CERIAS Tech Report 2002-55
Why NLP Should Move into IAS
by Mikhail J. Atallah
Center for Education and Research
Information Assurance and Security
Purdue University, West Lafayette, IN 47907-2086

WHY NLP SHOULD MOVE INTO IAS

Victor RASKIN
CERIAS and NLP Lab, Purdue University
1356 Heavilon Hall 324
W. Lafayette, IN 47907-1356 USA
vraskin@purdue.edu

Sergei NIRENBURG
CRL, New Mexico State University
286 New Science Building
Las Cruces, NM 88003 USA
sergei@crl.nmsu.edu

Mikhail J. ATALLAH
CERIAS and Department of Computer Science
Purdue University
1315 Recitation Hall 215
W. Lafayette, IN 47907-1315 USA
mja@cs.purdue.edu

Christian F. HEMPELMANN
CERIAS and NLP Lab,
Purdue University
1356 Heavilon Hall 324
W. Lafayette, IN 47907-1356 USA
hempelma@purdue.edu

Katrina E. TRIEZENBERG
CERIAS and NLP Lab,
Purdue University
1356 Heavilon Hall 324
W. Lafayette, IN 47907-1356 USA
kattriez@purdue.edu

Abstract

The paper introduces the ways in which methods and resources of natural language processing (NLP) can be fruitfully employed in the domain of information assurance and security (IAS). IAS may soon claim a very prominent status both conceptually and in terms of future funding for NLP, alongside or even instead of established applications, such as machine translation (MT). After a brief summary of theoretical premises of NLP in general and of ontological semantics as a specific approach to NLP developed and/or practiced by the authors, the paper reports on the interaction between NLP and IAS through brief discussions of some implemented and planned NLP-enhanced IAS systems at the Center for Education and Research in Information Assurance and Security (CERIAS). The rest of the paper deals with the milestones and challenges in the future interaction between NLP and IAS as well as the role of a representational, meaning-based NLP approach in that future.

1 Introduction

With new applications, NLP sees new challenges and has to develop additional functionalities. For a few decades, it was driven predominantly, if not exclusively, by MT. This application, while emphasizing certain functionalities, has a limited use for a reasoning functionality. Increasingly, the current applications, such as data mining and question answering bring reasoning to the front of NLP. Applications come, for the most part, from real life, and in real life, computer systems keep getting attacked by hackers and industrial or political adversaries and need to be protected with the help of automatic systems. Information security provides this protection by preventing unauthorized use and detecting intrusions. Information assurance guarantees the authenticity of transmitted and stored information. In the last five years, since the inception of CERIAS with the help of a massive grant from the Eli Lilly Foundation, two of the co-authors have led a pioneering effort in exploring the possibility of applying the methods and resources of NLP to IAS. Another co-author has led a decade-long effort in developing the resources of ontological semantics and testing them in various implementations of NLP applications. This paper is the result of all these efforts as well as of the excellent work of the participating and actively contributing graduate and undergraduate research assistants.

2 Basic Premises

Nirenburg and Raskin (2002) views NLP as an application of both linguistics and cognitive science. This application is a theory of itself, which defines the format of its descriptions, e.g.,

meaning representations for texts (TMRs). The theory is associated with methodologies to produce these descriptions. Applications tend to dictate the content of the descriptions they need in order to be successfully implemented and thus, to a large extent, the methodology of implementation, which is, thus, arrived at systematically and not by just trial and error and guesswork, as Chomskian linguistics would have us believe.

In general, one of the choices in NLP is the method-driven vs. the problem-driven approach. The former espouses the use of a particular method in as many applications as possible. The danger here is that both the applications and the level of results that is declared satisfactory are molded to what is allowed by the method: "To a hammer, everything looks like a nail."

Problem-oriented NLP chains back from the needs of an application and happily accepts eclectic or pipelined approaches if this arrangement promises better results.

We approach IAS from the problem-oriented point of view. It is a growing family of applications that society needs to protect its computer systems and databases from unauthorized use and destructive attacks. It is the goal of NLP to serve the existing IAS needs as well as helping the IAS community to discover new ways to adapt the existing NLP resources and to order the development of new resources.

3 NLP Applications to IAS

3.1 IAS Needs

Most generally, IAS develops software to:

- encrypt and decrypt data;
- preclude unauthorized use of computer systems and data with a vast array of protective measures;
- detect intrusion, including virus recognition and anti-virus protection.

Much of IAS deals with signals and information other than texts in natural language (NL) but there are enough applications for textual data, and this is where the methods and resources of NLP come into the picture.

3.2 NLP/IAS Interface

CERIAS has taken a leading role in investigating how NLP can be utilized for IAS, and the initial efforts, as early as 1998, were devoted to identifying the text-based subtasks in IAS. To date, the following applications have been recognized and addressed, in chronological order:

- using machine translation for an additional layer of encryption;
- generating mnemonics for random-generated passwords;
- declassification or downgrading of classified information;
- NL watermarking;
- preventing theft of intellectual property;
- forensic IAS, specifically, tracing leaks in divulging protected information;
- tamperproofing textual data;
- enhancing the acceptance of IAS products by the users with the help of computational humor.

In the rest of the section, we will characterize these tasks briefly, with an emphasis on the NLP contribution to their solution, a contribution which is largely constitutive in nature in the sense that they would probably not exist if NLP could not offer the know-how to implement them.

3.2.1 MT for Encryption

Inspired by the most obvious connection between encryption and NL, the largely apocryphal World War II episode, when instead of an elaborate code, the American and British General Headquarters in Europe used the native speakers of Navajo (Shawnee, in another version, involving the Pacific theater) to communicate in open, uncoded language and were never "decoded," the idea was to use a family of existing or rapidly deployable MT systems (see Nirenburg and Raskin 1998) to add a level of encryption in an "exotic" language.

Once proposed (Raskin *et al.* 2001), the idea failed to catch and has never been implemented, partially because there was no research challenge in that, but also because it would involve the "security by obscurity" principle disdained by IAS: one should assume that the adversary is at least as smart and knowledgeable as we, the good guys, are. Also, an MT system, even if publicly available, is too long and messy a "key," another IAS no-no.

3.2.2 Mnemonics for Random-Generated Passwords

Passwords are sometimes dismissed in IAS as too weak and ineffective a protection measure. Reality is, however, that for an absolute majority of computer users, this remains the only protection against unauthorized use and abuse and the loss of data, and the users weaken it considerably by changing the passwords randomly generated for them by the computer at the time the accounts are created to something that is easy for them to remember. The weakness of such passwords is that they can be vulnerable to a brute-force attack because the space of possible passwords to be tried by the attacker becomes much smaller than that for random-generated ones. Here and elsewhere, IAS measures hardly ever exclude the possibility of a successful attack (e.g., using a random generator to try every possible alphanumeric combination to access the account) but rather "raising the ante" for the adversary, making the attack costlier and more complicated.

We implemented Versions 1 and 2 of the automatic mnemonic text (jingle) generator (AMTG). Both versions take a randomly generated alphanumeric password as input and generate a funny and memorable two-line text (jingle). AMTG-1 implemented after the first 6 months of research limited the input to 8-letter (no digits) case-insensitive passwords and generated a rigidly formatted, uniform-meter, single-tune jingles, whose funniness depended on the verb antonymy between the first and second lines (here and throughout this section, see Raskin *et al*. 2001 for examples and further discussion). AMTG-2 removes the rigid limitation on the password format and accepts 3-8-symbol alphanumeric, case-sensitive input while generating two lines of purported political satire (see McDonough 2000). The proof-of-concept software was implemented by McDonough and is in preparation for patenting.

3.2.3 Natural Language Downgrading

Increasingly, in interagency exchanges in the government, international coalition communication, and exchanges among business partners, there has been a need to develop an intricate architecture for combining a "high" network and a "low" network. Authorized users, with access to the high network, where sensitive data is stored and exchanged, must have access to the low network, but not the other way around. If this is all there is to it, the communication between the two networks is assured with the help of a variety of switches and one-way filters: the low-network information can propagate up but the high-network information must not leak down. There are enough technical and conceptual problems with such one-way filters, but they are multiplied manifold if there is also a need to share some high-network information with the low-network users in a way that removes all the sensitive data. In this context the essentially semantic ability to recognize a sensitive message comes into play. We are focusing only on sanitizing textual information. In other words, for each classified text *T* there must be generated a sanitized, downgraded text *T'*, from which all sensitive data are removed according to a certain list of criteria. We are doing this by utilizing the NLP resources developed by the ontological-semantic approach (Nirenburg and Raskin 2002), which allows deep-meaning penetration and, as a result, much enhanced sensitive information detection and removal (see Mohamed 2001) than that allowed by any keyword-based approach, straightforward or statistical.

3.2.4 Intellectual Property Protection

Essentially the same methods of detection and seamless replacement developed for downgrading can be used to intercept and prevent deliberate or inadvertent divulging of proprietary and/or classified information. This is much easier to do offline, of course, but there is also an increasing need in inconspicuous

interception and sanitizing of e-mail online. Here, somewhat less than in straightforward downgrading, which can all be done offline, a half-way solution may be best: instead of letting the system detect the sensitive information and replace it, all fully automatically, a simpler and coarser-grain-sized system can only flag possible violations to a human, who makes the final determination.

3.2.5 Natural Language Watermarking

We have developed software capable of embedding a hidden textual watermark in a textual message without changing the meaning of the text at all and the wording only slightly if necessary. Let *T* be a NL text, and let *W* be a string that is much shorter than *T*. We wish to generate NL text *T'* such that: *T'* has essentially the same meaning as *T*; *T'* contains *W* as a secret watermark, and the presence of *W* would hold up in court if revealed (e.g., *W* could say, "This is the Property of *X*, and was licensed to *Y* on date *Z*"); the watermark *W* is not readable from *T'* without knowledge of the secret key that was used to introduce *W*; for someone who knows the secret key, *W* can be obtained from *T'* without knowledge of *T* (so there is no need to permanently store the original, non-watermarked copy of copyrighted material); unless someone knows the secret key, *W* is difficult to remove from *T'* without drastically changing the meaning of *T'*; the process by which *W* is introduced into *T* to obtain *T'* is not secret, rather, it is the secret key that gives the scheme its security. We developed a technique (Atallah *et al*. 2001, 2002) which embeds portions of *W*'s bitstring in the underlying syntactic and semantic (TMR) structures, respectively, of a selection of sentences in a text by manipulating those sentences slightly with the help of meaning-preserving syntactic and semantic information. The semantic technique is much more complex and allows for a much wider bandwidth, i.e., the use of much fewer watermark bearing sentences, thus making the later technique usable for such short sentences as wire agency releases. It also furthers that advantage by making it unnecessary to double the number of engaged and manipulated sentences and disposing of the marker-bearing sentences that precede each watermark-bearing sentence in the earlier, syntactic approach.

3.2.6 Tracing the Leaks

By embedding different, personalized watermarks in different copies of the same document, we can trace a leak to a particular recipient of classified or proprietary information. Thus, the watermark may state something like, "Copy #47 issued to Jane Smith." An additional research problem that needs to be addressed in such a system is the adversary collusion: the watermark should be such that the comparison of two differently watermarked copies of the same document not lead to the discovery and removal of the watermarks.

3.2.7 Tamperproofing as Extensions of Watermarking

The watermarking technique can be interestingly reversed from the search for the most robust, indestructible watermark to that for the most brittle one, so that any tampering with a document would invariably lead to the removal of the watermark (see Atallah, Raskin *et al.* 2002) and thus signal the tampering. The initial research in this area demonstrates, interestingly and not quite unexpectedly, that designing the most brittle watermark is as challenging as designing the most robust and resilient one.

3.2.8. Enhancing Customer Acceptance of IAS Products with Computational Humor.

One of the biggest issues in IAS has been the refusal to deploy the acquired IAS products because of the reluctance to learn, install, and debug the developed systems. One approach to resolving this very real problem is to reward the system administrators (sysadmins) for making the effort by entertaining them throughout the process of installing and maintaining the product with the help of humor-generating intelligent embodied agents (see Nijholt 2002, Stock and Strapparava 2002). The current state of the art in computational humor is rapidly making it increasingly feasible. The idea does have a shock value to it, both for the better and for the worse: some hard-core techies in IAS, and, as a matter of fact, in NLP, think that computational humor is a hoax. Usually, a little homework

changes this attitude (see Raskin 1996, 2002; Raskin and Attardo 1994).

4 Perspectives, Challenges, Milestones

NLP deals with texts in NL, and in Section 3.2.1 above, we clearly stated that the applicability of NLP to IAS depends on the use of textual data in IAS systems. This statement was, actually, a considerable simplification.

For lower end, non-semantic NLP methods, those dependent on Boolean keywords, syntax, and/or statistics, the presence of textual data is indeed essential. For ontological semantics, which is a system of text meaning representation, the "text" itself may be in any non-natural-language format, including any scientific or logical formalism, as long as it has conceptual content. That content is directly representable with the help of the ontology, bypassing any NL lexicon if necessary. In other words, the ontology is equally applicable to a formal language as it is to a NL if a lexicon for the former is accessible.

Nevertheless, what applications of ontological semantics can contribute most obviously and on a broader scale, is extending research and application paradigms in IAS by including NL data sources and adapting the appropriate NLP applications, their goals and results to them. These include:

- inclusion of NL data sources as an integral part of the overall data sources in information security applications, and
- formal specification of the information security community know-how for the support of routine and time-efficient measures to prevent and counteract computer attacks

Where does NL data play a role in IAS? The applications listed in Section 3.2 provide the obvious examples. In addition, system administrator (sysadmin) logs, the standard object of data-mining efforts in IAS with the purpose of intrusion detection, are written in a sublanguage of a NL and can be allowed to contain more complex language if the processing systems are capable of treating it; however, all the pre-NLP studies ignore the NL clues in the logs and thus miss out on a great deal of important content. Similarly, to use another example, if an InfoSec task involves human alongside software agents, NLP is the most efficient way of handling interagent communication (see Nirenburg and Raskin 2002, Ch. 1, and references there).

In the past, all the above tasks, if at all attempted, were supported by either keyword-based search technology or through stochastic mechanisms of matching and determination of differences between two documents. These approaches have approached the ceiling of their capabilities.

An ontology provides a new, content-oriented, knowledge- and meaning-based approach to form the basis of the NLP component of the information security research paradigm. The difference between this knowledge-based approach and the old "expert system" approach is that the former concentrates on feasibility, for example, by using a gradual automation approach to various application tasks. The ontological approach also deals, albeit at a much more sophisticated level, with encoding and using the community know-how for automatic training and decision support systems. The cumulative knowledge of the information security community about the classification of threats, their prevention and about defense against computer attacks should be formalized, and this knowledge must be brought to bear in developing an industry-wide, constantly upgradeable manual for computer security personnel that may involve a number of delivery vehicles, including an online question-answer environment and a knowledge-based decision support system with dynamic replanning capabilities for use by computer security personnel. The underlying knowledge for both of these avenues of information security paradigm extension can, as it happens, be formulated in a single standard format. The knowledge content will readily enjoy dual use in both NL data inclusion and decision support, and it is made possible through the use of ontologies. Fig. 1 below shows a generic scheme of interaction of the ontological resources applied to a conceptual domain, such as

information security. The language-independent single ontology defines the content of most lexical entries in the lexicon and in the onomasticon (proper noun lexicon) of each NL. The fact database contains all the remembered event instances, and text meaning representations (TMR) are automatically generated for each text by the analyzer part of the processing system. The output, whether in NL or any other knowledge representation system, is produced by the generator from the TMRs. Some other static and dynamic resources are left out of the figure for simplification.

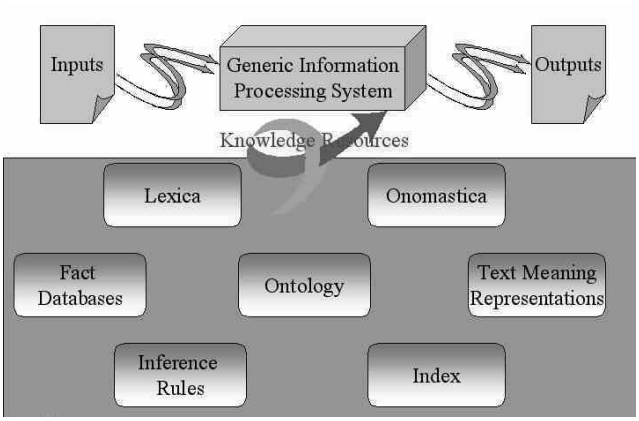

Figure 1. Application of the Ontological Paradigm to a Domain (e.g., IAS)

The attraction of using ontology, a conceptual structure for a domain of inquiry, is penetrating the IAS community only slightly more slowly than other disciplines. Since Raskin *et al.* 2001 and, especially, Raskin *et al.* 2002, the prospect of having a tangled hierarchy, or a lattice, bringing together all the main concepts in IAS, with a convenient public Web interface has found considerable support. The most practical interest has so far been along the lines of standardizing the IAS terminology. Research-wise, this is not the most challenging ontology-related issue among the ones listed above but, as many IAS gatherings amply demonstrate, different terminological dialects confuse and slow down many professional discussions. Much more practically and damagingly, the non-standard use of terms makes rapid responses to infections by CERT much more difficult because additional exchanges with the authors of reports are necessary to establish what is actually being reported.

Ontological semantics can develop as many useful tools to support the common language project, the standardization initiative in the IAS community (see Howard and Meunier 2002), with Web-interfaced, public-access ontological-semantic tools, as the implemented resources and their enhancements in this project will allow (e.g., dictionaries, both standard and dialectal; terminological ambiguity checker and corrector; mini-machine-translator from non-standard to standard usage).

Starting with such more or less obvious overlapping points, NLP can be used to enhance and enrich the IAS agenda by making many less obvious applications work in the domain. At the same time, the ever-changing and increasingly complex real-life and contentful needs of IAS will place demands on NLP, stimulating and guiding its development. We believe that content-, not formalism-oriented NLP approaches, such as ontological semantics, rather than non-meaning-based and/or non-representational approaches will be of most use to IAS. As in most fields populated by people trained in formalisms (and that includes both NLP and theoretical linguistics), there is a temptation to engage in a battle of formalisms to achieve maximum elegance, regardless of the formalized content—and, to add insult to injury, to be blissfully unaware of being not content-oriented. In linguistics, the practical task that used to provide a check against pure formalism-based approaches, the need to describe natural languages, has largely disappeared from the agenda. In NLP, there is more incentive to pay attention to content in contemporary applications, such as intelligent searches or question answering, than there was in MT, so the balance is changing in favor of content. In IAS, the practical task of preventing and countermanding hostile actions is fully dependent on understanding the content and goals of the actions, so the representation of meaning is a *sine qua non* of success, and this makes ontological semantics well suited for IAS applications. An ontological semanticist has the responsibility of identifying and sometimes discovering an IAS application of NLP

resources and of convincing the IAS community of the validity and importance of the application.

5 Conclusion

More and more interesting applications of NLP to IAS are being discovered, and the partial list above will be obsolete by the time this paper is presented. It is clear, therefore, that IAS is an important, enduring, and extremely well-funded field, whose needs NLP has every interest to serve and which will, therefore, determine, to an important extent, the development of NLP in the future. NLP, go for IAS!

Acknowledgments

The authors are grateful to CERIAS, with its pioneering multidisciplinary environment, and, especially, to its director, Eugene H. "Spaf" Spafford, for his vision in continuing to encourage and to support their work

References

Atallah, M., Raskin, V., Crogan, M., Hempelmann, C., Kerschbaum, F., Mohamed, D., and Naik, S. (2001). *Natural language watermarking: Design, analysis, and a proof-of-concept implementation*. In I. S. Moskowitz (ed.), "Information Hiding: 4th International Workshop, IH 2001, Pittsburgh, PA, USA, April 2001 Proceedings", I. S. Moskowitz, ed., Springer-Verlag, Berlin, pp. 185-199.

Atallah, M., Raskin, V., Hempelmann, C., Karahan, M., Sion, R., Topkara, U., and Triezenberg, K. E. (2002). *Natural language watermarking and tamperproofing.* Submitted to ih2002: Information Hiding Workshop 2002.

Howard, J. D., and Meunier, P. C. (2002). *Using a "common language" for computer security incident information.* In "Computer Security Handbook, 4th ed.", M. Kabay and S. Bosworth, eds., New York: Wiley.

McDonough, C. J. (2000). *Complex Events in an Ontological-Semantic Natural Language Processing System.* An unpublished Ph.D. thesis, Purdue University, W. Lafayette, IN.

Mohamed, D. (2001). *Ontological Semantics Methods for Automatic Downgrading.* An unpublished M. A. thesis, Purdue University, W. Lafayette, IN.

Nijholt, A. (2002). *Embodied agents: A new impetus to humor research.* In: Stock *et al.*, pp. 101-111.

Nirenburg, S., and Raskin, V. (1998). *Universal grammar and lexis for quick ramp-up of MT systems.* In "Proceedings of ACL/COLING '98. Vol. 2", Montreal: University of Montreal, pp. 975-979

Nirenburg, S., and Raskin, V. (2002). *Ontological Semantics.* Cambridge, MA: MIT Press (forthcoming).

Raskin, V. (1996). Computer implementation of the general theory of verbal humor. In: "Automatic Interpretation and Generation of Verbal Humor. International Workshop on Computational Humor, IWCH '96. Twente Workshop on Language Technology, TWLT 12", J. Hulstijn and A. Nijholt, eds., Enschede, NL: University of Twente, pp. 9-19.

Raskin, V. (2002). *Quo vadis computational humor.* In: Stock *et al.* 2002, pp. 31-46.

Raskin, V., Atallah, M. J., McDonough, C. J., and Nirenburg, S. (2001). *Natural language processing for information assurance and security: An overview and implementations.* In "NSPW '00: Proceedings of Workshop on New Paradigms in Information Security, Cork, Ireland, September 2000", M. Shaeffer, ed., New York: ACM Press, pp. 51-65.

Raskin, V., and Attardo, S. (1994). *Non-literalness and non-bona-fide in Language: An approach to formal and computational treatments of humor. Pragmatics and Cognition* 2/1, pp. 31-69.

Raskin, V., Hempelmann, C. F., Triezenberg, K. E., and Nirenburg, S. (2002). *Ontology in information security: A useful theoretical foundation and methodological tool.* In "Proceedings. New Security Paradigms Workshop 2001. September 10th-13th, Cloudcroft, NM, USA", V Raskin and C. F. Hempelmann, eds., New York: ACM Press, pp. 53-59.

Stock, O., Strapparava, C., and Nijholt A., eds. (2002), *Proceedings of The April Fools' Day Workshop on Computational Humour April 2002, Twente Workshop on Language Technology-TWLT 20, An Initiative of HAHAcronym, European Project IST-2000-30039*, Trento, Italy: ITC-irst.

Stock, O., and Strapparava, C. (2002). *Humorous agent for humorous acronyms: The HAHAcronym Project.* In: Stock *et al*. 2002, pp. 125-135.