DRDC-RDDC-2016-N037

The role of Social Media in the Intelligence Cycle

Bruce Forrester[a], Kees den Hollander[b]
[a] Defence R&D Canada – Valcartier, 2459 Pie-XI North, Quebec, QC, G3J 1X5; [b] TNO, Oude Waalsdorperweg 63, NL-2597 AK The Hague, the Netherlands

ABSTRACT

Social Media (SM) is a relatively new phenomenon. Intelligence agencies have been struggling to understand how to exploit the social pulse that flows from this source. The paper starts with a brief overview of SM with some examples of how it is being used by adversaries and how we might be able to exploit this usage. Often treated as another form of open source intelligence (OSINT), we look at some of the differences with traditional OSINT compared to SM then outline the possible uses by military intelligence.

The next section looks at how SM fits into the different phases of the intelligence cycle: Direction, Collection, Processing and Dissemination. For the first phase, Direction, a number of questions are identified that can be answered typically by SM. For the second phase, the Collection, it is explained how SM, as an asset, transfers questions into methods and the use of different SM resources (e.g. marketer, cognitive behavioral psychologist) and sources to seek the required information.

SM is exploited as a multi-intelligence capability. For the Processing phase some aspects are described in how to deal with this capacity (e.g. enabling other intelligence sources) and also which techniques are used to be able to validate the SM sources used.

Keywords: OSINT, Social Media Intelligence, INT Cycle

1. INTRODUCTION

> *Populations in non-democratic states will increasingly employ social media tools in pursuit of democratic agendas. However, these governments are and will continue to develop more nuanced, insidious and effective mechanisms for exploiting social media while maintaining already pervasive control over traditional media sources. For these reasons, this analysis recommends that the Intelligence Community increase its attention to developing tools to observe, measure and report on the complex and evolving use of social media both by citizens and governments in largely closed societies. [1]*

Such is the conclusion of an article produced for the Office of the Director of National Intelligence's 2010 Summer Hard Problem Program. A similar conclusion was also reached by an international group of scientists and OSINT practitioners looking into the intelligence uses of SM for NATO [2]. However with hundreds of millions of people around the world using SM daily, navigating this ocean of activity is complicated. For instance, just one site Twitter, with over 500 million active users as of 2012, generates over 340 million tweets and 1.6 billion search queries per day (http://en.wikipedia.org/wiki/Twitter). It is clear that we need to better understand how we can tap into this wealth of data and information.

During the Arab Spring there was a significant rise in the volume of tweets. The first conclusive report of the role of SM during the Arab Spring states "Over the course of a week before Mubarak's resignation, the total rate of tweets from Egypt —and around the world —about political change in that country ballooned from 2,300 a day to 230,000 a day. Interestingly, the relative contribution of people not living in the region diminished significantly over this period" [3]. During this same period, the concerned governments were frantically trying to shut down access to the sites and were arresting identified SM activists [3]. It was also the first times that major news agencies significantly increased their use of reports and videos that were produced by the local populations within these countries. Such amateur and mostly

Next-Generation Analyst IV, edited by Barbara D. Broome, Timothy P. Hanratty, David L. Hall, James Llinas,
Proc. of SPIE Vol. 9851, 98510G · © 2016 SPIE · CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2242530

unconfirmed sources are commonplace in today's mass media, used for both speed of reporting but also where media access is limited or prohibited [4]. So it seems promising that SM in general, and Twitter more specifically, could be used to help understand populations and governments in countries of interest – a novel sensor for instability[5].

There is a degree of maturity in the non-military applications of analytical methods, used by industry that monitors and analyses SM content [6], which can be used as a basis for the development of specific algorithms used during the processing phase by the intelligence community. Research has been conducted using twitter for such purposes as election prediction [7-9], finding influential users [10-12], determining how information flows within the network [13-17], and earlier work on producing meaningful metrics [11, 18-20]. However there are some distinct differences between the military and civilian target populations that will require research and modification of current algorithmic models.

2. OVERVIEW OF SOCIAL MEDIA FOR INTELLIGENCE PURPOSES

2.1.1 What makes SM different from other OSINT sources?

Traditional OSINT [21, 22] comes from sources that are for the most part orderly. They contain content that has likely been edited, produced by professional authors, and use formal easily understood language. In large part, these sources are revenue generating and hence are easy to find, download and search. Access to and use of such sources is well defined under existing laws. Relative to SM, there are a limited number of sources to scour; SM data can be found just about anywhere and everywhere on the Internet. SM can be defined as "online communications delivered and interacted with, via text, audio and or video (e.g. Facebook, YouTube, weblogs and micro-blogs)" [23]. As shown in Table 1, SM sources have quite a different set of characteristics. These differences make it very difficult if not impossible to apply most standard data treatment tools to SM data.

Table 1. A Comparison of Traditional OSINT Sources to SM Sources

Traditional OSINT Sources	**SM Sources**
Academic research, books, encyclopaedias, business and government documents, grey literature, images, journals, periodicals, broadcast media, maps, newspapers, radio	Blogs and micro-blogs, Internet forums, user-generated FAQs, Chat, podcasts, online games, tags, ratings, comments, social networking sites, online video, wikis, search engines, social bookmarking
Edited	Not-edited
Written by professional authors	Written by anyone and everyone
Use of proper grammar, spelling and punctuation	Anything goes
Minimal use of sarcasm, street language, profanity	Anything goes
Usually well catalogued and contains standardized metadata	Could be found anywhere, with non-standard folksonomies and tags
Easy access	Must be sought out
Some requiring subscription fees	Generally free. However some have limited access due to API restrictions
Well-defined use of acronyms	High use of text and chat acronyms – Netlingo [24]
Use of most common language dialect	Local dialects and special use of words that have different generalized meanings (i.e. "that is sick" meaning "that is really cool", could also mean that "this is really hot")
Audio is clear, audible and generally of high quality	Audio quality is highly variable
Video and photos are generally of high quality	Video and photo quality is highly variable

An appreciation of the differences was highlighted in Proceedings of the International Conference on Web Search and Web Data Mining 2008 [25]:

> *The main challenge posed by content in social media sites is the fact that the distribution of quality has high variance: from very high-quality items to low-quality, sometimes abusive content. This makes the tasks of filtering and ranking in such systems more complex than in other domains. However, for information-retrieval tasks, social media systems present inherent advantages over traditional collections of documents: their rich structure offers more available data than in other domains. In addition to document content and link structure, social media exhibit a wide variety of user-to-document relation types, and user-to-user interactions.*

2.1.2 What the experts say about potential intelligence uses

This section was derived from three international meetings involving scientists and OSINT practitioners from nine different NATO countries [2]. The basic question asked was: How could you envision using SM information and data from open sources? Table 2 summarises the answers.

Table 2 – Potential uses of SM Sources for Intelligence

Phenomenon	**Military/Intelligence impact**	**Intelligence Product**
1. Potential social uprisings: • What is the stability of current government in country? • What are the issues with the people? • Are things escalating? • What are the trends?	• Strategic and Operational • Contingency plans • Operational Plans • Peacekeeping	• Early warning and indicators • Trend watch • Response to standing RFI • Alerting service • Basic intelligence (baseline) • Threat assessment • Country studies
2. What is happening in remote areas where there are few other sources of information available?	• Current up-to-the-minute SA (situational awareness) of a particular area • Enables operational planning • Tactical threat assessment	• Response to targeted RFI • Alert service • Threat assessment • Information bulletin
3. Monitoring and pattern analysis looking for criminal / terrorist / insurgent activities.	• Cyber issues (taking down subversive sites) • Targeting • Understanding ECOA • Planning (collection, ops, tactical etc.) • Disrupting the insurgency cycle before the ACT stage	• Response to RFI • Threat assessment • Standing products
4. Targeting (non-kinetic) (i.e. profiling); identifying and getting information about particular person of interest, groups, organizations.	• Targeting • Understanding ECOA –Enemy Courses of Action • Planning (collection, ops, tactical etc.) • Understanding the ideology	• Response to RFI • Threat assessment • Profile • Structure of orgs

As an example, here are some sample questions concerning a potential social uprising that could likely be answered at least in part by SM:

- What is happening in country X?
- What is the population of country X talking about online?
- What are the hot topics?

- How are the grassroots discussions different from the mainstream media or government discussion?
- What are the issues that other countries are discussing with respect to the country X?
- What is the sentiment of the discussions?
- Is sentiment changing? - getting more aggressive or passive?
- What opposition exits?
- What organizations are acting within country X?
- Is there any religious polarization?
- Who are the thought leaders that are emerging in discussions?
- What is the rest of the world saying about country X?
- Where is the discourse community? – Who are they?

2.1.3 Uses of SM for Modern Warfare

We define a simple but inclusive and effective definition of SM: “any Internet based application that permits user-generated content to be shared.” There are thousands of such sites available via the Internet. A somewhat dated (circa 2013) but still operative snapshot of the diversity of SM platforms is shown in Figure 1. This figure shows a distinct separation between functionality and purpose of sites. Since 2013, platforms have been incorporating useful and popular functionality found on other sites and now most platforms have a social networking, photo & video sharing, and instant-messaging functionality. Some platforms that were once popular have seen a decline with migration of users to different sites (the MySpace to Facebook migration is a good example).

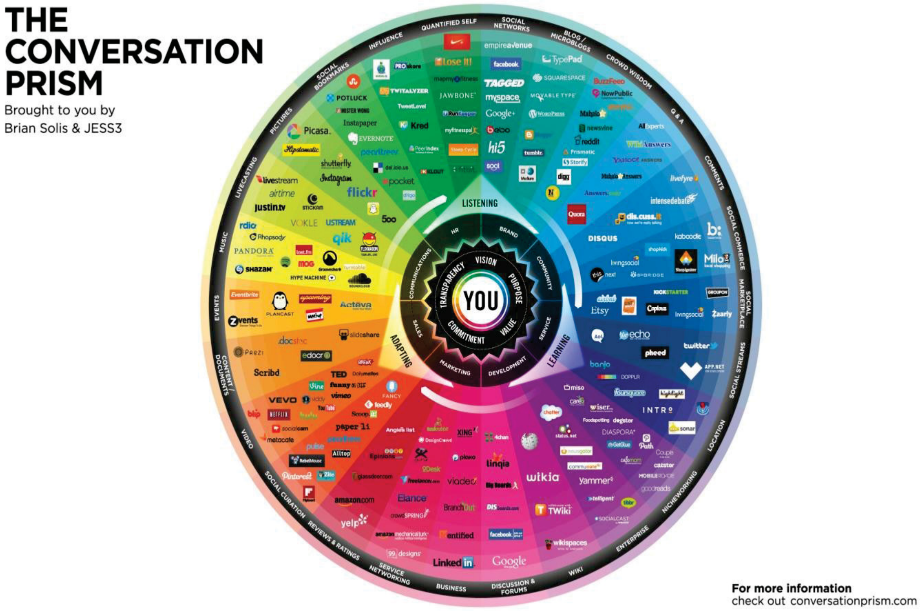


Figure 1. The diversity of SM circa 2013.

There is abundant evidence of the use of cyber influence activities in Jolicoeur and Seaboyer [26] who reported on the Russian cyber influence activity before and during the annexation of the Crimea in March 2014. In addition to many traditional cyber techniques (Website defacements, Distributed denial of service method, Intrusions and Infiltrations) there were significant SM campaigns by the Russians targeted at Ukrainians and to inform Russians. When directly targeting Ukrainians, the Russians launched massive campaigns to crowd out or counter the Ukrainian narrative. Here the Russians were portrayed as brining peace and protection. As for the Russian audiences, the campaign used VKontake (the major Russian social networking platform) for fundraising and coordinating volunteer organizers to help the annexation effort. These organizers in turn would contribute to the discussion on SM or even launch cyber-attacks (to facilitate, it is believed that Russia provided links to DDoS software and instructions on how to employ). It was used to help silence opposition and boost Putin's image as a hero [27]. Russia essentially conducted operations with minimal troop involvement and casualties with the positive result of annexation of the Crimea. However, the Ukraine fought back online. The Ukrainian protest movement was active in this SM battlefield by posting pictures of former president Yanukovych receiving the "bird" on Facebook and Twitter. This was meant to show disrespect and disdain for the pro-Russian president.

Given the fluidity of users and the diversity of SM platforms it is challenging to decide where to focus online activities for detecting influence activities. One way to choose platforms on which to focus is by determining which ones your target audiences use. A 2011 study of SM exploitation tools concluded that social networking platforms "represent a very powerful tool for use by legitimate political groups and lawful dissent, as well as those espousing insurgency and insurrection." [6]. Time, as well as reports and studies [3, 26, 28-30] have shown that indeed, SM is an ideal platform for organizational activities such as recruiting, fund raising, radicalization and broadcasting videos of acts.

A follow on study by Forrester et al. [31] expanded upon Labrèque [6] and put more context around the use of SM by insurgents. An important framework used in both studies follows what is called the insurgent wheel, reproduced in Figure 2. In the study it is demonstrated that this planning cycle can be typically applied to asymmetrical warfare. Each segment represents activities that insurgent groups must undertake if they are to conduct those actions necessary to achieve their objectives. Moreover, these phases and activities are not unique, but are typical of those undertaken by any group involved in organizing and executing operations. Military intelligence efforts are typically aimed at identifying activities in each of these phases, and their associated key indicators, to develop a series of "signatures" that point to the group's intentions. This in turn may point toward potential insurgent targets and what stage they are at in preparing to act.

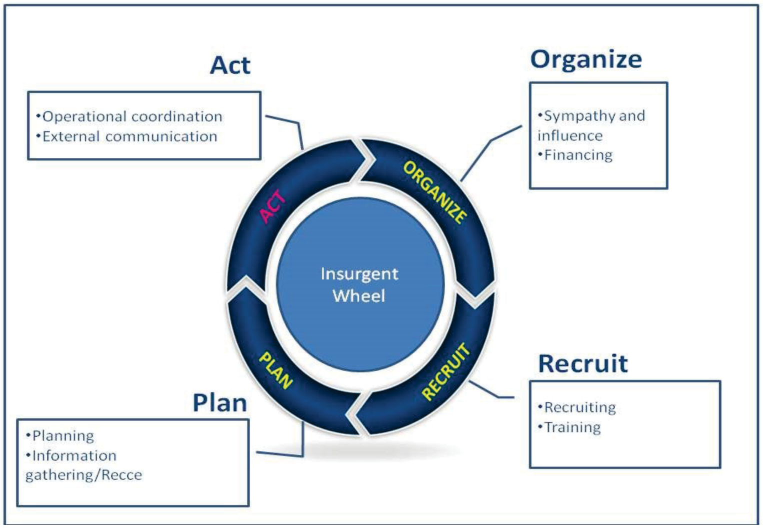


Figure 2: Insurgents activities of interest and the areas for investigation. In order to Act, insurgents need to first Organize, then Recruit and Plan. Indeed, these steps can be generalized to any organization's actions.

Perhaps the most significant finding of [6] is that the "Organize" phase has the most number of SM categories available for use. Placing emphasis and resources on this phase, before the deployment of troops, makes sense as this is the best place to intervene in a potential insurgency. While the insurgent case was used in this past research, the insurgent wheel is generic and can be used to describe the phases and steps that would be required of any organization conducting an act. So the above results are likely applicable to more general cases, for example a state's military would need to use the same basic steps for force generation and deployment.

3. SOCIAL MEDIA AND THE INTELLIGENCE CYCLE

3.1.1 Introduction

Social Media is a useful capability for intelligence to be combined with other sources, but needs to be processed as a separate source as well (Single Source Analysis) in order to be:

- able to provide information in time due to incoming real time data; and
- aware of to which information own commanders and troops are exposed.

Since SM can be obtained in near-real time, the latter effect requires an extra effort from intelligence to assess situations more quickly in certain cases to be able to withstand certain sentiments: commanders (actually: "everybody") have direct and easy access to SM information and are susceptible to influences by this medium.

Because many people have access to SM (and internet in general), it is used also for verification for HUMINT sources: do HUMINT sources give better information than already is known from the internet?

Several unique characteristics of SM fit in very well in the search for information to produce intelligence. This section gives a brief overview of some aspects of SM in the different phases of the intelligence cycle: Direction, Collection, Processing and Dissemination. Described are the usage and relevance of SM within the Intelligence Cycle. The driving force of the Intelligence Cycle is the Command & Control process with its commander/decision maker that requires intelligence to fulfil its tasks. Therefore this part is taken into account as well. Further, the perspective of SM as a component of the Intelligence Capability is taken into account; the different aspects that contribute to the capability of the intelligence as a whole. First the Intelligence Capability and its components are described in general. These components will be integrated in the second part of this section in which the contribution of SM in the Intelligence Cycle is described.

Some general remarks:

For SM reliability of the source and the evaluation (accuracy) of the information content (data) are uncertain aspects, similar to other collection methods (HUMINT, SIGINT, IMINT, etc., etc.). But SM in particular adds another abstraction as it is virtual and not a real world. Whether the information is true or not: the intelligence cell must be able to recognize its effect or impact on the online community (in the general population and on its own decision makers and troops).

The condition is that SM as a source is used sufficiently and containing relevant information regarding the Area of Intelligence Interest. Not in all circumstances is SM used as intensively as in most of the Western World.

3.1.2 The Intelligence Capability

The Intelligence Capability consists of the following seven components:

- Taskability, requestability of assets or outsourcing;
- The experience and education level of personnel;
- The thematic expertise of intelligence personnel;
- Freedom of Movement;
- Characteristics during the mission preparation;
- Characteristics during the mission execution; and
- Characteristics of the information exploitation.

3.1.3 Taskability, requestability of assets or outsourcing

This component is about the fact whether collection and processing assets are under command of the unit that requires intelligence or not. If they are *taskable*, then information needs can be (partially) fulfilled by directing own assets. *Requestable* assets (assets that are "borrowed" from other commanding units) can be deployed for its own purpose. Outsourcing is hiring or using other than your own assets to collect and/or process information for intelligence for your purpose.

For SM all three options occur, but in certain circumstances SM is suitable for outsourcing collection and/or processing of SM sources: many commercial services are available and even citizen participation can be an option in certain cases by requesting information about a broad or particular item. These services could bring in required experience, thematic expertise capabilities or even extra capacity in case there is a shortage of personnel.

Not only data but obtaining already available collection and processing products from SM sources could enrich one's own information position under particular circumstances. Research journalism (e.g. Bellingcat) and citizen participation ("*Be your own Sherlock Holmes"*) are domains and initiatives where citizens using and publishing their research on SM.

In some cases it is not possible to give specific directions to outsource agencies and therefore it might give poor results regarding the Intelligence's Priority Intelligence Requirements (PIRs), i.e. the important themes to be addressed by the Intelligence Cell to satisfy the Commander's information needs. Also the lack of awareness regarding the context of the information needs *("For what is it used?")* might give data/information and products that do not satisfy (one of) the PIRs.

Another disadvantage is that using external assets makes others aware of one's particular intentions, for both Western and non-Western services/platforms. Often in requests for that kind of service search terms are more broadly defined than one is interested in, however this costs more in time and budget.

For specific subjects/projects and also for (near) real time intelligence support during missions, one's own (taskable) and other (requestable) assets are required and cannot be outsourced as time is critical for a running operation.

3.1.4 The experience and education level of personnel

In general, this Intelligence Capability is highly dependent of the task to fulfil by the personnel.

For simple monitoring, standard search and processing to look for trends and Warning & Indicators could be accomplished by general personnel, supported by modest interactive collection and analysis tools. In case of more complex monitoring, the generation of Warnings & Indicators should be done by the organization's analysis capable staff. Their products and knowledge then must be implemented in this type of tooling. This also counts for SM.

We consider only SM exploitation where no interaction takes place between analysts and users in order to collect data. This situation is different in cases like Cyber HUMINT operations. In these operations active interaction or intrusion in SM takes place in order to collect information. Personnel performing such a task must be highly trained, must have expertise and a way of interacting within certain communities that could be expected from their (virtual) profiles. They also must have knowledge about effects of their particular interaction with or influencing of the environment. Furthermore, its activities require a solid legal basis but this paper is restricted to passive collection only.

3.1.5 The thematic expertise of intelligence personnel

Examples of themes or topics are: military aspects, security, geo spatial aspects, anthropology, demography, culture, history, economy, governance and religion. Information containing all these aspects can be obtained from many SM sources, and therefore require a very broad expertise or a combination with expertise from other collection methods (SIGINT, IMINT, HUMINT, etc.) or external sources and agencies (universities, industry, commercial parties, NGOs, etc.).

3.1.6 Freedom of Movement

This component has to do with the level of Freedom of Movement or restrictions of assets for the collection of information during their deployment. Options are: total freedom, required coordination with other troops (Ops Order), operating only along own troops, operating only within the own Area of Operation (AoO).

Working with SM has no limitation in the physical Freedom of Movement: an internet connection, anonymous surfing capabilities and some other arrangements are required. The location from where this operation takes place is independent from a Headquarter or AoO of a mission that is to be supported. In case there is no need to be on a site in the mission area to be able to give direct support like (near) real time support, this capability could remain at an operating base or compound.

There are situations where a legal permission is demanded (up to Ministerial Approval) and give constraints to what actions are allowed on the internet (interacting, influencing). The Freedom of Movement can also be limited by rules for labour circumstances (number of working hours a day, additional payment for abnormal working hours, etc.) if there is a difference between the location where the SM support takes place and the AoO.

3.1.7 Characteristics during the mission preparation

During the mission preparation collection and processing units require a certain amount of time to plan, prepare and physically transfer (and recover) assets.

Little preplanning is required to start the collection and processing on SM data (assuming that such a capability already exists within the organization), other than defining indicators to look for, derived from the Commander's Intend (CI), Commander's Critical Information Requirement (CCIRs), etc. In general, no transfer of assets is required to be able to perform collection and processing.

If the AoO or Area of Intelligence Interest (AoII) is fairly new and needs to be explored for the first time, the mission is not started yet and/or there is no possibility or permission to enter the AoO, SM has the advantage to collect and analyze data of the perceived Area even before the mission preparation starts or when the rules of engagement prohibits the presence of other intelligence collection means. Also is it possible to retrieve historical data, i.e. data posted and put on SM back in time (long) before there was a need to collect information for intelligence.

If this historical data has to be obtained through commercial services, there might be a disadvantage that these external services could figure out one's particular intentions. The same problem for outsourcing collection and processing. This might be an even bigger challenge when obtaining data from non-Western services/platforms.

3.1.8 Characteristics during the mission execution

During the mission execution the deployment, the agility and performance of collection assets are dependent on range/standoff, endurance/sustainability, terrain characteristics/line-of-sight, available flying or operating hours/number of sorties, sensor quality, etc.

Regarding SM, these (mostly physical) characteristics could be translated as follows: the range to be able to collect (detect) data (activities) is almost infinite as long as there is access to particular sources, networks and non-Western services/platforms. SM as a source produces a potential infinite stream of data and information; however this big data source is a real challenge as well for networks, data storages, availability of personnel etc.

Again, particularly in case of near-real time support, collection and processing should preferably be carried out in the AoO to be able to respond to quick requirements of the Commander and troops in the field and changes in the environment.

3.1.9 Characteristics of the information exploitation

This component is about near-real time transfer of data from platforms, standards data storage and retrieval, standardized structure and lay-out, data fusion capacity, processing speed, etc. Most of these aspects have been described in the above sections.

For SM there is a real challenge to cope with different data and product formats, structures and lay-outs (pictures, videos, unstructured texts and metadata) that change in time. Also SM non-Western sources, services and platforms of (and their relevance) might appear, disappear and change in time.

Near-real time processing and transfer of data is a challenge: data enrichment and filtering by collection and processing systems are therefore necessary to be able to select and use the most relevant data.

Data enrichment like time and space references in the content of messages, sometimes can be found in the metadata, but in a lot of cases these type of data need to be extracted in another way. For example, by using techniques that combine selected wording in the content indicating time and/or space aspects.
The increasing commercial interest of SM causes a tendency to add more and more metadata in messages, which is convenient from an analysis perspective, but there might be some pressure to slow down this tendency due to privacy issues.

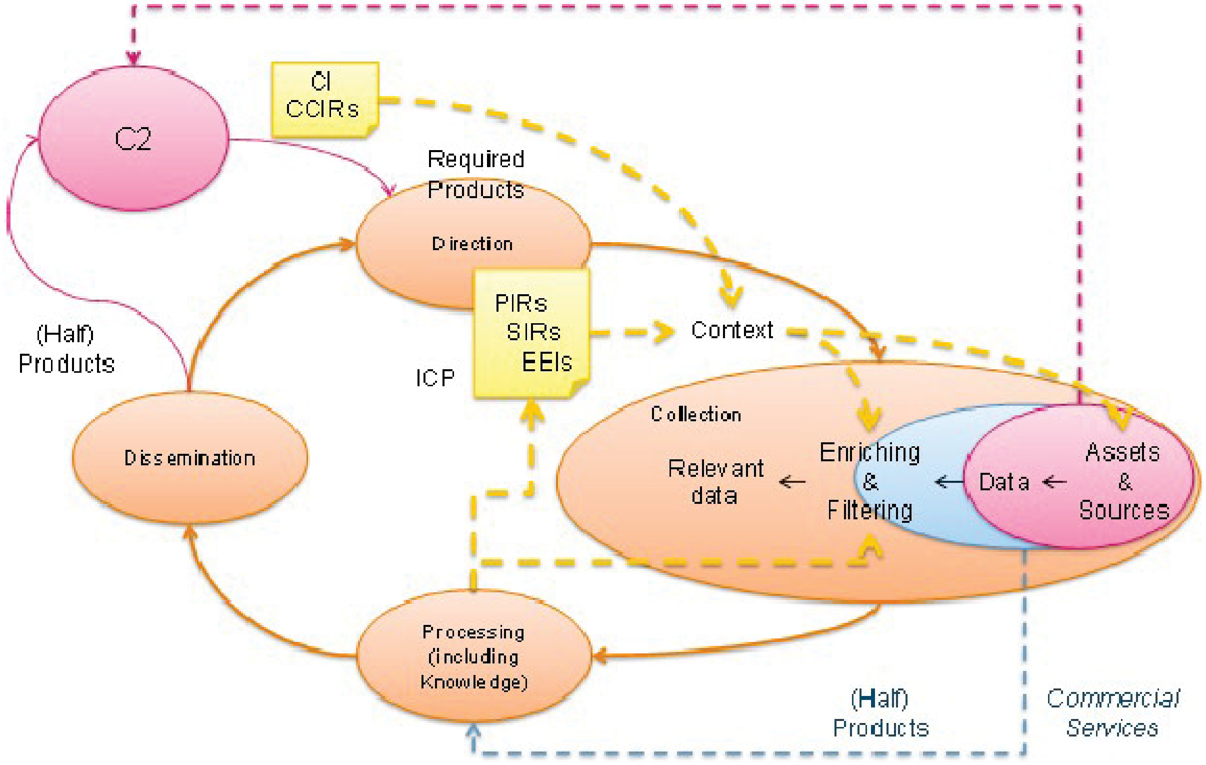
This example of enriching data (time and space) supports the filtering of relevant data. There are many different enrichment and extraction techniques (Text parsing, Image analysis, Auto Translation, Natural Language Processing and Machine learning). Other examples of filters can be derived from the information requirements in the Intelligence

Collection Plan (ICP). This means that analysts must be have wide access to filter settings in acquired collection and processing systems. However, many commercial products have only predefined filters with no user control. Regarding filtering, there was a variety of functionality, but in most of these cases filtering of data streams were not possible at all, due to lack of filter setting functionality or only possible with the intervention of the company.

3.1.10 Command & Control and the Intelligence Cycle

The Intelligence Cycle is the sequence of activities whereby information is obtained, assembled, converted into intelligence and made available to users.

The Intelligence Cycle consist of the four sub steps Direction, Collection, Processing and Dissemination, but is initiated and maintained by the Command & Control (C2) process. Figure 3 gives an overview of the different items regarding these mechanisms.



3.1.11 Direction

Direction is the determination of intelligence requirements, planning the collection effort, issuance of orders and requests to collection agencies and maintenance of a continuous check on the productivity of such agencies.

Again, Intelligence starts with the Commander and its intend or Commander's Intend (CI) and more narrowly: the Commander's Critical Information Requirements (CCIRs). Derived from the CCIRs a number of information and intelligence products are defined and transferred into particular information requirements. These are then transformed into an Intelligence Collection Plan (ICP) (or Intelligence Collection eXploitation Plan (ICXP)), in which information needs are prioritized and collection means are planned to be deployed such that the collection takes place in an optimized way.

This ICP is split up into three different levels that worked out the information requirements in more and more detail: the Priority Intelligence Requirements (PIRs), Specific Information Requirements (SIRs) and Essential Elements of Information (EEIs).

Examples of typical general PIRs are: Governance, Culture, Economics and Security. More specific subjects are related to particular events, e.g. IEDs and Elections. The priority of each PIR and subsequent detailed questions vary in time and location during the total mission. For example: the PIR on Elections is active two months before and a month after an election takes place in a particular area of the AoO

The dynamics of the steps from CI to the EEIs comprise the Context of what to look for given the task of the Commander. Actually, here lies the biggest challenge: how to define and couple context in general and context of the Commander and Intelligence in particular to incoming data streams? In a simple way, these elements from CI to the EEIs need to direct the Assets & Sources as well as Enriching and Filtering of data coming in from the next step in the cycle: Collection.

Another input for the assets & sources and enriching and filtering with the same challenge comes from the Knowledge gained further up in the Intelligence Cycle (Processing). This knowledge could be used directly in the f assets & sources and enriching and filtering, or via (an adaption of) the ICP.

To define EEIs and define and/or translate them in such a way that systems could understand them (Technical EEIs) is not an easy job. One reason is that the developers of an ICP and the people who use the collection and processing systems (including the filter settings) speak different "languages"; it takes a lot of effort to understand each other. Advantage of such a process is that answering EEIs in such a way there is a direct link and awareness with fulfilling the ICP.

The Information Requirements Management & Collection Management (IRM&CM) functions derive the best available assets and sources & agencies to collect the required data. SM as a capability and source is one of these assets. Since SM is, in and of itself, a wide variety of sources, content and methods, the SM (or OSINT) Cell has to support the IRM&CM function in how the collection plan gets produced.

Table 3 shows an example of SM contributions towards the ICP. The first PIR is about Security at the tactical level.

Table 3 Example of SM contributions towards the ICP (Tactical Level)

PIR	**SIRs**	**Social Media contributions**
PIR: Security (Tactical Level)	Who are the groups operating in the area? How are they influencing each other? Where are they positioned?	Basic intelligence (baseline): Clusters based on SM Activity Social relations / Sentiment between groups based on SM activity Geographical presence based on SM activity
	What threats (Social Unrest, Strikes) to the force can be expected during the fulfilment of operations at the tactical level (i.e. (near) real time threat assessment)?	EEIs: Planning directions are shared via SM Inflammatory content is shared via SM Inflammatory content and the amount of SM; Interactions (retweets etc.) are rapidly increasing Social groups are created online to facilitate communication Encrypted communication means are used Steganography is being used to transfer messages (Preparations for) life casting are in place/being made. Calls for action are posted on SM

Note: detecting the EEIs in Table 3 during the next step in the Intelligence Cycle (Collection) is just one aspect to identify if there is any threat to the force. The real contribution of each EEI is another aspect of context to be able to determine the actual level and type of threat. This contribution is derived from knowledge gained in Processing.

3.1.12 Collection

Collection is the exploitation of sources by collection agencies and the delivery of the information obtained to the appropriate processing unit for use in the production of intelligence.

An advantage of SM as a source is the information of new areas or topics can be explored even before the mission preparation starts or rules of engagement prohibits the presence of other intelligence collection means. Also is it possible to retrieve data posted and put on SM back in time (long) before there was a need to collect information for intelligence.

As stated in above, given the information requirements derived from the ICP, the SM Cell has to look at what sources and expertise should be brought to bear (marketer, cognitive behavioural psychologist, etc. and relating sources) in order to seek the required information, and process it after collection. Also outsourcing collection (and/or processing) should be looked at to bring in required experience, thematic expertise capabilities and/or extra capacity. All this leads to the direction on selected assets and sources, and the direction on the enrichment and filtering of data coming from these sources.

Reliability and representativeness of the source and the evaluation (accuracy) of the information content partly determine the value of (part of) the result of an information or intelligence product. If known, this could influence the selection of assets and sources as well.

3.1.13 Processing

Processing is the conversion of information into intelligence through Collation, Evaluation, Analysis, Integration and Interpretation. The first four activities can be (semi-)automatic, the last step is seen as a human activity only.

On one hand, the intension is to select sources and filter the information that is relevant to the mission or tasks to fulfil. This (selected) information is then used to be processed for intelligence. As figure 3 shows the CI, CCIRs and the (elements in the) ICP can support the determination of the correct sources and filtering of data coming from these sources (Collation, Evaluation and Analysis). Integration of the incoming information with the information already available is the last step before being able to interpret the meaning and relevance for the Commander.
On the other hand, one wants to be aware of potential threats or other events that is not part of the ICP. This could be done by general monitoring of trends and anomalies (Warning & Indicators), although the problem still is: What is a trend? What is an anomaly? To be able to automate this, machine learning methods should be combined with knowledge gained from experience.

For SM a number of techniques are available for extraction of the data (Collation, Evaluation and Analysis). Examples are Text parsing, Image analysis, Auto Translation, Natural Language Processing (NLP), Machine learning, facial recognition, Video extraction, speech to text, and semantic searching. Other analysis techniques exist that deal with the determination of validity and veracity of SM sources.

As stated above, there are more factors that need to be taken into account as part of the data collected based on the identified context. Identifying EEIs and the collection of data based on the EEs are not sufficient. In case of the determination of threats, one needs to identify the contribution of each EEI towards a particular threat. Table 4 shows an example of the contributions (diagnostic values) of each EEI to the threats *Social Unrest* and *Strike*.

Table 4 Diagnostic value towards threats: Social Unrest and Strike (example)

Social Media Indicators	Social Unrest	Strike
Inflammatory content is shared via Social Media	Medium	Weak
Inflammatory content and the amount of Social Media. Interactions (retweets etc.) are rapidly increasing	Medium	Weak
Planning directions are shared via Social Media	Strong	Strong
Social groups are created online to facilitate communication	Medium	Weak
Encrypted communication means are used	Weak	Medium
Steganography is being used to transfer messages	Weak	Medium
(Preparations for) life casting are in place/being made.	Medium	Medium
Calls for action are posted on Social Media	Medium	Weak

Based on J. Wevers MSc./Dr. B. van der Vecht, Dr. A.C. van den Broek [TNO 2016 R10171)], with key sources: Thomas D. Mayfield III, The impact of social media on the nature of conflict, and a commander's strategy for social media (2010); - Dr. B. Forrester, Social Media Exploitation Tools: Understanding Where and How to Look, NATO, RTO-MP-HFM-201

3.1.14 Dissemination

Dissemination is the timely conveyance of intelligence, in an appropriate form and by any suitable means, to those who need it. Some forms of SM are coming in near-real time, an extra effort from intelligence is required to assess situations more quickly in certain cases to be able to withstand certain sentiments. Although SM is thought of as an open source, there are some restrictions in dissemination the results. Reasons could be not to reveal one's information position or one's intentions (the same reason as with directing SM sources, outsourcing). Sharing SM information could also be limited due to subscription limitations (contract, copyrights).

3.1.15 Example (near) real time tactical level

Table 3 showed typical examples of SM contributions towards the ICP regarding near-real time situation awareness at the tactical level. This information could be used as showed in this example to determine threats based on SM inputs.

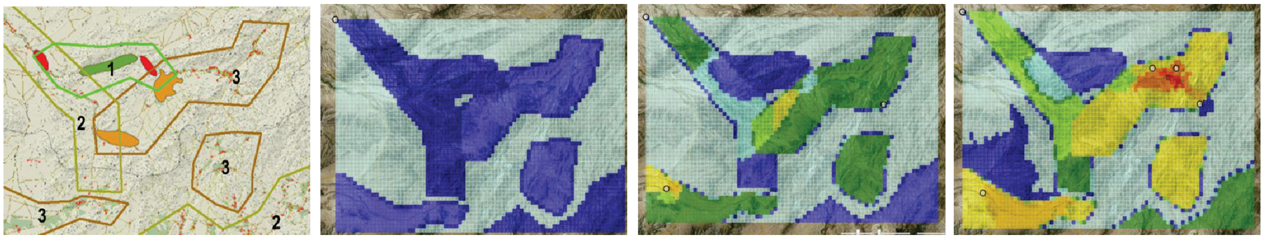


Figure 4 Threat developments extrapolated in time based on Social Media events

Figure 4 shows a visualisation of threat development in time regarding Social Unrest. The visualisation comes from a demonstrator NRTSA (Near Real Time Threat Situation Awareness (Ref. TNO 2016 R10171) to extrapolate ("forecast") different potential threats in time by using a Bayesian Belief Model, combining a-prior knowledge and (near) real time observations to support running tactical operations. One threat model deals with SM inputs and calculates the spread of effects in the Human Environment. In the screenshots an example is shown how populations (1, 2 and 3) are dispersed in an area of 100 x 200 km^2. The model takes into account a-prior knowledge and (near) real time observations about how populations are liaised to each other (positive, neutral, negative) and are developing in time. From left to right it can be seen what the effect is regarding the security situation. In the most right picture parts of the map are red, indicating a deteriorating situation. In case own forces are in that part of the area or are heading towards that area, commanders can be warned, so they will be able to take decisions with respect to this ("forecasted") situation.

Other sources provide information as well to determine threat levels, each producing their own layer. Combining the layers gives an impression of the overall threat level at a particular time and location.

4. CONCLUSION

While SM is a relatively new phenomenon there has been a huge uptake by society. More significantly, SM is being used with great success by terrorist groups for the purposes of fund raising, recruiting and garnering support for their causes. As a result, intelligent agencies have become increasingly interested in understanding and exploiting the openly available information provided by this source. However, this has been an uphill battle within many militaries that are by nature conservative and slow to adapt to new (civilian) technologies; SM has been no exception. In 2015 it seems that the tide has changed and much emphasis is now being made to understand this source and its potential uses.

This paper explored the nature of SM and looked at potential military intelligence uses. Briefly covered were some of the ways SM is being exploited by adversaries in modern conflicts. Such deceptive and influence uses are ideal exploits in a medium where anonymity is easy and the validation and verification of source is difficult. Section 3 concentrated on situating SM within the intelligence cycle. The cycle itself is not altered, but the analyst must allow for greater flexibility when including SM sources. The use of the Internet poses great challenges to collection and processing in particular. Protecting one's intent and collection patterns is very hard on the open Internet. Finally, SM represents a social sensor and hence requires careful interpretation of data; quantitative analysis is not enough, a social science perspective is required.

References

[1] Helen, B. and P. Benjamin, *Stop looking for the Next Twitter Revolution*, D.o.N. Intelligence, Editor 2010.

[2] ET.BY, *Technical Activity Proposal - Intelligence Exploitation of Social Media*, 2012, NATO RTO.

[3] Howard, P.N., A. Duffy, D. Freelon, M. Hussain, W. Mari, and M. Mazaid, *Opening Closed Regimes What Was the Role of Social Media During the Arab Spring?*, N.S. Foundation, Editor 2011, The Project on Information Technology and Political Islam: Washington.

[4] UN. *Freedom of the Press: in the Middle East, widely curtailed and often violated*. 2012 [cited 2013 11 Jan]; Available from: http://www.ohchr.org/EN/NewsEvents/Pages/FreedomofthePressintheMiddleEast.aspx.

[5] Estabrooke, I. and D.J.Y. Combs, *Social Media Defining the Problem: A Research Perspective*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, R. NATO, Editor 2012: Tallinn, Estonia.

[6] Labrèque, A., *Study of Social Networking Exploitation Tools*, B. Forrester, Editor 2011, Defence Research and Development Canada: Quebec City.

[7] Gayo-Avello, D., *A Balanced Survey on Election Prediction using Twitter Data.* arXiv, 2012.

[8] Choy, M., M. Cheong, M.N. Laik, and K.P. Shung, *US Presidential Election 2012 Prediction using Census Corrected Twitter Model*, 2012.

[9] Yu, S. and S. Kak, *A Survey of Prediction Using Social Media*, 2012, Oklahoma State University: Stillwater, Oklahoma.

[10] Leavitt, A., E. Burchard, D. Fisher, and S. Gilbert, *The Influentials: New Approaches for Analyzing Influence on Twitter*, in *Web Ecology Project*2009.

[11] Bongwon, S., H. Lichan, P. Peter, and H.C. Ed, *Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network*, in *Proceedings of the 2010 IEEE Second International Conference on Social Computing %@ 978-0-7695-4211-9*2010, IEEE Computer Society. p. 177-184.

[12] Jianshu, W., L. Ee-Peng, J. Jing, and H. Qi, *TwitterRank: finding topic-sensitive influential twitterers*, in *Proceedings of the third ACM international conference on Web search and data mining %@ 978-1-60558-889-6*2010, ACM: New York, New York, USA. p. 261-270.

[13] Tinati, R., L. Carr, W. Hall, and J. Bentwood. *Identifying Communicator Roles in Twitter*. in *WWW2012 - MSND'12 Workshop*. 2012. Lyon, France.

[14] Shakarian, P. and D. Paulo. *Large Social Networks can be Targeted for Viral Marketing with Small Seed Sets*. in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM-2012)*. 2012.

[15] Medoza, M., B. Poblete, and C. Castillo, *Twitter Under Crisis: Can we trust what we RT?*, in *1st Workshop on Social Media Analytics (SOMA'10)*2012: Washington, DC.

[16] Fink, C., J. Kopecky, and N. Bos, *Evaluating Social Media as a Source of Public Opinion in the Developing World*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, N. RTO, Editor 2012, RTO NATO: Tallinn, Estonia.

[17] Haewoon, K., L. Changhyun, P. Hosung, and M. Sue, *What is Twitter, a social network or a news media?*, in *Proceedings of the 19th international conference on World wide web %@ 978-1-60558-799-8*2010, ACM: Raleigh, North Carolina, USA. p. 591-600.

[18] Java, A., X. Song, T. Finin, and B. Tseng, *Why we twitter: understanding microblogging usage and communities*, in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*2007, ACM: San Jose, California. p. 56-65.

[19] Asur, S. and B.A. Huberman, *Predicting the Future with Social Media*, 2009, Social Computing Lab HP Labs: Palo Alto.

[20] Rao, D. and D. Yarowsky, *Detecting Latent User Properties in Social Media*, 2009.

[21] *The Next Generation in Open Source Intelligence*, 3i-MIND, Editor 2011.

[22] NATO, *Open Source Intelligence Gathering Handbook*, NATO, Editor 2001.

[23] Reynolds, W.N., M.S. Weber, R.M. Farber, C. Dorley, A.J. Cowell, and M. Gregory, *Social Media and Social Reality Theory, Evidence and Validation*, in *ISI 2010,* 2010, IEEE: Vancouver, BC.

[24] *Netlingo*. 2012 [cited 2011 18 December]; Available from: http://www.netlingo.com/acronyms.php.

[25] Agichtein, E., C. Castillo, D. Donato, A. Gionis, and G. Mishne, *Finding high-quality content in social media*, in *Proceedings of the international conference on Web search and web data mining* 2008, ACM: Palo Alto, California, USA. p. 183-194.

[26] Jolicoeur, P. and A. Seaboyer, *The Evolution of Russian Cyber Influence Activity: A Comparison of Russian Cyber Ops in Georgia (2008) and Ukraine (2014)*, DND, Editor 2014: Royal Military College of Canada.

[27] Berzins, J., *Russia's new generation warfare in Ukraine: Implications for Latvian defense policy*, C.f.S.a.S. Research, Editor 2014, National Defence Academy of Latvia. p. 15.

[28] Lever, R. *Social media a key element for terror groups: study*. 2014 [cited 2015 3 June]; Available from: http://phys.org/news/2014-05-social-media-key-element-terror.html.

[29] Eriksson, M., U. Franke, M. Granasen, and D. Lindahl, *Social media and ICT during the Arab Spring*, S. Defence, Editor 2013, FOI: Stockholm.

[30] Dewey, T., J. Kaden, M. Marks, S. Matsushima, and B. Zhu, *The impact of social media on social unrest in the Arab spring*, 2012, Stanford University: Stanford, CA.

[31] Forrester, B., A. Frini, and R. Lecocq. *Understanding the Role of Social Media in a Counter-Insurgency Context*. in *NATO IST-099 RSY-024 Emerged/Emerging "Disruptive" Technologies*. 2011. Madrid, Spain.