proliferant preferences in the choice of fissile material. *The Nonproliferation Review*, volume 2, no. 1:pp. 1–15, 1994. URL <https://doi.org/10.1080/10736709408436562>.
- [44] Erickson, S.A. Economic and technological trends affecting nuclear nonproliferation. *The Nonproliferation Review*, volume 8, no. 2:pp. 40–54, 2001.
- [45] Ogilvie-White, T. Is there a theory of nuclear proliferation? An analysis of the contemporary debate. *The Nonproliferation Review*, volume 4, no. 1:pp. 43–60, 1996.
- [46] International Atomic Energy Agency. *Implementation of the NPT Safeguards Agreement and relevant provisions of Security Council resolutions in the Islamic Republic of Iran (GOV/2013/56)*. GOV/2013/56. IAEA, 2013.
- [47] International Atomic Energy Agency. *Implementation of the NPT Safeguards Agreement and relevant provisions of Security Council resolutions in the Islamic Republic of Iran (GOV/2014/58)*. GOV/2014/58. IAEA, 2014.
- [48] EU/E3+3 and The Islamic Republic of Iran. Joint Statement by EU High Representative Catherine Ashton and Iran Foreign Minister Zarif, 2013. URL http://www.eeas.europa.eu/statements/docs/2013/131124_02_en.pdf.
- [49] Walley, P. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, 1991.
- [50] Hacking, I. *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press, 2006.
- [51] Ferson, S. and Ginzburg, L.R. Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, volume 54, no. 2-3:pp. 133–144, 1996.
- [52] Jo, D.J. Determinants of Nuclear Weapons Proliferation. *Journal of Conflict Resolution*, volume 51, no. 1:pp. 167–194, 2001.
- [53] Gartzke, E. and Kroenig, M. A strategic approach to nuclear proliferation. *Journal of Conflict Resolution*, volume 53, no. 2:pp. 151–160, 2009.
- [54] International Atomic Energy Agency. *International Target Values 2010 for Measurement Uncertainties in Safeguarding Nuclear Materials*. STR Series STR-368, IAEA, 2010.
- [55] International Atomic Energy Agency. *IAEA Safeguards Technical Manual: part F - Statistical Concepts and Techniques*. IAEA-TECDOC TECDOC-261, IAEA, 1982.
- [56] Rabinovich, S.G. *Measurement errors and uncertainties: theory and practice*. Springer Science & Business Media, 2006.
- [57] Cui, W. and Blockley, D.I. Interval probability theory for evidential support. *International Journal of Intelligent Systems*, volume 5, no. 2:pp. 183–192, 1990.

- [58] Blockley, D. Analysing uncertainties: Towards comparing Bayesian and interval probabilities'. *Mechanical Systems and Signal Processing*, volume 37, no. 1:pp. 30–42, 2013.
- [59] Keet, C.M. Open World Assumption. In W. Dubitzky, O. Wolkenhauer, K.H. Cho, and H. Yokota (Editors), *Encyclopedia of Systems Biology*, pp. 1567–1567. Springer New York, New York, NY, 2013. URL https://doi.org/10.1007/978-1-4419-9863-7_734.
- [60] Shafer, G. *A mathematical theory of evidence*. Princeton university press Princeton, 1976.
- [61] Yager, R.R. On the Dempster-Shafer Framework and New Combination Rules. *INFORMATION SCIENCES*, volume 41:pp. 93–137, 1987.
- [62] Hacking, I. Combined evidence. In *Logical theory and semantic analysis*, pp. 113–123. Springer, 1974.
- [63] Kuhn, T.S. *The structure of scientific revolutions*. University of Chicago Press, 2nd edition, 1970.
- [64] Defense Systems Management College. *Systems Engineering Fundamentals*. Defense Systems Management College, Fort Belvoir, Virginia, 1999.
- [65] Bachastow, T.S. *Geospatial Intelligence & the Geospatial Revolution*. Penn State University, 2015.
- [66] United Nations Security Council. Resolution 2231 (2015), 2015. URL <http://unscr.com/en/resolutions/doc/2231>.
- [67] Renda, G., Cojazzi, G.G.M., Cagno, S., and Kim, L.K. A Systems Thinking Approach to Open Source Information Analysis for Nuclear Non Proliferation. In *Proceedings of the 14th International Workshop on Separation Phenomena in Liquids and Gases (SPLG 2017), 4-9 June 2017, Stresa, Italy*, JRC105992. 2017.
- [68] Coblenz, D. and Pabian, F. Revised Geologic Site Characterization of the North Korean Test Site at Punggye-ri. *Science & Global Security*, volume 23, no. 2:pp. 101–120, 2015. URL <http://dx.doi.org/10.1080/08929882.2015.1039343>.
- [69] Atkinson, M. and Van der Goot, E. Near real time information mining in multilingual news. In *Proceedings of the 18th international conference on World wide web*, pp. 1153–1154. ACM, 2009.
- [70] Steinberger, R., Pouliquen, B., and Van Der Goot, E. An introduction to the europe media monitor family of applications. In *Proceedings of the SIGIR 2009 Workshop - Information Access in a Multilingual World, Boston, Massachusetts, USA*. 2013.
- [71] Heuer, R.J. *Psychology of intelligence analysis*. Center for the Study of Intelligence, 1999.
- [72] Lovins, A.B. Nuclear weapons and power-reactor plutonium. *Nature*, volume 283, no. 5750:pp. 817–823, 1980.
- [73] Pellaud, B. Proliferation aspects of plutonium recycling. *Comptes Rendus Physique*, volume 3, no. 7:pp. 1067–1079, 2002.
- [74] Kessler, G., Höbel, W., Goel, B., and Seifritz, W. Potential nuclear explosive yield of reactor-grade plutonium using the disassembly theory of early reactor safety analysis. *Nuclear Engineering and Design*, volume 238, no. 12:pp. 3475–3499, 2008.
- [75] Bathke, C.G., Ebbinghaus, B.B., Collins, B.A., Sleaford, B.W., Hase, K.R., Robel, M., Wallace, R.K., Bradley, K.S., Ireland, J.R., Jarvinen, G.D., Bradley, K.S., Ireland, J.R., Johnson, M.W., Prichard, A.W., and Smith, B.W. The attractiveness of materials in advanced nuclear fuel cycles for various proliferation and theft scenarios. *Nuclear Technology*, volume 179, no. 1:pp. 5–30, 2012.
- [76] Follath, E. Evidence points to Syrian Push for Nuclear Weapons. Spiegel Online Interational, 2015. URL <http://www.spiegel.de/international/world/evidence-points-to-syria-still-working-on-a-nuclear-weapon-a-1012209.html>.

- [77] Ruhle, H. Assad "Nuclear Factory" claim a Hoax. *Asian Times*, 2015. URL http://www.atimes.com/atimes/Middle_East/MID-01-130215.html.
- [78] Kim, L.K., Jungwirth, R., Renda, G., Wolfart, E., and Cojazzi, G.G.M. Potential Signatures and the Means of Detecting a Hypothetical Ground Source Cooled Nuclear Reactor. *Science & Global Security*, volume 24, no. 2:pp. 92–113, 2016. URL <http://dx.doi.org/10.1080/08929882.2016.1184529>.
- [79] Renis, T., Yudin, Y., and Hori, M. Conducting Acquisition Path Analysis for Developing a State-Level Safeguards Approach. In *Proceedings of the 55th INMM Annual Meeting*. 2014.
- [80] Panel on Reactor-Related Options for the Disposition of Excess Weapons Plutonium and Committee on International Security and Arms Control. *Management and Disposition of Excess Weapons Plutonium - Reactor Related Options*. Technical report, National Academy of Science, 1995.
- [81] Kemp, R.S. The Nonproliferation Emperor Has No Clothes. *International Security*, volume 38:pp. 39–78, 2014. URL <http://hdl.handle.net/1721.1/89182>.
- [82] Kim, L.K., Renda, G., and Cojazzi, G.G.M. Methodological Aspects on the IAEA State Level Concept and Related Acquisition Path Analysis. *ESARDA Bulletin*, volume 53:pp. 100–110, 2015.
- [83] Follath, E. and Stark, H. The Story of "Operation Orchard": How Israel Destroyed Syria's Al Kibar Nuclear Reactor. *Spiegel Online International*, 2009. URL <http://tinyurl.com/kcdsjbh>.
- [84] AAVV. Background Briefing with Senior U.S. Officials on Syria's Covert Nuclear Reactor and North Korea's Involvement, 2008. URL http://www.dni.gov/files/documents/Newsroom/Speeches%20and%20Interviews/20080424_interview.pdf.
- [85] Broad, W.J. Syria Rebuilds on Site Destroyed by Israeli Bombs, 2008. URL http://www.nytimes.com/2008/01/12/world/middleeast/12syria.html?_r=0.
- [86] Albright, D. and Brannan, P. *The Al Kibar Reactor: Extraordinary Camouflage, Troubling Implications*. Technical report, Institute for Science and International Security, 2008.
- [87] Pabian, F.V. Strengthened IAEA Safeguards-Imagery Analysis: Geospatial Tools for Nonproliferation Analysis. In *Nuclear Nonproliferation Safeguards and Security in the 21st Century, 2012-06-11/2012-06-29 (Brookhaven, New York, United States)*, LA-UR-12-24104. 2012.
- [88] International Atomic Energy Agency. *Implementation of the NPT Safeguards Agreement in the Syrian Arab Republic (GOV/2011/30)*. GOV/2011/30. IAEA, 2011.
- [89] Pabian, F.V. Open Source Tools for Nonproliferation Application. Lecture notes for the 12th ESARDA Course on Nuclear Safeguards and Non-Proliferation, JRC-Ispra, Italy, March, 2014.

List of figures

1 System’s hard, soft and context layers 7

2 Multi-satellite, multi-sensor analysis for verification applications 9

3 Social Network integral signal layout on two different nuclear research centres 11

4 Example of the analysis of Social Network posts temporal line 12

5 Example of Social Network posts content analysis 12

6 Snapshot of the Iranian nuclear fuel cycle in 2013 and in 2014 13

7 Hard-soft layers integration to locate an Ebola cemetery in Liberia 16

8 Example of puzzle-solving OS analysis 18

9 Example of data foraging OS analysis: Iranian centrifuges 19

10 Example of data foraging OS analysis: geo-localization of the first four DPRK nuclear tests 20

11 Example of model building OS analysis 21

12 Example of model building applied to an “enigma” 22

List of tables

1 Research dimensions for OS analysis applied to nuclear non-proliferation . . . 17

(Page intentionally left blank)

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest to you at <http://europa.eu/contact>

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: <http://europa.eu/contact>

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: <http://europa.eu>

EU publications

You can download or order free and priced EU publications from EU Bookshop at:

<http://bookshop.europa.eu>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see <http://europa.eu/contact>).

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



EU Science Hub

ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub



Publications Office

doi:10.2760/87306

ISBN 978-92-79-98286-6