When is Lying the Right Choice? *

Federico Cerutti†[1], Artemis Parvizi[1], Alice Toniolo[1], Dave Braines[2], Geeth R. de Mel[3], Timothy J. Norman[1], Nir Oren[1], Jeff Z. Pan[1], Gavin Pearson[4], Stephen D. Pipes[2] and Paul Sullivan[5]

[1]University of Aberdeen
[2]Emerging Technologies, IBM
[3]T. J. Watson Research Center, IBM
[4]DSTL
[5]INTELPOINT, Inc.

Abstract

Restricting the spread of sensitive information is important in domains ranging from commerce to military operations. In this position paper, we propose research aimed at exploring techniques for privacy enforcement when humans are the recipient of — possibly obfuscated — information. Such obfuscation could be considered to be a *white lie*, and we argue that determining what information to share and whether it should be obfuscated must be done through *controlled query evaluation*, which depends on each agent's risk/benefit evaluation. We concentrate on machine-human interactions, and note that appropriate specific natural language interfaces need to be developed to handle obfuscation. We propose a solution for creating controlled query evaluation mechanisms based on robust approaches for data representation under uncertainty, viz. $\mathcal{S}$DL-Lite, and propose using ITA Controlled English for natural language generation so as to handle subjectivity. We present the general architecture for our approach, with a focus on the relationship between formal ontology, controlled query evaluation, and natural language.

1 Introduction

Information is valuable, and controlling its spread is critical in domains where partially trusted parties interact with each other, especially in cases where information leakage can have serious economic or life-threatening repercussions. Simultaneously however, information must often be exchanged when co-operation is required to achieve a goal. In this paper, we examine how a balance between these two different pulls can be found through the use of *obfuscation* — the replacement of sensitive data with related information, thereby providing an entity with enough knowledge to achieve a goal while preventing sensitive inferences from being made. The problem of information leakage has already been widely studied within the area of autonomous agents. However, while much work has examined machine-machine contexts, the problem becomes more difficult when humans form part of the system. The ability of people to exploit background knowledge to make additional inferences can allow for more sensitive information to be discovered. However, it also provides more scope for misinformation to propagate.

The problem of obfuscating information for humans thus revolves around determining what natural language information to provide to the user in order to have them make desirable inferences, while

*This research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

†Corresponding author: f.cerutti@abdn.ac.uk

preventing undesirable inferences from being made. In this paper, we concentrate on deciding how to translate (portions) of an ontology to natural language. Furthermore, we concentrate on the domain of military intelligence analysis, where probabilities, uncertainty and impreciseness are a natural feature. We thus pay additional attention to the issue of how to express, and obfuscate, uncertainty in natural language. We address both of these issues through the use of $\mathcal{S}$DL-Lite — an extension of a tractable subset of description logics which utilises Dempster-Shafer's theory of evidence to encode uncertainties. (Şensoy et al., 2013) — as underlying ontology language. We then make use of ITA Controlled English (ITA-CE) (Mott, 2010) to translate concepts from $\mathcal{S}$DL-Lite into a natural language form. ITA-CE can easily represent Dempster-Shafer's theory of evidence in semi-structured form (Arunkumar et al., 2013), making it a natural candidate to encode concepts from $\mathcal{S}$DL-Lite. Furthermore, Kuhn (2014) has described ITA-CE as a general purpose language, reliable enough for automatic interpretation but with dominant natural elements and medium expressiveness.[1]

In this paper, we make use of provenance analysis as an exemplar domain. Provenance is becoming increasingly important with the growth in sharing and processing of scientific data. The most widely used representation of provenance is through the W3C standard PROV-O ontology,[2] which encodes provenance through a directed acyclic graph where nodes represent entities, activities or agents. Edges are labelled, and encode how nodes interact with each other. Idika et al. (2013) propose a probabilistic extension of traditional provenance graphs and Acar et al. (2013) introduce a core calculus for provenance which can be exploited for privacy enforcement. As discussed below, we seek to identify how information about provenance should be exchanged so as to achieve some (human or artificial) agent's goals, and demonstrate how and why obfuscation might be required.

The remainder of the paper is organised as follows. In Section 2, we present a motivational example, while in Section 3 we discuss the state-of-the-art with regards to approaches for controlled query evaluation and ontological reasoning under uncertainty. We then outline the main questions we will address in this research (Section 4), sketching the theoretical contribution required to make an impact when addressing machine-to-human interaction, our ultimate goal. In Section 5, we conclude the paper by discussing the evaluation procedures we will adopt for ensuring the soundness and correctness of our approach.

2 Motivation

Miles and Ella are two analysts operating as part of an Italy-UK-US coalition. They are tasked with determining the causes of an unidentified illness affecting livestock in an area under coalition control. They have been provided with a report indicating engineered non-waterborne contamination in the local water system. This report originated from an analysis by a UK based lab. However, this analysis was — to the best of the analysts' knowledge — performed by a third-party (non-governmental organisation) chemist. Additional uncertainties in the report arise as it could be the case that the chemist utilised data from what was assumed to be a water sample, which was expected to be generated by the sampling activity of what was claimed to be an Italian Water Sensor. This provenance chain is depicted in Figure 1 via the solid black lines.

Now assume that Miles works for the UK army and has access to the data presented above, while Ella is an analyst who is working for the US army. Miles shares only the final result with Ella the information "the water is contaminated", avoiding the sharing of provenance information.[3] However Ella can ask Miles specific questions regarding provenance, such as "who are the actors involved in this analysis?" (**Q1**). The answer can be "the involved actors are Italian sensors (likely), NGO chemist (possibly), and UK lab (almost certain)." Choosing the best term among *possibly*, *certain*, ..., can affect further inferences by Ella, and is one of the main problems we want to address.

[1]Classification $P^3E^3N^3S^3$ — reliably interpretable languages (P^3) with medium expressiveness (E^3); dominant natural elements (N^3); and lengthy descriptions (S^3) (Kuhn, 2014, Table 2).

[2]`http://www.w3.org/TR/prov-o/`.

[3]See `http://www.smh.com.au/articles/2003/07/02/1056825430340.html` for a real world example.

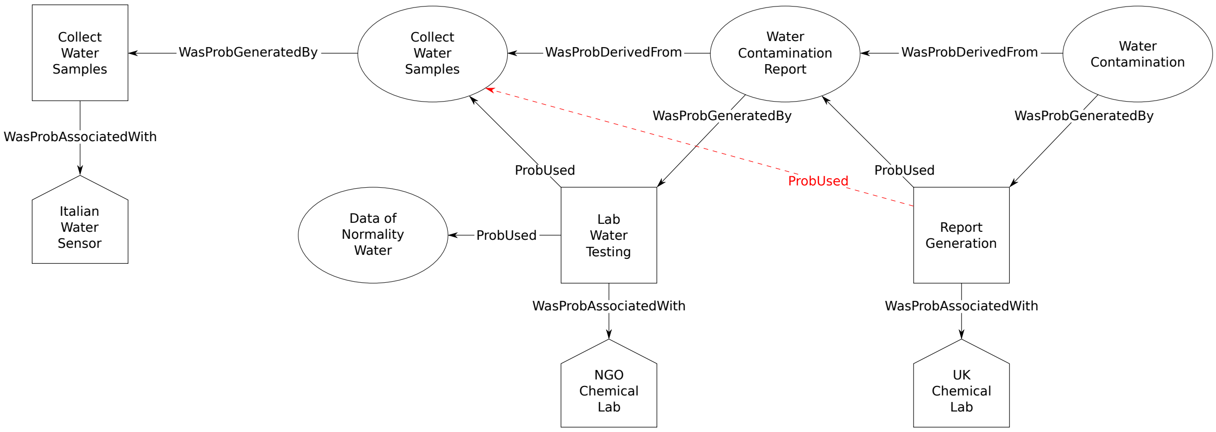

Figure 1: Provenance associated with the statement “the water is contaminated”. The dashed red line encodes the added relation (aimed at maintaining consistency with the white lie) “the UK did not use only the results coming from the NGO chemist.” Round nodes represent Entities, Squares represent Activities, and Pentagonal shapes represent Agents.

Moreover, Ella can in turn ask “did the UK lab use only the results coming from the NGO chemist?” (**Q2**). Miles knows that the results provided by the NGO are reliable, but Ella does not. Moreover, Ella might consider “collateral knowledge” or biases (e.g., “by default every NGO is unreliable”), and thus a positive answer would result in a loss of trust. Miles therefore has two choices for a response:

- abstain from providing the information: e.g. “I cannot tell you this”; or
- tell a white lie: e.g. “the UK did not only use the results coming from the NGO chemist.”

Identifying the best strategy is difficult; abstaining from answering can lead Ella to derive what should be kept confidential (Sicherman et al., 1983). At the same time, obfuscating information is a complex activity as Miles must maintain consistency during his interaction with Ella. For instance, since Miles already shared the information “The NGO chemist is possibly involved in the analysis,” he cannot retract this information. Therefore, one approach to maintaining consistency is to change the provenance graph by showing that the UK final report has been generated both from the NGO analysis and from the original water sample, as indicated by the red dashed arrow in Figure 1.

3 Background

3.1 Controlled Query Evaluation

Controlled Query Evaluation (CQE) is a framework for confidentiality enforcement, first proposed by Sicherman et al. (1983), and then developed in several other works, e.g. (Cuenca Grau et al., 2013; Biskup et al., 2014). We describe the main elements for a CQE framework based on the approach proposed by Biskup et al. (2014).

First of all, a *world view* is presented as a sequence $\langle\Gamma; \varphi_1, \ldots, \varphi_n\rangle$, where $\varphi_i \in \mathcal{L}$ are observations from a language $\mathcal{L}$, and Γ is the *background knowledge* (e.g., collateral knowledge, or even biases) over a language $\mathcal{L}_\Gamma$, which extends $\mathcal{L}$ to express rules.

Agents can form *beliefs* about $\varphi \in \mathcal{L}$ using the Bel operators, which belongs to a family of operators Ξ, such that $Bel : 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^* \mapsto 2^{\mathcal{L}}$. We say that the agent believes φ using Bel if $\varphi \in Bel(\langle\Gamma; \varphi_1, \ldots, \varphi_n\rangle)$.

Each belief operator $Bel \in \Xi$ satisfies:

- *consistency* — $\nexists \varphi \in \mathcal{L}, W \in 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^*$ s.t. $\varphi \in Bel(W)$ and $\neg\varphi \in Bel(W)$;
- *propositional consequence* — $\forall \varphi, \psi \in \mathcal{L}$, if $\varphi \vdash \psi$, $\forall W \in 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^*$, $\varphi \in Bel(W)$ implies $\psi \in Bel(W)$.

Moreover $\preceq_{cred}$ is an ordering over Ξ such that if $Bel \preceq_{cred} Bel'$ (Bel' is at least as credulous as Bel), then $\forall W \in 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^*$, $Bel(W) \subseteq Bel'(W)$.

In particular, let $\Xi^{RW} = \{Bel_p \mid p \in (0.5, 1]\}$, where $Bel_p(\langle \Gamma; \varphi_1, \ldots, \varphi_n \rangle) = \{\varphi \in \mathcal{L} \mid \delta(\varphi, \langle \Gamma; \varphi_1, \ldots, \varphi_n \rangle \geq p)\}$, with $\delta(\varphi, \langle \Gamma; \varphi_1, \ldots, \varphi_n \rangle) = \frac{|\mu(\varphi) \cap \mu(\Gamma \cup \{\varphi_1, \ldots, \varphi_n\})|}{|\mu(\Gamma \cup \{\varphi_1, \ldots, \varphi_n\})|}$, assuming that $\Gamma \cup \{\varphi_1, \ldots, \varphi_n\}$ is consistent and $\mu(\cdot)$ is a model operator (Bacchus, 1996). $Bel_p \preceq^{RW}_{cred} Bel_{p'}$ iff $p' \leq p$.

When an agent becomes aware that sentence $\varphi \in \mathcal{L}$ holds, i.e., when $\tau(\varphi) = \top$ with $\tau(\cdot)$ denoting a truth function, as the result of the speech act $inform(\varphi) \in Act$, it appends φ to its observation using the operator $+ : 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^* \times Act \mapsto 2^{\mathcal{L}_\Gamma} \times \mathcal{L}^*$ s.t. $\langle \Gamma; \varphi_1, \ldots, \varphi_n \rangle + inform(\varphi) = \langle \Gamma; \varphi_1, \ldots, \varphi_n, \varphi \rangle$.

A secrecy policy is then a set of secrecy constraints such that each constraint is of the form $\langle \varphi, Bel \rangle \in \mathcal{L} \times \Xi$ which expresses the desire to avoid that an agent believes φ when using the operator Bel.

In sharing a sentence, each agent must make assumptions about the recipient's world view and its choice of Bel operator for that sentence. This is encoded via a non-empty set of world views denoted the *postulated world view*.

Biskup et al. (2014) provides principles for a secrecy reasoner by taking into consideration a set of postulated world views and a set of possible actions to be classified:

- I.1: (*avoid potential violations*) it is desirable to avoid that in some of the postulated world views more secrecy constraints become violated;
- I.2: (*mitigate potential violations*) if a sentence cannot be protected against inferences with Bel as desired, it should at least be protected against inferences with operators less credulous than Bel;
- II: (*minimise classification*) a classification should be as minimally restrictive as possible w.r.t. other desires such as cooperative information sharing;
- III: (*be cautious towards credulous reasoners*) the more credulous the belief operator is postulated to be in a secrecy constraint, the more cautious an agent has to be while acting;
- IV: (*be more cautious the more uncertain*) the more world views a recipient holds according to the postulated world views, the more uncertain the situation is.

Moreover, a generic algorithm for implementing a secrecy reasoner is also provided in (Biskup et al., 2014, § 4) which we do not review here due to space constraints. It is worth noticing that Biskup et al. (2014)'s approach is propositional: each sentence φ is either true ($\tau(\varphi) = \top$) or false ($\tau(\varphi) = \bot$) and $inform(\varphi)$ implies that the recipient knows that φ holds ($\tau(\varphi) = \top$). For our scenario, discussed in Section 2, this is not enough as we need to consider uncertainties and probabilities.

3.2 Subjective DL-lite

Şensoy et al. (2013) introduce the $\mathcal{S}$DL-Lite formalism, which is an extension of a tractable subset of Description Logics with Dempster-Shafer Theory (DST) of evidence. In DST, a *binomial opinion* — or *SL opinion* — about a proposition ϕ is represented by a triple $w_\phi = \langle b(\phi), d(\phi), u(\phi) \rangle$, where $b(\phi)$ is the belief about ϕ — the summation of the probability masses that entail ϕ; $d(\phi)$ is the disbelief about ϕ — the summation of the probability masses that entail $\neg\phi$; $u(\phi)$ is the uncertainty about ϕ — the summation of the probability masses that entail neither ϕ nor $\neg\phi$; and $b(\phi) + d(\phi) + u(\phi) = 1$.

A DL-lite knowledge base $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. Axioms are either

- class inclusion axioms: $B \sqsubseteq C \in \mathcal{T}$ where B is a basic class $B := \mathsf{A} \mid \exists R \mid \exists R^-$ (A denotes a named class, R a named property, and R^- the inverse of R) and C is a general class $C := B \mid \neg B \mid C_1 \sqcap C_2$; or
- individual axioms: $B(\mathsf{a}), R(\mathsf{a}, \mathsf{b}) \in \mathcal{A}$ where a and b are named individuals.

$\mathcal{S}$DL-Lite (Şensoy et al., 2013) is an extension of DL-lite with subjective opinion assertions of the form $\mathcal{B} : w$ where w is an opinion and $\mathcal{B}$ is an ABox axiom. Let $\mathcal{W}$ be the set of all possible subjective binary opinions. A subjective interpretation is a pair $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$, where the domain $\Delta^\mathcal{I}$ is a non-empty set of objects, and $\cdot^\mathcal{I}$ is a subjective interpretation function, which maps:

- an individual a to an element of $\mathsf{a}^\mathcal{I} \in \Delta^\mathcal{I}$;

- a named class A to a function $\mathsf{A}^{\mathcal{I}} : \Delta^{\mathcal{I}} \mapsto \mathcal{W}$;
- a named property R to a function $R^{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mapsto \mathcal{W}$.

The complete specification of the semantics of $\mathcal{S}$DL-Lite is given in Table 1.

Syntax	Semantics	Syntax	Semantics
$\top$	$\top^{\mathcal{I}}(o) = \langle 1, 0, 0 \rangle$	$B_1 \sqsubseteq B_2$	$\forall o \in \Delta^{\mathcal{I}}, b({B_1}^{\mathcal{I}}(o)) \leq b({B_2}^{\mathcal{I}}(o))$ and $d({B_2}^{\mathcal{I}}(o)) \leq d({B_1}^{\mathcal{I}}(o))$
$\bot$	$\bot^{\mathcal{I}}(o) = \langle 0, 1, 0 \rangle$		
$\exists R$	$b((\exists R)^{\mathcal{I}}(o_1)) \geq \max \bigcup_{\forall o_2} \{b(R^{\mathcal{I}}(o_1, o_2))\}$ and $d((\exists R)^{\mathcal{I}}(o_1)) \leq \min \bigcup_{\forall o_2} \{d(R^{\mathcal{I}}(o_1, o_2))\}$	$B_1 \sqsubseteq \neg B_2$	$\forall o \in \Delta^{\mathcal{I}}, b({B_1}^{\mathcal{I}}(o)) \leq d({B_2}^{\mathcal{I}}(o))$ and $b({B_2}^{\mathcal{I}}(o)) \leq d({B_1}^{\mathcal{I}}(o))$
$\neg B$	$(\neg B)^{\mathcal{I}}(o) = \neg B^{\mathcal{I}}(o)$	$B(\mathsf{a}) : w$	$b(w) \leq b(B^{\mathcal{I}}(\mathsf{a}^{\mathcal{I}}))$ and $d(w) \leq d(B^{\mathcal{I}}(\mathsf{a}^{\mathcal{I}}))$
R^-	$(R^-)^{\mathcal{I}}(o_2, o_1) = R^{\mathcal{I}}(o_1, o_2)$	$R(\mathsf{a}, \mathsf{b}) : w$	$b(w) \leq b(R^{\mathcal{I}}(\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}))$ and $d(w) \leq d(R^{\mathcal{I}}(\mathsf{a}^{\mathcal{I}}, \mathsf{b}^{\mathcal{I}}))$

Table 1: Semantics of $\mathcal{S}$DL-Lite.

4 CQE Procedures for Provenance and ITA-CE Interfaces

The ultimate goal of this research is to provide computationally effective strategies for controlled query evaluation over uncertain knowledge bases with (controlled) natural language interfaces, thereby allowing the information recipient to be a human user. In particular, we focus on provenance due to its importance in domains such as intelligence analysis, and because it has a constrained domain ontology. Previous works, e.g. (Acar et al., 2013), examine privacy enforcement techniques for provenance, which provide a useful baseline for comparison.

There are three main concepts in the ontology presented in Figure 1, namely E (Entity), Ac (Activity), and Ag (Agent); and a variety of properties, such as R_{Assoc} (WasProbAssociatedWith), R_{Gen} (WasProbGeneratedBy), R_U (ProbUsed), R_{Der} (WasProbDerivedFrom). Thus, the answer to the first query (**Q1**) of Section 2 — "the involved actors are Italian sensors (*likely*),..." — could be the result of the following fragment in the knowledge base:

- $Ag(\mathsf{itwatsensor}) : \langle 1, 0, 0 \rangle$ (it is "certain" that "Italian Water Sensor" is an Agent);
- $Ac(\mathsf{collectwater}) : \langle 1, 0, 0 \rangle$ (it is "certain" that "Collect Water Samples" is an Action);
- $R_{Assoc}(\mathsf{collectwater}, \mathsf{itwatsensor}) : \langle 0.7, 0.01, 0.29 \rangle$ (it is "*likely*" that the Italian sensor is related to the water sampling activity).

The relationship between the concept *likely*, defined as "Such as well might happen or be true; probable," and the opinion $\langle 0.7, 0.01, 0.29 \rangle$, which could be encoded as "being certain at 70%, and admitting ignorance for 29% of cases," is clearly far from being universally acceptable. One of the difficulties of dealing with uncertain knowledge bases and humans, and also one of our main research questions, is *which are the suitable fuzzy categories for representing uncertainty in a machine-to-human dialogue*?

Moreover, in order to deal with subjective logic opinions, we extend the CQE framework by considering $\mathcal{S}$DL-Lite knowledge bases. Given a $\mathcal{S}$DL-Lite knowledge base $\mathcal{K}$, a sentence φ has an SL opinion associated to it — i.e., $\tau(\varphi) = w \in \mathcal{W}$. We can identify some limit-cases w.r.t. the propositional approach described in Section 3, namely $\top = \langle 1, 0, 0 \rangle$ and $\bot = \langle 0, 1, 0 \rangle$. Furthermore, as we already consider SL opinions associated to sentences, we will also extend the definition of Ξ^{RW} to $\Xi^{SL} = \{Bel_w \mid w \in \mathcal{W}\}$ thus applying opportune thresholds.

In addition, as a human engages in a querying process which might involve a white lie, coherence among the information shared during the dialogue must be ensured. This can be partially addressed by using belief revision approaches (Gärdenfors, 2005) in conjunction with plausibility metrics (Ma and Liu, 2008). This highlights another interesting question — similar to the one investigated by Cerutti et al. (2014) — viz. *are theoretically sound plausibility metrics useful in interacting with human users*?

Finally, formal results regarding privacy enforcement in the context of provenance, might be reused here (Acar et al., 2013). These approaches adopt very general languages to represent provenance traces but do not consider uncertainty. A way to abstract our representation of provenance data, to be able to reuse some of the results of Acar et al. (2013), is (1) to identify the actions in a provenance graph; (2) to create clusters around the action nodes; (3) to analyse the resulting clustered graph — potentially via a (probabilistic) finite automaton — as a probabilistic Kripke structure and represent it using probabilistic computational tree logic (CTL) (Kleinberg, 2012, Chapter 3). In this way, we could easily provide an answer to query (**Q2**), viz. "did the UK lab use only the results coming from the NGO chemist?" This thus raises the following question: *how to support a human user in querying provenance data*?

In our proposed use of ITA-CE as the Controlled Natural Language representation we also recognize that a well-known trait of Controlled Languages and Controlled Vocabularies is that they are "easy to read, hard to write" (Preece et al., 2014; Braines et al., 2014). Traditional solutions involved training and tooling to assist the user in constructing valid and meaningful sentences in these languages. In the work proposed here we envisage a system that provides maximum utility to untrained users in situations where they may have very limited support from tooling or custom applications. Example usage could include an SMS text message interface, or a Siri-like spoken conversation interaction. Therefore in related research (Preece et al., 2014; Braines et al., 2014) we are investigating human-machine conversational interaction where the human agent is able to express their input (assertion, question etc) in Natural Language, with the system providing an answer or interpretation in ITA-CE, thereby mitigating the "hard to write" aspect by enabling the user to provide Natural Language input, and confirming the machine interpretation by showing the human user the ITA-CE interpretation which we assert is "easy to read."

From a technical perspective, Ibbotson et al. (2012) develop a representation of the PROV-O using ITA-CE, and discuss how this can provide a narrative of the provenance documentation. While the model is fully integrated, the query procedure is still quite preliminary. Instead, Ibbotson et al. (2012) focus more on the relationship with trust and exploit one of the capabilities of ITA-CE, namely its ability to *explain* query results through chains of "because" statement (*rationale*). However, Ibbotson et al. (2012) do not consider neither uncertain knowledge, nor the need for privacy.

5 Conclusion

In this paper we outline the main elements for a research program aimed at ensuring privacy enforcement in shared information when humans are in the loop. As a case-study we consider the provenance record associated with information. Some approaches for privacy enforcement of provenance have recently been proposed although they focus on representation models quite distant from ontology-based ones, which are those of interest for us. Indeed, we want to develop an approach as general as possible — exploiting results related to controlled query evaluation approaches — thus relying on general ways to represent domain knowledge. To this end we chose to adopt an advanced description logic, namely $\mathcal{S}$DL-Lite, which is built on top of DL-lite, extending it with Dempster-Shafer's theory of evidence.

This brings us to one of the main components of our architecture, namely the choice of the ITA Controlled English ITA-CE (Mott, 2010). ITA-CE has the unique capability to handle subjective logic opinions (Arunkumar et al., 2013), and it has been shown to be able to encompass the PROV-O model of provenance (Ibbotson et al., 2012) with some basic query process. Our goal is to develop a natural language interface for our CQE approach. This is a key element as we plan to adopt it for evaluating this work through experiments involving human users, which, hopefully, will in turn guide the development of the underlying theory. Theoretical properties of intermediate results will be also evaluated either in connection with the relevant literature or comparing them against clear and intuitive desiderata.

Finally, additional interesting elements will be investigated, in particular the connection between time and privacy. There are naturally situations where the need of privacy is strictly related to time. For instance, to successfully rescue hostages, their position must be kept secret until the rescue operation is concluded (Cerutti et al., 2014). After the hostages are safe, there is no longer a need for keeping such a secret. Moreover, time and trust are strictly related as "distrust may be expected to retard the rate of information dissemination through a network" (Huynh et al., 2010). Similarly, a time-varying sharing effort can affect the trust perceived by the other members of a coalition.

References

Acar, U. A., A. Ahmed, J. Cheney, and R. Perera (2013). A core calculus for provenance. *J. of Comp. Sec. 21*(6), 919–969.

Arunkumar, S., M. Srivatsa, D. Braines, and M. Sensoy (2013). Assessing trust over uncertain rules and streaming data. In *FUSION*, pp. 1–8.

Bacchus, F. (1996). From statistical knowledge bases to degrees of belief. *Artif. Intell. 87*(1-2), 75–143.

Biskup, J., G. Kern-Isberner, P. Krümpelmann, and C. Tadros (2014). Reasoning on Secrecy Constraints under Uncertainty to Classify Possible Actions. In *FoIKS 2014*, pp. 97–116.

Braines, D., A. Preece, G. de Mel, and T. Pham (2014). Enabling CoIST users: D2D at the network edge. In *FUSION*, pp. 1–8.

Cerutti, F., G. R. de Mel, T. J. Norman, N. Oren, A. Parvizi, P. Sullivan, and A. Toniolo (2014). Obfuscation of Semantic Data: Restricting the Spread of Sensitive Information. In *DL2014*.

Cerutti, F., N. Tintarev, and N. Oren (2014). Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In *ECAI*, pp. 207–212.

Şensoy, M., A. Fokoue, J. Z. Pan, T. J. Norman, Y. Tang, N. Oren, and K. Sycara (2013). Reasoning about uncertain information and conflict resolution through trust revision. In *AAMAS*, pp. 837–844.

Cuenca Grau, B., E. Kharlamov, E. V. Kostylev, and D. Zheleznyakov (2013). Controlled Query Evaluation over OWL 2 RL Ontologies. In *ISWC*, pp. 49–65.

Gärdenfors, P. (2005). *The Dynamics of Thought*, Volume 300 of *Synthese Library*. Springer Netherlands.

Huynh, T. D., P. R. Smart, D. Braines, N. Shadbolt, and K. Sycara (2010). The Cognitive Virtues of Dynamic Networks. In *ACITA'10*.

Ibbotson, J., D. Braines, D. Mott, S. Arunkumar, and M. Srivatsa (2012). Documenting Provenance with a Controlled Natural Language. `https://www.usukita.org/node/2186`.

Idika, N., M. Varia, and H. Phan (2013). The Probabilistic Provenance Graph. In *SPW*, pp. 34–41.

Kleinberg, S. (2012). *Causality, Probability, and Time*. Cambridge University Press.

Kuhn, T. (2014). A Survey and Classification of Controlled Natural Languages. *Comp. Ling. 40*(1), 121–170.

Ma, J. and W. Liu (2008). A General Model for Epistemic State Revision using Plausibility Measures. In *ECAI*, pp. 356–360.

Mott, D. (2010). Summary of ITA controlled English. Technical report.

Preece, A., D. Braines, D. Pizzocaro, and C. Parizas (2014). Human-machine conversations to support multi-agency missions. *MC2R 18*(1), 75–84.

Sicherman, G. L., W. De Jonge, and R. P. Van de Riet (1983). Answering queries without revealing secrets. *ACM Trans. on Database Syst. 8*(1), 41–59.