Application of a military data dissemination standard in a civil context

Daniel Haferkorn[a], Philipp Klotz[b], and Roland Rodenbeck[c]

[a,b,c]Fraunhofer IOSB, Fraunhoferstrasse 1, Karlsruhe, Germany

ABSTRACT

Nowadays, ever larger amounts of data are being generated, processed and linked. This enables to share data with other people or communities and to work collaboratively and evaluate data. Depending on the use case, environment and domain there are different aspects to consider regarding data security, availability, data protection etc.

In the military environment, a concept and derived specifications for data distribution were standardized as STANAG 4559 and are already used operationally. The advantages of such a solution can also be of interest for other domains with similar needs.

A possible use case is in the context of research data. Especially in areas where huge amounts of data with specific features are needed, it is often difficult to access (enough) research data and as a result, the outcome of the research is of limited quality. As every research institution creates its own data it would be helpful to have a possibility to share data and information amongst each other in a standardized way. The possibility of the aggregation from individual authorities results in a joint data pool. The research based on such a data pool can be more (cost) efficient, the quality increases due to the broader data sets and aspects like anomaly detection could be enforced. We present an idea of a concept to use a military data distribution standard for civil applications by defining data model extensions and considering security aspects and obstacles that may occur from various aspects such as the military characteristic and inflexibility of the standard and the data model.

Keywords: Interoperability, STANAG, Standard, Coalition Shared Data, Distribution, Data Sharing, Data Model

1. INTRODUCTION

Many different applications rely on the presence of useful data in meaningful amounts. Such data is usualy specific to the domain of the respective application. Some machine learning techniques especially rely on the presence of larger amounts of data and to be able to draw the right conclusion, it might be of interest to be able to make use of data produced externally. Depending on the specific use case it might even be of interest to be able to retrieve data from outside of the own domain. However the dissemination between partners should always be demand-oriented.

Over time, the amount of data that is generated or collected within various contexts has grown steadily. This is advantageous for a wide range of applications. For example, the processing of data with a larger number of surveys can potentially become more accurate. This would be the case, for example, if a measurement result was repeatedly confirmed in different measurement runs. Furthermore, larger amounts of data also allow easier assumption of general commonality. This would be the case if data are collected in different places and a statement on the basis of these data can be generalized, which is confirmed by the different collected data. By avoiding unnecessary redundant data collection, (cost) efficiency may be improved.

If artificial intelligence or machine learning is used to automatically work with the data, it is useful to be able to access as many different data as possible. These could be available as training data as well as test data.

Further author information:
Daniel Haferkorn: E-mail: daniel.haferkorn@iosb.fraunhofer.de
Philipp Klotz: E-mail: philipp.klotz@iosb.fraunhofer.de
Roland Rodenbeck E-Mail: roland.rodenbeck@iosb.fraunhofer.de

For the use of larger amounts of data, a meaningful and structured data distribution is necessary. If the data is to be shared with other potential users, a certain level of security is also required. Not all data can be shared with all possible users. This is the case in both civil and military environments. In the civil environment, the GDPR[1] should be mentioned here. This regulation requires that no personal data may be passed on for processing unless authorized. Other possible limitations for sharing of data include it-security concerns and intellectual property rights. These factors should already be taken into account at an early stage when planning the data dissemination process.

Other fields that rely on data processing are for example the fields of medicine or chemistry, where fundamental conclusions can be drawn from certain data. There are similar restrictions in the military environment. Not all generated data or collected data can be safely shared. See the work "Intelligence Information: Need-to-Know vs. Need-to-Share"[2] for more information about sharing data safely.

For a possible data distribution to different users, a concept is necessary, which covers all these aspects. This also comes into play, for example, when data is to be exchanged across domains. One example would be data that has been generated in a military environment but is to be made available for research purposes.

In the military environment, there are already data models and approaches that make such a security respecting distribution of data possible. Such an international military standard is the NATO STANAG 4559.[3] It describes the interfaces and data model for an information system that is capable of disseminating and distributing data within a coalition, including security labels and with strategies for bandwidth reduction. With this work, we want to show that such an application can also be used profitably in a civilian environment and that it brings advantages for all involved parties.

In the following section 2 we show related works, which have similar approaches. Afterwards in section 3 we first talk about the STANAG 4559 3.1 and provide a description of the CSD concept 3.2. We show possible use-cases in section 3.3. In section 3.4 and 3.5, we discuss the possible changes on a modification of the original standard. Future work and our conclusion are shown in section 4.

2. RELATED WORK

In their paper "Adaptation of interoperability standards for cross domain usage"[4] our colleagues discussed interoperability in the context of cross-domain data sharing. They describe and apply four levels of interoperability to both the civil and the military domain. These four levels of interoperability are technical, syntactic, semantic and pragmatic interoperability. They also discuss the benefits and a possible generic adaption of standards without a specific scenario, as well as an approach for cross-domain usage of applications and standards.

Additionally, in their paper "Interoperability of heterogeneous distributed systems"[5] our colleagues discussed aspects of interoperability within heterogenous distributed systems. They have addressed the different types of data and information. Among other things, they considered 'real time' and 'non-real time' data and the different formats that occur in data distribution. The main point is that the data in the discussed military context is very sensitive and limited in terms of distribution. Therefore it was examined how the right data can be made available to the right user. Thus, the use of an existing standard for the distribution of data was taken into account.

In their work "Collective data mining: A new perspective toward distributed data analysis",[6] the authors deal with how data from heterogeneous distributed systems can be successfully used for data mining without coming up with unjustified results because of too much oversimplification. They present a framework that allows a correct analysis of data and collection of local data models with minimal data communication.

Our colleagues addressed in their work "Evolution of the Coalition Shared Data concept in Joint ISR"[7] the origin of the data distribution that we mentioned here.

According to Peter Rhese and his work "CIMIC: Concepts, Definitions and Practice",[8] the cooperation of military and civil forces/authorities in various conflicts is becoming increasingly important, as the humanitarian missions in East Timor, Kosovo and Bosnia have shown us. He calls this cooperation Civil-Military Cooperation (CIMIC). In his work he describes and analyses such a cooperation between NATO and ICRC (International Committee of the Red Cross). Concepts, definition and practice are listed. This work shows that it is becoming

increasingly important to advance the cooperation between military and civilian applications for data exchange and so on.

The publication "The Influence of Noise, Vibration, Cycle Paths, and Period of Day on Stress Experienced by Cyclists"[9] is to be seen as an example for the necessity of large heterogeneous data sets, which need to cover a large area. In this work, the authors discuss the stress factor of cyclists in road traffic. The aim of the paper is to investigate the extent to which stressful situations prevent cyclists from using bicycles as means of transport. For this, data is collected at different locations. Using a heat map, the influence on the stress level can be investigated. In order to obtain a good and generalizable statement about possible stress factors, as many different situations and places as possible have to be covered.

In its publication on "Biodiversity Data Access",[10] the Scottish Wildlife Trust looks at how its collected data can be shared and what needs to be considered. Therefore, they issue a policy together with their work. This shows how useful it is to use metadata and security domain transfers so that only authorized data is actually distributed. Such data can be narrowed down by classification, as is often the case with intelligence information.

Our colleagues described a specific use case for a multi-national coordinated data collection, exploitation and dissemination in their "Tailored information provision for multinational naval operations".[11]

In our approach, we want to discuss ideas for possible scenarios as the basis for a civil application of a military standard.

3. OUR APPROACH

The foundation of our approach is the STANAG 4559 (described in the section 3.1). The STANAG 4559 specifies an interface, which enables access to so-called "Intelligence, Surveillance, Reconnaissance" (ISR) products. The CSD-Server was developed according to this standard and thus offers uniform access to the ISR products (described in the section 3.2).

3.1 Introduction to the STANAG 4559

In order to simplify cooperation between NATO nations as much as possible, various "Standardization Agreements" (STANAGs) have been defined. The aim of these agreements is to ensure that all NATO troops are equipped as uniformly as possible, thus enabling the exchange and procurement of all kinds of materials and equipment in larger and thus more cost-effective quantities. In addition to equipment, communication between NATO partners is also defined in STANAGs. Uniform radio frequencies, transmission methods for radio equipment and various rules for data sharing are just a few examples of standardization through STANAGs.

The STANAG 4559 describes the process of sharing NATO ISR products (imagery, video, reports, task, etc.) and other documents between NATO partners. If there are different systems from different vendors in a network a general concept needs to be agreed upon and implemented. Therefore, the standard defines common interfaces, use cases and processes to enable the basic data dissemination between the nodes. By following these basic definitions and agreements the interoperability inside a coalition network could be increased.

The output of projects or trials, change requests by the nations and further knowledge is constantly evolving the standard and pushing it forward. The current version of the standard is Edition 4 which is, in contrast to the previous edition (Edition 3 Amendment 2), divided into three documents - called the AEDPs (Allied Engineering Documentation Publication) 17,[12] 18[13] and 19.[14] These documents describe the three main parts and data types provided by the standard.

The complete topic covered by this standard is called the Coalition Shared Data (CSD) concept.

3.2 Description of the CSD concept

In recent years, the interest in surveillance tasks in both the civilian and the military environment has increased more and more. Surveillance of buildings (private and public), public places and public transport are common practice today. Also the military forces and the law enforcement authorities monitor their adversaries more and more by drone or with satellite pictures.[15] Certain collected data, however, are not only relevant for single

persons but may contribute to the solution of problems of others for whose purpose they were not actually collected. In addition, in combination with other information of higher value, data fusion may take place.

This collection of information usually takes place via separate systems and is therefore not centrally available for shared use. Thus it can occur that certain information is collected twice or even several times and thus a multitude of redundant data and information is generated. As a result, effective and successful data acquisition and data usage is not always possible.

In order to solve this problem, a concept for the demand-oriented distribution of data was developed and implemented in the military environment. The "Coalition Shared Data" - concept offers the storage of standardized data according to user requirements, the applicable network and security guidelines. The concept provides this securely and demand-oriented for different systems. A metadata concept is used to store further information (creation date, location, sensor type, etc.) in addition to the captured products (images, videos, etc.). The metadata generated by sensor, tasking and exploitation systems are collected and stored on the common data server or services. Connected clients thus have the option of accessing this data, making changes, supplementing existing data and entering new data.

The following figure 1 shows an exemplary structure of a CSD coalition consisting of four different networks. It is noticeable that not every subnetwork has to provide all components; it may only be a subset of them. For example, network D does not have its own data sources or data storage units, but only client systems for evaluating the content provided by the other networks. In addition, the connection of the individual networks should demonstrate that no fully meshed network is necessary to access content from other nodes. Network D can therefore also, for example, access the contents that have been stored in network C. The individual network parts are connected via so-called network domain gateways, which ensure that the data exchange between the networks can be carried out according to previously defined guidelines (e.g. security guidelines).

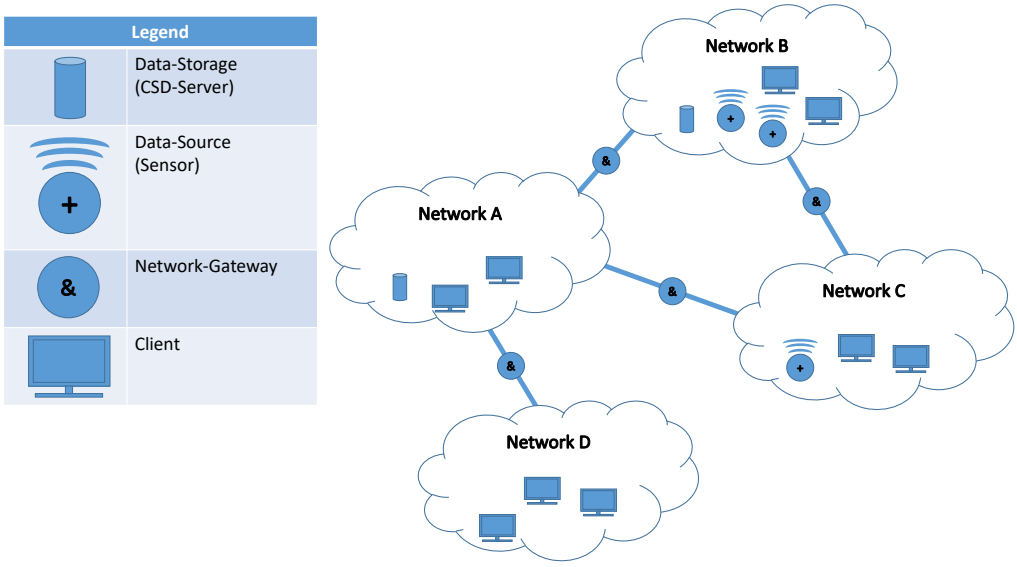

Figure 1. An Example of a Network

3.2.1 Basic functionality of the CSD-Server

Data is stored and distributed in accordance with STANAG 4559 AEDP-17.[12] This allows for a large number of data types to be stored, such as reports, videos and images. Furthermore query and subscription of CSD metadata is possible. Subscriptions make it possible to be automatically informed when new interesting data is available to the user. Synchronization of CSD metadata across multiple CSD-Servers is done over the WAN (Wide Area Network) according to bandwidth and task constraints, allowing accurate bandwidth adaptation to

the restrictions imposed by the environment. It is also scalable and in general adaptable to different network topologies. However our colleges discussed in their work[16] several limitations to the network topology.

Enhanced features and functions include support for CSD data model extensions. Validation is done by validating metadata, related files, and XML validation with schemas. The Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB) has developed an operationally usable implementation of the standard. In addition to the specification of the standard, this implementation contains a role-based user management with support of different access roles and security settings for each user. The Lightweight Directory Access Protocol (LDAP) can also be used to authenticate and authorize users. Data backup is covered by backup management and may be upgraded to hot standby functionality. All system status and security-related audit data can be logged and can be forwarded to a log server. An administrator may use graphical user interfaces to visualize the system status and track past events. It is also straightforward to make adjustments to system parameters and general configuration. With these extensions operational usability is improved.

Further possible extension to the CSD concept are:

- The integration of other reconnaissance types, such as HUMINT (Human Intelligence), SIGINT (Signals Intelligence), ELINT (Electronic Intelligence) und MASINT (Measurement and Signature Intelligence).
- The integration of other sensor types.
- Intelligence Surveillance Reconnaissance (ISR) using other services (SOA).
- Querying of semantic models.
- Usage of data fusion services.
- The support of cross-domain (e.g. high/low) communication using a security gateway.

3.3 Possible environments for this application

In the following section, we talk about the usage of the concept and the standard in different environments. For that matter, we consider different scenarios in a civil and civil-military domain (CIMIC).

3.3.1 Application in civil environments

The CSD concept has already proven itself in the military environment and is in operational use. Standardized data exchange between cooperation partners of different nations and their implementations is possible. Now large amounts of data can collected and made available to all partners.

A similar approach would be conceivable in a civil environment where data on a subject area is compiled and is needed by several parties.

Possible fields of application are:

- Joint acquisition and usage of measurement data (e.g. weather conditions, sea levels, earthquakes, etc.)
- Joint collection and usage of research results (e.g. data, information, etc.)
- Cataloguing of objects (e.g. natural history museum, archaeological data, etc.)
- Digital health records to transfer and share patient information between hospitals or doctors and patients (see[17])
- Digital school records
- Digital refugee records

Much of this can already be implemented (at least in part) with the CSD concept. Nevertheless, a number of adjustments have to be made. (See section 3.4 - Possible Adaptation and Necessary Adjustments)

3.3.2 Application in CIMIC environments

A possible scenario for a reasonable application of the concept in a CIMIC environment is disaster response. Other possible CIMIC environments are border control or coastguard duties where we also see this application as suitable. The advantage of the application is to provide the user with all the data he might need to fulfill his work in the most effective way.

In all kinds of examples for CIMIC environments, the "by-catch" data from military operations can be transferred to the civilian sector. It should be noted that the following considerations are only intended to represent initial thoughts and ideas in this context. However, in the future process, these still have to be considered further in regard to security concerns and general feasibility.

An example of such by-catch data could be satellite images from military satellites that are provided in lower resolution to solve problems in the civil sector. In the disaster example a new satellite image of the affected region could help to know where disaster response team is needed at most.

A military network and a civil network are usually not allowed to be connected directly as they are naturally networks of different classifications. By introducing a security gateway between for example the military and the civil network it can be ensured that no security-relevant information is transmitted.

3.4 Possible adaptions and necessary adjustments

As mentioned earlier, the CSD concept has already been implemented several times and is already in operational use. Through the use in different projects and scenarios, baseline knowledge could be acquired and problems already encountered in the use of the concept could be eliminated. By using an existing and established concept, certain basic problems and considerable development expenditure can be avoided and existing groundwork can be used. For this purpose, the intersection of functions between the existing solution and the new intended use has to be formed. Some functionalities can be completely or partially adopted, while others must be completely redesigned.

The metadata model forms the basis of what kind of data the CSD-Server holds. It defines additional information to describe the existing content. For example, information such as time and location of coverage, resolution, format, and more can be provided for a stored imagery product. It is also possible to extend the metadata model with extensions. Domain-specific attributes can thus be included in the metadata model and subsequently used.

The standard (STANAG 4559) provides a synchronization concept for distributing the catalog entries between the existing CSD nodes. Individual systems request new catalog entries from other systems and integrate them into their own collection. Using cross-domain transceivers, data exchange can even be ensured across domain borders through security gateways. This ensures that a shared data set can be accessed by all nodes in the entire network and that relevant information is also available across network boundaries. However, it must be ensured that certain security guidelines are adhered to at the junctions of two differently classified networks. Only those data may be exchanged between the networks which have the necessary security label and releaseability information. This concept would also be conceivable in a civilian or mixed civilian/military environment (see section 3.3.1 and 3.3.2). When distributing research results, it may be the case that not all data is allowed to be distributed to all cooperation partners in the network. If, for example, personal data are collected and processed, these may only be distributed if the distribution does not violate the GDPR.[1] Under certain circumstances, the distribution of data in a civilian/military environment must also be monitored and restricted. For example, a sharing of maritime data could be envisioned, where content is collected from both the coastguard and the navy. The navy would therefore have access to all data records, while the coast guard would have only a limited view of the records.

3.5 Standardization of data distribution

In addition to the standard (STANAG 4559) we mentioned, there are other possible standards for data distribution, such as OGC Sensor Web Enablement (SWE),[18] Electronic Business Registry Information Model (ebRIM, ISO 19135),[19] All Purpose Structured Eurocontrol Surveillance Information Exchange (ASTERIX).[20] However, since we have already gained experience in the handling and implementation of the standard, we have taken it as

the basis for our considerations. Furthermore, the military domain already demands certain security requirements which are already supported by the standard. In order to enable a preferably standardized data distribution of e.g. research and measurement data, a standard similar to the STANAG 4559 should be designed. This could also be based on the STANAG 4559 and extend it. The advantage of a distinct (civil) standard would be that it would be open and would be independent of possible further development of the STANAG 4559. However, it could be re-aligned again in case of developmental changes to the STANAG 4559.

4. CONCLUSION

In this paper, we have presented a theoretical approach in which an existing military standard is to be used for data distribution in a civil or mixed (CIMIC) environment. The basic idea of the approach was introduced and presented in the section 1, while various related works in the context of data distribution have been presented in section 2. In the context of data distribution using a military standard, the basic CSD concept and the STANAG 4559 standard were outlined. On the basis of this analysis new fields of application for a structured and standardized data distribution in the civil / mixed environment were discussed.

We have presented various possible ideas that show different approaches of a concept to use a military data distribution standard for civil applications. We assume that the STANAG 4559 cannot be adopted in full and that certain adaptations are necessary. Such an adaptation concerns above all the data model, which must be extended and adapted with domain-specific aspects. One essential aspect would be the definition of suitable metadata to enable relevant data distribution. It could also be an advantage to exchange data and information between different domains. We consider concepts such as security gateways to limit distribution as a feasible basis for a possible solution.

4.1 Future work

The preliminary work so far shows that it makes sense to continue the work on this. One possible approach is to first create concrete use cases from the presented ideas. These can then be used as the basis for creating a proof-of-concept.

For a possible implementation of the concept, concrete domains have to be considered and possible applications have to be defined. The corresponding domain-specific data models must be evaluated and transferred to the underlying model of the standard. Depending on the domain, a more or less complex mapping has to be developed. For a concrete adaption, a suitable scenario for a possible application should be specified and then implemented in practice.

REFERENCES

[1] Commission, E., "The General Data Protection Regulation." Data protection in the EU https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en (Feb. 2019). (Accessed: 19 February 2019).

[2] Best Jr, Richard A, [*Intelligence information: Need-to-know vs. need-to-share*], DIANE Publishing (2011).

[3] NATO Standardization Office (NSO), "STANAG 4559." https://nso.nato.int/nso/nsdd/stanagdetails.html?idCover=8838 (2018). (Accessed: 19 February 2019).

[4] Essendorfer, B., Kerth, C., and Zaschke, C., "Adaptation of interoperability standards for cross domain usage," in [*Next-Generation Analyst V*], **10207**, 102070E, International Society for Optics and Photonics (2017).

[5] Zaschke, C., Essendorfer, B., and Kerth, C., "Interoperability of heterogeneous distributed systems," in [*Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XV*], **9825**, 98250Q, International Society for Optics and Photonics (2016).

[6] Kargupta, H., Byung-Hoon, D. H., and Johnson, E., "Collective data mining: A new perspective toward distributed data analysis," in [*Advances in distributed and parallel knowledge discovery*], Citeseer (1999).

[7] Essendorfer, B., Kerth, C., and Zaschke, C., "Evolution of the Coalition Shared Data concept in Joint ISR," in [*Proceedings of the NATO IST/SET-126 Symposium on Information Fusion (Hard and Soft) for Intelligence, Surveillance & Reconnaissance (ISR), Joint Symposium IST-106 and SET-189*], (2015).

[8] Rehse, P., "Cimic: Concepts, definitions and practice," (2005).

[9] Nuñez, Javier and Teixeira, Inaian and Silva, Antônio and Zeile, Peter and Dekoninck, Luc and Botteldooren, Dick, "The Influence of Noise, Vibration, Cycle Paths, and Period of Day on Stress Experienced by Cyclists," *Sustainability* **10**(7), 2379 (2018).

[10] Trust, S. W. L., *Policy – Biodiversity Data Access* (Dec. 2005). Manual note.

[11] Müller, W., Reinert, F., Haferkorn, D., Essendorfer, B., Arecchi, A., Svensson, K., Rieter-Barelld, Y., and Ditzel, M., "Tailored information provision for multinational naval operations," To be published at the conference: SPIE Defense + Commercial Sensing (2019).

[12] NATO Standardization Office (NSO), "NATO STANDARD ISR LIBRARY INTERFACES AND SERVICES - AEDP-17." "https://nso.nato.int/nso/nsdd/apdetails.html?APNo=2272 (2018). (Accessed: 15 March 2019).

[13] NATO Standardization Office (NSO), "NATO STANDARD ISR Streaming SERVICES - AEDP-18." "https://nso.nato.int/nso/nsdd/apdetails.html?APNo=2273 (2018). (Accessed: 15 March 2019).

[14] NATO Standardization Office (NSO), "NATO STANDARD ISR WORKFLOW ARCHITECTURE - AEDP-19." "https://nso.nato.int/nso/nsdd/apdetails.html?APNo=2274 (2018). (Accessed: 15 March 2019).

[15] Rieger, F., "Das Gesicht unserer Gegner von morgen." Frankfurter Allgemeine Zeitung, 20 September 2012 http://www.faz.net/aktuell/feuilleton/debatten/krieg-mit-drohnen-das-gesicht-unserer-gegner-von-morgen-11897252.html (Sept. 2012). (Accessed: 11 February 2019).

[16] Kerth, C., Klotz, P., and Essendorfer, B., "A new approach for information dissemination in distributed JISR coalitions," To be published at the conference: SPIE Defense + Commercial Sensing (2019).

[17] Commission, E., "Commission makes it easier for citizens to access health data securely across borders." European Commission - Press release http://europa.eu/rapid/press-release_IP-19-842_en.htm (Feb. 2019). (Accessed: 28 February 2019).

[18] OGC, "Sensor Web Enablement (SWE)." opengeospatial http://www.opengeospatial.org/ogc/markets-technologies/swe (2016). (Accessed: 11 March 2016).

[19] OASIS, "ebXML Registry Information Model Version 3.0." OASIS http://docs.oasis-open.org/regrep/regrep-rim/v3.0/regrep-rim-3.0-os.pdf (May 2005). (Accessed: 11 March 2016).

[20] EUROCONTROL, "ASTERIX." EUROCONTROL http://www.eurocontrol.int/asterix (2016). (Accessed: 11 March 2016).