Ontology for the Intelligence Analyst

Barry Smith, University at Buffalo and National Center for Ontological Research

Tatiana Malyuta, Data Tactics Corp. and City University of New York

David Salmen, Data Tactics Corp.

William Mandrick, Data Tactics Corp.

Kesny Parent, Intelligence and Information Warfare Directorate

Shouvik Bardhan, High Performance Technologies, Incorporated

Jamie Johnson, EOIR Technologies

Abstract. As available intelligence data and information expand in both quantity and variety, new techniques must be deployed for search and analytics. One technique involves the semantic enhancement of data through the creation of what are called ontologies or controlled vocabularies. When multiple different bodies of heterogeneous data are tagged by means of terms from common ontologies, then these data become linked together in ways that allow more effective retrieval and integration. We describe a simple case study to show how these benefits are being achieved, and we describe our strategy for developing a suite of ontologies to serve the needs of the war-fighter in the ever more complex battlespace environments of the future.

New Demands for Intelligence Analysts

Intelligence analysts are trained to use their knowledge of available sources to enable querying across huge quantities of rapidly changing data. Already the richness and diversity of these sources makes it very difficult for human analysts, even with the most powerful software tools, to leverage their knowledge for analytic purposes. But their problems will only get worse. For while conventional intelligence processes have been focused primarily upon enemy units and on the effects of terrain and weather on military operations, new strategic guidance will require the intelligence community to focus also on disciplines such as cyberwarfare and civil information management [1, 2], and this will imply a massive expansion of the types of information relevant to analysis. The complex operations in which the warfighter of the future will be involved will require not only the mastery of vast quantities of network data but also information pertaining to the entire ecology of daily life in the areas of operation for asymmetric warfare, including information regarding religion, leadership, economics, culture, disease, food, water and other natural resources, and many more. All of this will go hand in hand with a vast expansion of the range of opportunities for the enemy to exploit weaknesses on the side of the warfighter—including weaknesses in our own understanding of this expanded environment of civil/military operations.

This increase in data diversity and volume, and in the velocity of change of data sources will pose an entirely new set of challenges for intelligence analysts, bringing the need for an approach to automated analytics that can solve the problem of rapid integration of heterogeneous and rapidly changing data in a way that can be reapplied in agile fashion to each new domain. This problem is analogous in some respects to the problem faced by warfighters of previous generations, who were attempting to develop the capability for massing timely and accurate artillery fires by dispersed batteries upon single targets. For massed fires to be possible dispersed artillery batteries needed the capacity for communication in real time of a sort that would create and sustain a common operational picture that could be constantly updated in light of new developments in the field. A way needed to be found, in other words, to transform dispersed batteries into a single system of what we might today call interoperable modules. The means to achieve this capability through a new type of governance and training, and through the creation of new doctrine in the field of artillery, were forged only in the early years of the last century at Ft. Sill, Oklahoma [3].

Today, we are facing the problem of *massing intelligence fires*—of bringing all relevant intelligence capabilities to bear on a target of interest in such a way that they, too, can serve as interoperable modules contributing to the development of a single shared evolving operational picture. In what follows we describe a strategy that is designed to address just one part of this problem—a strategy that is already being applied in the field to aid intelligence analysts working with a very large dynamic (cloud-based) data store to support operational decision-making [4]. The approach is of interest not least because it can be applied not merely to enhance existing data sources but also to build new representations in situ to serve analysts in the field.

Military Ontology

An ontology, in brief, is a set of terms and definitions representing the kinds and structures of entities and relations in some given area of reality. An ontology is thus comparable to a computerized dictionary. But it differs from a dictionary in being built around a logically robust classification of the entities in its domain, of a sort that can be used to enhance computer-based retrieval and integration of salient data.

The methods used today in ontology building include getting clear about what the types of entities are in a shared domain of interest, and also getting clear about the sorts of relations between these entities, methods which have been used by commanders and war-planners since the dawn of organized warfare in order to represent the tactical, operational, and strategic-level realities that make up the battlespace (see Figure 1).

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 2012		**00-00-2012 to 00-00-2012**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Ontology for the Intelligence Analyst	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University at Buffalo and National Center for Ontological Research,126 Park Hall,Buffalo,NY,14260**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

As available intelligence data and information expand in both quantity and variety, new techniques must be deployed for search and analytics. One technique involves the semantic enhancement of data through the creation of what are called ontologies or controlled vocabularies. When multiple different bodies of heterogeneous data are tagged by means of terms from common ontologies, then these data become linked together in ways that allow more effective retrieval and integration. We describe a simple case study to show how these benefits are being achieved, and we describe our strategy for developing a suite of ontologies to serve the needs of the war-fighter in the ever more complex battlespace environments of the future.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

The Strategy of Semantic Enhancement (SE)

In the data sources available to the analyst, multiple different terms, formats and data models are used to describe the data. The strategy of SE [6] is a response to the problems created by this diversity resting on the use of simple ontologies whose terms are used to tag (or 'annotate') source data artifacts in a consistent way. Ontologies built for SE purposes provide a restricted vocabulary that will enable analytics tools to see through the inconsistencies and redundancies in the data. This means: providing one term ('preferred label'), and one definition, for each salient type in each domain [7].

As illustrated in Table 1, the terms in an SE ontology are connected together in a simple hierarchy by means of the "is_a" (or subtype) relation. Each term appears only once in this hierarchy, and is associated in a stable way with its parent and child terms in the hierarchy even when new terms or even whole new branches are added to the ontology in the course of time. This stability is important, since the success of the strategy requires ontologies that can be repeatedly reused to annotate many different kinds of data in ways that then serve multiple different analyst communities and thereby contribute to the creation of an ever more comprehensive common operational picture. SE is thus designed to be at the same time more stable and more flexible than the traditional harmonization and integration approaches that, because they are typically based on ad hoc mappings amongst data models, often rapidly degrade in their effectiveness over time.

On the other hand, however, ontology is no panacea. Indeed, the increasing popularity of ontologies in the wake of the Semantic Web [8] has meant that ontologies, too, are now frequently being created in ad hoc fashion to address specific local data integration needs with little or no attention to the issues of consistency and stability. For SE to work, however, it is important that we find a way, through governance, training and doctrine, to counteract this tendency to ad hoc ontology development by bringing it about that a single evolving suite of consistent ontologies is created through the coordinated effort of multiple communities. Already the return on investment from the initial phase of the work described here has shown that such coordinated effort can bring significant benefits by making visible connections between data that had hitherto been walled off in separate siloes.

The Architectural Approach

To this end, the SE ontologies are organized on three levels, with successively greater degrees of flexibility:

- A single, small, domain-neutral Upper-level Ontology (ULO), for which our selected candidate is the Basic Formal Ontology [9].
- Mid-level Ontologies (MLOs), formed by grouping together terms relating to specific domains of warfare, or to specific tasks such as inter-agency information sharing [10].
- Low-level Ontologies (LLOs) focusing on specific domains, for example: EyeColor, HairColor, Name.

The terms used in these ontologies represent what is general or repeatable in reality at successively more specific levels. The level of an ontology is determined by the degree of generality of the types in reality which its nodes represent.

Figure 1: "Rakkasan" Commander Col. Luong issues an opening statement at the start of a sand table briefing. The pieces on the sand table are the result of an ontological process of categorization of the entities in the relevant domain [5].

⌊ = *is_a* (or subtype)

vehicle =def: an object used for transporting people or goods

⌊ **tractor** =def: a vehicle that is used for towing

⌊ **artillery tractor** =def: a tractor that is used to tow artillery pieces

⌊ **wheeled artillery tractor** =def: an artillery tractor that runs on wheels

⌊ **tracked artillery tractor** =def: an artillery tractor that runs on caterpillar track

Table 1: Examples of definitions used in SE ontologies.

The ULO is maximally general; it provides a high-level categorization relating to distinctions such as that between an object and a process, or between an object and its qualities (for example temperature), roles (for example, commander), and spatial locations.

The MLOs are general representations formulated using terms (such as database, person, organization) which will be needed by specific communities of SE users and developers.

At the bottom of the hierarchy are the LLOs, each representing some narrow homogeneous portion of reality. In the SE approach, the LLOs represent reality in such a way that:

1. For each salient domain, exactly one LLO is constructed that is in conformity with the settled science or military doctrine in that domain.
2. The LLOs are orthogonal (they do not share any terms in common).
3. They are designed to reduce the need for (typically fragile, and costly) mappings between ontologies covering the same or overlapping domains.
4. They are able to be used as reliable starting points for the development of cross-domain ontologies needed for all of intelligence and for specific areas of intelligence analysis.

Figure 2: Human Anatomical Property Ontology

An example SE LLO is illustrated in Figure 2. Other examples are:

- PersonName (with types: FirstName, LastName, Nickname, …)
- PersonIdentification (with types: SocialSecurityNumber, DriverLicenseNumber, …)
- PersonDate (with types: BirthDate, DeathDate, …)
- InformationProvenance (with types: Origin, Credibility, Confidence, …)
- Evidence (with types: ConfirmingEvidence, ContravertingEvidence, ...)

The SE approach is designed to be of maximal utility to intelligence analyst users of data. Ontology content is created only in response to identified situational needs of analysts, and architectural requirements are designed to ensure coherent evolution of the SE resource without sacrificing the flexibility and expressivity needed in actual deployment in the field. As more experience is gained using SE ontologies, intelligence analysts will uncover new ways to exploit the SE resource, and new groups of users will begin to see the benefits to be gained from developing their own complementary ontology resources in a way that is compliant with the SE architecture. Their data will then progressively become integrated with existing SE resources, bringing benefits through increase in the amount, variety and quality of data upon which intelligence analysts can draw [11]. In this way—following a pattern that has been realized already in biology and other domains [12]—the SE strategy will engender collaborative ontology development and re-use over multiple data collection endeavors, both internal and external.

The Discipline of Intelligence Analysis

Joint doctrine [13] defines multiple hierarchically organized disciplines, for example, intelligence, information operations, cyberspace operations; the discipline of Intelligence in its turn has doctrinally defined sub-disciplines such as Human Intelligence (HUMINT), Signals Intelligence (SIGINT), and imagery intelligence [14].

On the typical approach to intelligence analysis, each new set of analytical problems rests on its own collection of data sources, which must be identified and integrated in ad hoc fashion through manual effort by the analyst. A typical analyst may be working with some 100s of data sources, with each source coming from a particular discipline such as HUMINT or Geospatial Intelligence (GEOINT). For an analyst to come to a conclusion or decision, he has to verify each particular piece of information in 3 distinct disciplines. For example, if a GEOINT source says that location X is 'bad', then there has to be something in, say, a HUMINT and a SIGINT source that confirms this statement.

Already here we see the vital need for integration of heterogeneous data for purposes of intelligence analysis. The SE approach has evolved in response to the general recognition that traditional approaches to such integration, both physical and virtual, are increasingly failing in the face of the scale, diversity, and heterogeneity of many data sources and data models. Such traditional approaches fail where they do not address the following requirements:

- Integration must occur without the need for heavy pre-processing of the data artifacts which need to be integrated.
- Integration must occur without loss or distortion of data.
- The integration approach must be able to evolve to accommodate highly heterogeneous and rapidly evolving data.

Already the tagging of intelligence data in consistent fashion by drawing on a simple ontology for describing the different kinds of sources brings benefits to the analyst in a way that meets all of these requirements.

Case Study Illustrating the Benefits Brought by SE to Intelligence Analysis

In what follows we illustrate how these benefits are realized in terms of a simple case study in which the SE approach is applied to a set of cloud-based data sources, including text, images, audio, and signals, as described in [3]. These data sources are stored together with structured descriptions of their associ-

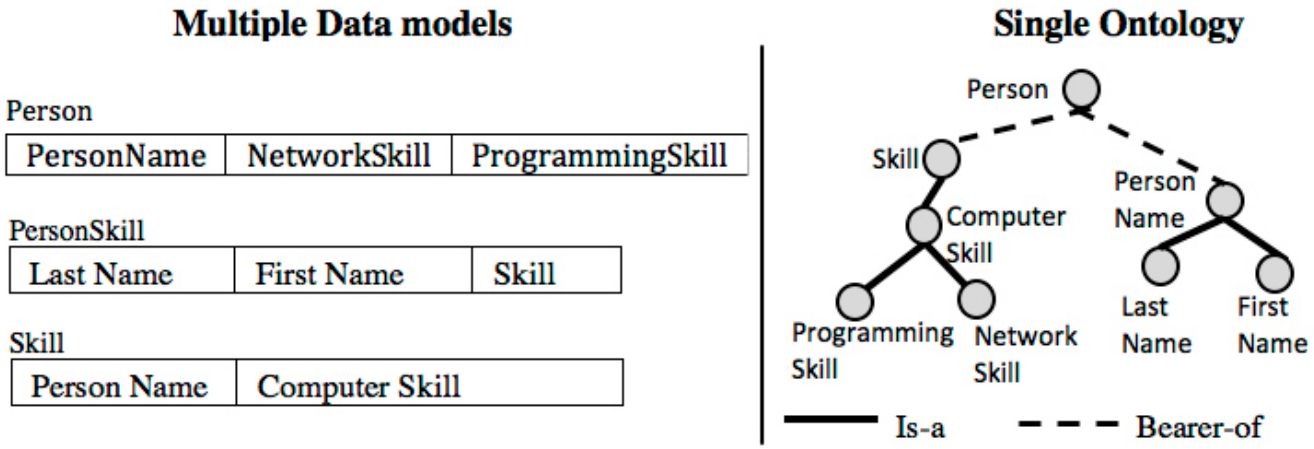


Figure 3: Samples of data models, in which arbitrary combinations are allowed (LEFT), vs. SE ontologies, with their constrained hierarchies (RIGHT)

ated data models. The problem that SE is designed to solve arises because different data models can present data about the same entities in arbitrarily many different ways, as illustrated on the left of Figure 3. The SE ontology content illustrated on the right, in contrast, employs simple terms in a stable fashion to ensure that entities of the same types are represented always in the same way.

SE terms are associated with the labels used in the native data model descriptions, as in Tables 2 and 4. To enable benefits from this association in the form of efficient search, the entire aggregated content of our data sources, both structured and unstructured, is indexed, using a Lucene index [15] distributed over Solr [16]. This Index, which is continuously being re-created to ensure synchronization with newly posted data, is a result of pre-materialization; that is, it reflects pre-calculations of the answers to sets of the most common queries posted by analysts.

We consider a simplified example using three native data sources, Db1-3, which we illustrate in each case by column labels and a single row of sample data. To see the sorts of problems we face compare how, in Db1, 'Java' is used elliptically to mean 'Java programming skill', while 'Name' is used to mean 'Name of skill'.

**Source database Db1, with tables Db1.Person and Db1.Skill, containing person data and data pertaining to skills of different kinds, respectively.*

PersonID	**SkillID**
111	222

SkillID	**Name**	**Description**
222	Java	Programing

* *Source database Db2.Person, containing data about IT personnel and their skills:*

ID	SkillDescr
333	SQL

* *Source database Db3.ProgrSkill, containing data about programmers' skills:*

EmpIID	SkillName
444	Java

Second, we use SE ontologies as illustrated in Figure 3 to annotate the data from these databases. Sample results of this annotation are illustrated in Tables 2-4, which are representative of the kinds of tables contained in our aggregated store.

Table 2 contains sample labels used in annotations. The rows of Table 3 represent sample annotations using SE ontology terms. The rows of Table 4 consist of sample statements of the sorts used both in storing native data and in generating the Index.

Label	Source
PersonID	Db1.Person
SkillID	Db1.Skill
Name	Db1.Skill
Description	Db1.Skill
ID	Db2.Person
SkillDescr	Db2.Person
EmpIID	Db3.ProgrSkill
SkillName	Db3.ProgrSkill

Table 2. Sample labels from source data models and from the SE ontologies

Source Artifact Label	SE Label
Db1.Name	SE.Skill
Db2.SkillDescr	SE.ComputerSkill
Db3.SkillName	SE.ProgrammingSkill
Db1.PersonID	SE.PersonID
Db2.ID	SE.PersonID
Db3.EmpIID	SE.PersonID

Table 3. Sample annotations of labels in source artifacts

To begin to see the benefits of SE for data integration, note how three distinct items in the first column of Table 3—PersonID from Db1, ID from Db2, and EmpIID from Db3—are all annotated with the same SE expression, namely PersonID from the PersonIdentification LLO.

Data Value and Associated Label	Relation	Data Value and Associated Label
111, Db1.PersonID	Db1.hasSkillID	222, Db1.SkillID
222, Db1.SkillID	Db1.hasName	Java, Db1.Name
222, Db1.SkillID	Db1.hasDescription	Programming, Db1.Description
333, Db2.ID	Db2.hasSkillDescr	SQL, Db2.SkillDescr
444, Db3.EmpIID	Db3.hasSkillName	Java, Db3.SkillName

Table 4. Statements illustrating the sorts of source data used in compiling the Index

The process of annotation proceeds manually as follows. The annotator is required to apply to each label in the target data model the term at the lowest level in the SE hierarchy whose application is still warranted (1) by the meaning of the label and (2) by information the annotator has about the database in question, including (3) information concerning the data values labeled. For example, Db1 contains data about skills in many areas; its label Skill must therefore be annotated with the general term Skill and not with any more specific term. Db2 is known to contain only data about skills in the area of IT; this warrants the use of ComputerSkill in annotating its label SkillDescr.

The Index contains entries of various sorts, as represented in Table 4. Which sorts of entities we index is determined by the ontologies for Person, Place, and so on. The subservient LLOs, which provide the SE labels to be used in annotations for different sorts of data, are used in formulating the field value pairs associated with Index entries.

Currently, the SE Index incorporates the results of inferences over an initial tranche of semantically enhanced content. In Table 5 we see how the Index looks when it is able to incorporate the results of integration over the SE annotations. These inferences rest on the logical structure of the SE ontologies and of their constituent definitions. For example, the term Programmer is defined as Person with programming skill and the Skill LLO incorporates an inferred subclassification of persons, which is represented in the Index using the Subtype field (see the entry for PersonID=444 in Table 5).

When creating the Index, the indexing process crawls statements of the sorts shown in Table 4 and uses SE labels for the Index fields wherever these are available. Thus, as Table 5 illustrates, we obtain fields carrying terms from the LLO Skill and LLO PersonIdentification, as follows:

Index Entry	Associated Field-Value
111, PersonID	Type: Person
	Skill: Java
	Db1.Description:Programming
333, PersonID	Type: Person
	ComputerSkill: SQL
444, PersonID	Type: Person
	SubType: Programmer
	ProgrammingSkill: Java

Table 5. Sample Entries of the Dataspace Index based on the SE

Some native content is not (or not yet) covered by the SE (the Description label from Db1.Skill in our example), reflecting the incremental nature of the SE process. Indexing in such cases is effected using native labels. In this way, incomplete SE coverage of native models does not entail unavailability of the corresponding data to analysts' searches.

A Sample Query Illustrating the Advantages Brought by SE

Suppose the analyst needs to use the Index in order to find, for example, all instances of the type Person referenced in the Dataspace as having some predefined set of skills. When addressed to the sample entries in Table 5, this will yield results as in Table 6.

To see the advantages that have been brought to the human analyst by the SE process, contrast now Table 7, which shows Index entries corresponding to those of Table 5 as they would have been generated prior to SE. Table 7 reveals two sorts of obstacles faced by the analyst using pre-SE data. First: because person IDs and names of skills in the native sources are listed under many different headings, querying these sources without SE, even for simple person ID or skill information, requires knowledge on the part of the analyst of the idiosyncrasies of each data source. Second: because data models are flat, in the sense that they do not define hierarchical relations between more general and more specific types, querying across sources that contain data at different levels of detail is virtually impossible.

Indeed, however much manual effort the analyst is able to apply in performing search supported by the Index entries illustrated in Table 7, the information he will gain will still be meager in comparison with what is made available through Table 5. Even if an analyst is familiar with the labels used in Db1, for example, and is thus in a position to enter Name = Java, his query will still return only: person 111. Directly salient Db4 information will thus be missed.

Conclusion

Analysts are of course trained to be aware of the types of information that are available in different sources. But in today's dynamic environment, in which ever more domains and ever more associated data sources become salient to intelligence analysis, it is practically impossible for any analyst to know the content of all sources. The likelihood that important data will be missed remains very high, and the need for agile support for retrieval and integration of the sort provided through the strategy of semantic enhancement becomes all the more urgent. This strategy was designed, in effect, to remedy some of the consequences of the inevitable lack of coordination in the development of information resources in the intelligence domain, and thereby to support massed informatics fires against ever-new types of intelligence targets.

Acknowledgements:

This work was funded by U.S. Army CERDEC Intelligence and Information Warfare Directorate (I2WD). The authors wish to thank Peter Morosoff of Electronic Mapping Systems, Inc., H. Shubert of Potomac Fusion, Inc. and C. Rosenwirth of MITRE for their contributions to the content of the SE and for helping us to understand the discipline of intelligence analysis. We also thank Dr. Guha and S. Patel of I2WD for their support and guidance on advancing this body of work.

- entering `Skill = Java` (which will be re-written at run time as: `Skill = Java` OR `ComputerSkill = Java` OR `ProgrammingSkill = Java` OR `NetworkSkill = Java`) will return: persons 111 and 444
- entering `ComputerSkill = Java` OR `ComputerSkill = SQL` will return: persons 333 and 444
- entering `ProgrammingSkill = Java` will return: person 444
- entering `Description = Programming` will return: person 111
- entering `SubType = Programmer` will return: person 444

Table 6: Sample queries over the Dataspace Index and their results with SE

Index Entry	Associated Field-Value
111, PersonID	Type: Person
	Name: Java
	Description: Programming
333, ID	Type: Person
	SkillDescr: SQL
444, EmpIID	Type: Person
	SkillName: Java

Table 7. Sample Entries of the Dataspace Index prior to SE

ABOUT THE AUTHORS

Dr. Barry Smith is a prominent contributor to both theoretical and applied research in ontology. He is the author of some 500 publications on ontology and related topics, with a primary focus on biomedical and defense and security informatics. He is director of the National Center for Ontological Research and University at Buffalo Distinguished Professor.

E-mail: phismith@buffalo.edu

Dr. Tatiana Malyuta is a Principal Data Architect and Researcher of Data Tactics Corporation and an Associate Professor of the New York College of Technology of CUNY. She is a subject matter expert in data design and data integration. Recently she has been working on integrated data stores on the Cloud. She received a Master's Degree in Applied Mathematics and a Ph.D. Degree in Computer Science from the State Polytechnic University in Lviv, Ukraine.

E-mail: tmalyuta@data-tactics.com

Dave Salmen is the Chief Technology Officer of Data Tactics Corporation, armed with over 20 years of extensive experience with full life cycle database system development with an emphasis on initiatives involving intelligence data. His recent work includes DCGS SIPR data cloud (Rainmaker), Information Integration Pilot (I2P), and Zones of Protection (ZoP). He has experience with cloud architecture, cloud data structure design, high volume data ingest, cloud deployment, and cloud security work.

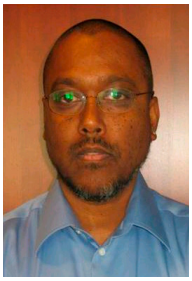
E-mail: dsalmen@data-tactics.com

Dr. Bill Mandrick is a Senior Ontologist at Data Tactics Corporation and an Adjunct Professor at the University at Buffalo. He is also a Lieutenant Colonel in the U.S. Army Reserves with deployments to Iraq and Afghanistan where he has commanded soldiers, planned for major operations, and served as the primary civil-military operations advisor to a Brigade Combat Team. Recently he has been working on intelligence related ontologies for the Intelligence and Information Warfare Directorate (I2WD).

E-mail: william.mandrick@us.army.mil

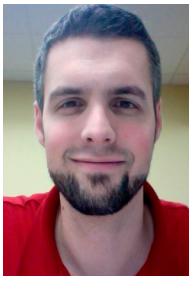
Kesny Parent is a Branch Chief in the Intelligence Information Warfare Directorate (I2WD) at the Communications-Electronics Research, Development and Engineering Center (CERDEC). He has worked in the Intelligence, Surveillance, and Reconnaissance (ISR) domain since 1989. He leads the Development and Integration for the DCGS-A Standard Cloud (DSC) project, a major Army initiative to integrate Cloud Computing Intelligence infrastructure across the entire Intelligence Community. In this capacity, he directed the design, development, and fielding of a highly complex cloud computing architecture with tools that greatly enhance the capabilities available to soldiers.

E-mail: kesny.parent@us.army.mil



Shouvik Bardhan has more than 25 years of experience in the field of complex software design and development and continues to be a hands-on developer on J2EE/PKI/Hadoop based enterprise software. He has managed, architected and delivered systems ranging from FISMA based Certification and Accreditation automation, supply chain management and financial applications to identity federation and document control. Most recently he has worked on U.S. Army's cloud project where as a part of the core development team he design and develops software for an Ultra Large Scale (ULS) Cloud computing environment. He holds a BS and MS in Computer Science (MS from Johns Hopkins University, MD) and is a Ph.D. student in the department of Computer Science in George Mason University, Fairfax, VA.

E-mail: sbardhan@drc.com



Mr. Jamie Johnson, is a Software Developer at EOIR Technologies. He has worked with the Intelligence Community for the past eight years as a Department of Defense Civilian Employee and as a Civilian Contractor. Most recently he has worked on cloud scale search and indexing technologies for the DCGS-A Standard Clouds integrated data store. He received a Masters in Computer Engineering from Stevens Institute of Technology and a Bachelor's Degree in Computer Engineering from Rutgers University.

E-mail: jjohnson@eoir.com

REFERENCES

1. Publication 2-01 Joint and National Intelligence Support to Military Operations, Chairman of the Joint Chiefs of Staff. Washington, DC. 05 January 2012: <http://www.dtic.mil/doctrine/new_pubs/jp2_01.pdf>
2. Strategic Guidance Document, Sustaining U.S. Global Leadership: Priorities for 21st Century Defense, Secretary of Defense. Washington DC. 05 January 2012: <http://www.defense.gov/news/Defense_Strategic_Guidance.pdf>
3. Boyd L. Dastrup, Cedat Fortuna Peritis: A History of the Field Artillery School, Combat Studies Institute Press, US Army Combined Arms Center, Fort Leavenworth, Kansas
4. Distributed Common Ground System - Army (DCGS-A), from 2011 Army Posture Statement, <https://secureweb2.hqda.pentagon.mil/VDAS_ArmyPosture Statement/2011/information_papers/PostedDocument.asp?id=151>
5. For more examples of the role of ontology in the history of military decision-making see <http://militaryontology.com/>.
6. David Salmen, Tatiana Malyuta, Alan Hansen, Shaun Cronen, Barry Smith, "Integration of Intelligence Data through Semantic Enhancement", Proceedings of the Conference on Semantic Technology in Intelligence, Defense and Security (STIDS), George Mason University, Fairfax, VA, November 16-17, 2011, CEUR, Vol. 808, 6-13.
7. Here 'type' is used to refer to what is general in reality (thus: military unit, vehicle, monsoon, headgear, and so on), as contrasted with particular instances (this military unit, that vehicle, last season's monsoon, Haneef's keffiyeh, and so on).
8. Tim Berners-Lee, James Hendler and Ora Lassila, "The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities", Scientific American Magazine, May 2001.
9. <http://ifomis.org/bfo.>
10. Barry Smith, Lowell Vizenor and James Schoening, "Universal Core Semantic Layer", Ontology for the Intelligence Community, Proceedings of the Third OIC Conference, George Mason University, Fairfax, VA, October 2009, CEUR Workshop Proceedings, vol. 555.
11. W. Brian Arthur, Increasing Returns and Path Dependence in the Economy, Ann Arbor, University of Michigan Press, 1994.
12. Barry Smith, et al., "The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration", Nature Biotechnology, 25 (11), November 2007, 1251-1255.
13. Joint Publication 1. Doctrine for the Armed Forces of the United States, Chairman of the Joint Chiefs of Staff. Washington, DC. 20 March 2009. <http://www.dtic.mil/doctrine/new_pubs/jp1.pdf>
14. Joint Publication 2-0 Joint Intelligence, Chairman of the Joint Chiefs of Staff. Washington, DC. 22 June 2007 <http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf>
15. <http://lucene.apache.org/java/docs/index.html>
16. <http://lucene.apache.org/solr/>