# Accessing the Web of Data through Embodied Virtual Characters

Philipp Cimiano [a] Stefan Kopp [b]

[a] *Semantic Computing Group, Cognitive Interaction Technology Excellence Center (CITEC), Bielefeld University*
*E-mail: cimiano@cit-ec.uni-bielefeld.de*
[b] *Sociable Agents Group, Cognitive Interaction Technology Excellence Center (CITEC), Bielefeld University*
*E-mail: kopp@cit-ec.uni-bielefeld.de*

**Abstract.** The amount of data published on the Semantic Web has witnessed a tremendous growth in the last years to which the Linked Open Data (LOD) project has contributed significantly. While the Semantic Web was originally conceived of as an extension to the Web by addition of machine-readable data allowing automatic processing by machines, the question how humans can benefit from all the data published on the Web is becoming an important one. In the light of this question it seems crucial to make accessing the data on the Web as easy and intuitive as possible by adapting to the cognitive and information processing capabilities of humans. In this short position paper, we argue that one interesting and promising approach in this direction is to allow people to access semantic data on the Web through multimodal interaction with embodied virtual characters.

Keywords: semantic web, interaction, multimodal access, virtual characters

## 1. Introduction

The amount of data published on the Web of Data (a.k.a. The Semantic Web) has witnessed a tremendous growth in recent years. The Linked Open Data (LOD) project[1] has contributed significantly to this growth. People are in fact massively following Tim Berners-Lee's advice to publish data on the Web following the "linked data" principles [3,5]. Linked Data is a term referring to the recommended best practices for exposing, sharing, and connecting RDF data via de-referenceable URIs on the Semantic Web. While the Semantic Web was originally conceived as an extension to the Web by the addition of machine-readable data allowing automatic processing by machines, the question how humans can benefit from all the data published on the Web is certainly becoming a more and more important one.

In the light of this question it seems crucial to make accessing the data on the Web as easy and intuitive as possible. One central concern is to adapt to the cognitive and information processing capabilities of humans by making interaction on the Semantic Web *"meaningful"* for the user. Clearly, this is not only a user interface question. While suitable user interfaces are definitely yet to be seen, appropriate interaction paradigms for the Web of Data need to provide answers to the following questions:

– How much of the RDF graph should remain visible to end users? Should we fully abstract from the RDF data model / graph? It has been argued that the graph is actually not particularly useful as a way to present semantic data [28]. While some tools for accessing the Web of Data stick quite closely to the (linked) data graph (see Tabulator [4]), it seems important to abstract from the data graph when interacting with end users. After all,

---

[1] http://linkeddata.org/

why should users care about the data model if all they want is relevant information?

– How should information be packaged? That means, which are the information units that users can handle optimally? At one end of the spectrum, we have a single triple (arguably the smallest information unit on the Web of Data), at the opposite end we can come up with complete (multimodal) presentations generated by integrating various resources, aggregating data, computing diagrams, etc.

– What is the ideal interaction paradigm to access the Web of Data? Keyword querying [33], browsing [4], query-by-example, natural language [26] or even by way of reciprocal conversation between the user and the interface [14]?

– How can users understand non-trivial concepts such as trust, provenance, confidence etc? What are appropriate metaphors to convey such meta-information?

Providing answers to the above questions does clearly not only pertain to research on mere user interface design, but rather constitutes a non-trivial and long-term endeavor for the Semantic Web and related fields of research. Developing new and effective paradigms for interacting with the Semantic Web has indeed been recognized as a key challenge in the field (see [18,16,17]).

## 2. Motivation

We argue that one interesting and promising approach is to allow people to access semantic data on the Web through multimodal communication with embodied virtual characters. Consider that you would like to get information about the relation between the two French painters *Claude Monet* and *Edouard Manet*. Suppose also that you would like to get an overview about all painters considered as impressionists or that you would like to receive information about US presidents in chronological order. A natural way to pose such queries to a system is by way of natural language. And, as the Web of Data is structured, in contrast to the traditional Web, we can indeed provide answers to such information needs by fetching and re-composing different pieces of data available. Such a composition of available information into new informational structures that meet a current information need would be much more difficult on the traditional Web as it requires to understand the textual content first.

The even more important question for our purposes here is: which kind of structure would we like to get back as answer to such a request? An unordered and unstructured set of triples crawled from the Web of Data? Certainly not. Rather it seems crucial to find approaches that allow to assemble the relevant triples into a logical and coherent structure that can be conveyed to users. Imagine that you have a virtual character as assistant that you have posed the above query to compare Monet and Manet. The character would provide the following spoken answer along with different non-verbal modalities (we highlight the ouput in non-textual modalities in bold font):

**[Agent displays three photos of Claude Monet, Edouard Monet and one of Paris in the 19th century, respectively]** *"Both Claude Monet* **[points to the photo of Claude Monet]** *and Edouard Manet* **[points to photo of Manet]** *were French painters born in Paris in the 19th century* **[points to photo of Paris]**. *Monet was born on the 14th of November 1840, whereas Manet was born earlier on the 23rd of January of 1832. While both are associated with the Impressionism movement, Monet is also considered to belong to the realism movement. The most important works of Monet include Impression Sunrise, Rouen cathedral, London Parliament, Water Lilies and Poplar Series* **[agent sequentially blends in pictures of all these works, synchronized with its speech]**. *The most important works of Manet include "The Lunch on the grass".*

The strengths of such an approach to accessing the Web of Data by virtual embodied characters can be clearly appreciated: by packaging information into different modalities and units (e.g. sentences in speech) that people are used to from everyday conversation, we can generate a structured, yet compact, concise and amenable presentation. There are a number of further benefits, which we discuss with respect to the issues raised in the introduction:

– **Abstracting from the RDF data model:** It has been argued that the RDF graph is actually not particularly useful as a way to present semantic data [28]. Virtual characters are a promising way to realize a human-tailored access that abstracts from the RDF data model in order to transform the information into units that can be presented via different natural modalities (speech, text, gestures etc.).

– **Multimodal communication as interaction paradigm** Conveying multimodal output (using speech, intonation, gestures or facial expressions) allows to package information more effectively and compactly as different types of information can be conveyed across different suitable channels in parallel. As a corollary, this will lead to information packages that are closer to the information units that people are used to process and assimilate in daily interaction with human partners.

– **Tangible notion of a mediator:** Virtual characters are known to be entertaining and to increase the motivation to interact with a system [24]. In the difficult situation of wanting (or having) to access the abstract body of knowledge contained in the Semantic Web, we hypothesize that the presence of a virtual character can be beneficial because it makes tangible the notion of an assistant who is there to help users in finding the relevant information. Also, being able to formulate an information need in natural language is a custom and natural way for humans (and is increasingly supported by search engines, e.g. Wolfram-Alpha[2], or by natural language interfaces to the Semantic Web, see [27],[20]). Virtual characters are known to be social actors in the sense that they elicit the willingness to apply natural interaction patterns [24,23,10].

– **Expressing meta-information:** An embodied virtual character can use gestures, appropriate facial expressions, prosody and intonation, or linguistic modifiers such as *possibly* or *probably* to make clear that trust in a certain bit of information is low. As trust is an important building block of the Semantic Web (see [30,12,13,9,32,1]), conveying trust levels to human users becomes a crucial issue. Using gestures along with appropriate linguistic modifiers enables natural communication of such qualifiers together with the actual content, in a way that is more intuitive than presenting lists of items ranked by confidence or other symbolic or numerical representations of confidence values or trust levels.

## 2.1. Related Work

Developing new and effective paradigms for user interaction with the Semantic Web has been recog-nized as a key challenge in the field. Heath et al. [18] have identified the following challenges that *"must be addressed if Semantic Web technologies are to enter into widespread usage"*: i) increasing awareness, ii) providing clear benefits and iii) delivering appropriate functionality, iv) giving guidance for users, v) improving usability, vi) ensuring coherence of Semantic Web applications and vii) creating a critical mass of participation. In fact, there have been several workshops on this topic since 2004, e.g. IDWS[3], EUSW[4], SWUI'07[5], SWUI'08[6], SWUI'09[7]. Some authors of papers at these workshops have already proposed that conversational interaction with the Web of Data is an important interaction paradigm to explore (see [14] and [11]). However, there have only been quite preliminary approaches [7,21], which allow such an access and even extend it to the use of embodied virtual characters. For example, Kimura and Kitamura [21] directly embed RDF queries into utterance rules specified in the chatterbot markup language AIML. This simple approach allows for responding to a certain input phrase with a fixed utterance in which predetermined parts are replaced with retrieved fragments, but it represents by no means a flexible and comprehensive method to collect semantic data and to turn this into coherent multimodal presentations to satisfy the user's information need.

It is important to emphasize that we are not stating that an approach to access the Semantic Web / Web of Data by way of embodied virtual characters will solve all of the challenges raised by Heath et al. [18]. However, as argued above, an approach based on embodied virtual characters has the potential to provide access to the Web of Data in an intuitive and natural manner and thus to *improve usability*. Given the massive amount of data available, techniques that gather this data and generate intuitive and appealing summaries are addressing a clearly defined user need and deliver a *clear benefit* and *appropriate functionality*. Heath et al. claim that in order for semantic technologies to *increase in awareness* and receive widespread adoption we would need to hide the label "Semantic Web" and convey the fact to users that technology is doing useful things. The interaction with embodied virtual characters as we propose here would indeed contribute to making the tech-

---

[2] http://www.wolframalpha.com/

[3] http://interaction.ecs.soton.ac.uk/idsw04/
[4] http://www.ifi.uzh.ch/ddis/iswc2005ws.html
[5] http://swui.semanticweb.org/swui2007/
[6] http://swui.webscience.org/SWUI2008CHI/
[7] http://swui.webscience.org/SWUI2009/

nology and data models used behind the scenes transparent to the user while focusing on the system's presence as a helpful and useful assistant. This capitalizes on the fact that an agent can naturally provide assistance and guidance to the user in case he/she is experiencing problems, and can provide an enjoyable interaction which has the potential to increase the participation toward a *critical mass.*

As argued for by Dickinson [11], many problems for which the Semantic Web or Linked Open Data is useful are inherently exploratory in nature, where the users start out with a vague idea of what to look for and develop further insights into the nature of the enquiry during the interaction. This is exactly the type of incremental interaction that we aim to support and elicit with an embodied conversational character. It has been further argued for by Dickinson that *"turn-taking dialogues are a natural fit for the iterative exploration moves in exploratory search"* [11]. We agree with Dickinson that access though conversation is the most natural interaction type possible for humans engaging in exploration tasks. The fact that conversation-based access to the Semantic Web is an alternative with high potential is also argued for by Golbeck and Mutton, who allow users to access services from Internet Relay Chat (IRC) [14].

## 3. Challenges and Roadmap

We have highlighted the benefits of accessing the Web of Data by way of embodied virtual characters. However, this is a daunting endeavor requiring remarkable progress in a number of areas. We can, however, point out some of the challenges that need to be addressed:

- **Selecting and packaging information intro narratives**: An important challenge is to develop approaches that i) select the right information from the Web of Data to satisfy a user's information need, ii) construct plans how to convey this information in different modalities and iii) generate coherent narrative structures as output. The latter requires the generation of discourses beyond single sentences, which has been partially addressed in the language generation community (see [34], [29] [19]). The biggest challenge is to accomplish this robustly without requiring a fixed data schema to allow for scaling up to the size and heterogeneity of the Web of data.

- **Verbalization of Information:** Conversational access to the Semantic Web requires that we are able to verbalize RDF data, possibly in different languages. We hence need generation algorithms that can exploit linguistic knowledge captured in models such as LexInfo [8] about how data elements are to be realized linguistically in order generate language output. First approaches to verbalize semantic data have been presented [6] but are restricted to very rigid schemas and require manual effort by the user to adapt the system.

- **Generation of appropriate non-verbal behavior:** Flexibly producing gestures, facial expressions, or head movements that can accompany other modalities (speech, audio, video, text) is a non-trivial problem and subject of ongoing research [2]. In our context, one key challenge is to extract the information from a semantic resource that allows for generating behaviors that communicate differently from speech, e.g. gestures modulating or complementing it with imagistic or indexical information.

- **Language-based interfaces:** We need to support language-based interaction between the user and the conversational agent that goes beyond mere question-answering functionality (see [26]). One key issue is to be able to interpret language input for the user's information need, which is most often only possible by embedding it in the context of previous requests and presentations.

- **Synchronizing different modalities:** While expressing information through different channels (modalities) allows to compactly and efficiently apportion and encode information, synchronizing the different channels becomes crucial and is a big challenge (see [22]).

- **Robust Dialog Management in large domains:** An important challenge for providing conversational access to the Semantic Web is to be able to implement robust dialog management strategies that do not follow a fixed schema but allow, e.g., for clarification requests or repairments of misunderstanding. Thus, slot-filling techniques that have been developed in the dialog system community [25] are less suited in this context. Dialog management systems that can flexibly cope with arbitrary domain data are needed here. A first approach in this direction proposing to use a multimodal dialog system to access semantic data can be found in [15].

– **User acceptance**: One challenge is that embodied characters bear the risk of raising expectations with the users that the systems cannot live up to and that creates annoyance or frustration of the user. We argue that by smart design of the character and its interaction capabilities, and by a user-centered approach to developing character-based interfaces to the Web of Data, it may be possible to find the balance between what users demand from the system (e.g. full natural language conversation), what the character evokes by its appearance and behavior, and what it actually delivers.

While all of these issues are open research questions, we think that they are worthwhile to explore. We can conceive of a step-wise development towards the ultimate vision of an embodied virtual character that takes a query and answers it like a human expert in the respective field. At first, we will see agents that understand relatively simple requests and can automatically generate a simple discourse (possibly applying a limited set of templates). This first generation of agents might be already able to generate simple non-verbal output, synchronizing it with the speech modality and have basic mechanisms for conveying trust levels. The interaction with the user will be most likely text-based rather than via speech and there will be no mixed-initiative interaction, i.e, the agents will merely react to the input of a user. The technology for such characters is already available and building them is mainly a matter of system construction and attunement to the Semantic Web domain. Then, we will see agents that implement simple patterns of interaction and are able to engage in clarification dialogues. Simple mixed-initiative dialogs in selected domains have been realized already (see for example the project Gossip Galore which aims at developing conversational agents providing users access to pop trivia [35]). Finally, we might have reached a state that allows us to engage in conversation in selected domains and to receive multimodal information presentations from the character that are informative and tailored to the context. Robustness might be achieved by data-driven techniques which acquire script knowledge via games with a purpose [31], by observation, or via trial-and-error.

## 4. Conclusion

Providing meaningful interaction paradigms to access the Web of Data is an important topic for the Semantic Web community. We have suggested that providing access to the Semantic Web through multimodal interaction with embodied virtual characters is an interesting avenue to explore. Besides preliminary case studies, there has not been extensive research on this topic so far. As is clear from the challenges mentioned in this article, developing systems that provide conversational access to the Web of Data requires techniques and knowledge from a number of disciplines (dialog management, natural language processing, information retrieval, multimedia processing, virtual agents, etc.). Thus, we see it as a vision to which different research fields could (and should) contribute to and can cross-fertilize each other by doing so.

Certainly, "the one" interaction paradigm which fits all purposes and users does not exist. We thus think that it is only through appropriate user studies that we will be able to find out which (combination of) interaction paradigms are suited for which purpose. Conversational access to the Semantic Web might be one of them, possibly most suitable for casual users wanting to explore the Web of Data.

## References

[1] D. Artz and Y. Gil. A survey of trust in computer science and the semantic web. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):58–71, 2007.

[2] K. Bergmann and S. Kopp. Gnetic - using bayesian decision networks for iconic gesture generation. In *Proceedings of Intelligent Virtual Agents (IVA09)*, pages 76–89. Springer-Verlag, 2009.

[3] T. Berners-Lee. Linked data. http://www.w3.org/DesignIssues/LinkedData.html, 2006.

[4] T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, and D. Sheets. Tabulator: Exploring and analyzing Linked Data on the Semantic Web. In *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.

[5] C. Bizer, T. Heath, and T. Berners-Lee. Linked data - the story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.

[6] K. Bontcheva. Generating tailored textual summaries from ontologies. In *Proceedings of the European Semantic Web Conference (ESWC)*, pages 531–545, 2005.

[7] A. Breuing, T. Pfeiffer, and S. Kopp. Conversational interface agents for the semantic web – a case study. In *Proceedings of the 7th International Semantic Web Conference (ISWC2008)*, 2008.

[8] P. Buitelaar, P. Cimiano, P. Haase, and M. Sintek. Towards linguistically grounded ontologies. In *Proceedings of the 6th Annual European Semantic Web Conference (ESWC)*, pages 111–125, 2009.

[9] J. Carroll, C. Bizer, P. J. Hayes, and P. Stickler. Named graphs, provenance and trust. In *Proceedings of the World Wide Web (WWW) Conference*, pages 613–622, 2005.

[10] D.M. Dehn and S. van Mulken. The impact of animated interface agents: A review of empirical research. *Int. J. Human-Computer Studies*, 52:1–22, 2000.

[11] Ian Dickinson. In favour of (more) intelligence in the semantic UI. In *Proceedings of the Semantic Web User Interaction Workshop (SWUI'09),*, 2009.

[12] R. Queiroz Dividino, S. Sizov, St. Staab, and B. Schüler. Querying for provenance, trust, uncertainty and other meta knowledge in RDF. *Journal of Web Semantics*, 7(3):204–219, 2009.

[13] J. Golbeck, P. A. Bonatti, W. Nejdl, D. Olmedilla, and M. Winslett, editors. *Proceedings of the ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*. CEUR-WS.org, 2004.

[14] J. Golbeck and P. Mutton. Semantic web interaction on internet relay chat. In *Proceedings of the Workshop on Interaction Design on the Semantic Web*, 2004.

[15] Y. He, T.T. Quan, and S.C. Hui. A multimodal restaurant finder for semantic web. In *Proceedings of the IEEE International Conference on Computing and Communication Technologies (RIVF)*, 2006.

[16] T. Heath. How will we interact with the web of data? *IEEE Internet Computing*, 12(5):88–91, 2008.

[17] T. Heath. A taskonomy for the semantic web. *Semantic Web Journal*, 2010. accepted for publication, this volume.

[18] T. Heath, J. Domingue, and P. Shabajee. User interaction and uptake challenges to successfully deploying semantic web technologies. In *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.

[19] E.H. Hovy. Automated discourse generation using discourse structure relations. *Artif. Intell.*, 63(1-2):341–385, 1993.

[20] E. Kaufmann and A. Bernstein. How useful are natural language interfaces to the Semantic Web for casual end-users? In K. Aberer, K.-S. Choi, N. Noy, D. Allemang, K.-I. Lee, L.J.B. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, and P. Cudré-Mauroux, editors, *ISWC/ASWC*, volume 4825 of *Lecture Notes in Computer Science*, pages 281–294. Springer, 2007.

[21] M. Kimura and Y. Kitamura. Embodied conversational agent based on Semantic Web. In *Agent Computing and Multi-Agent Systems*, LNCS 4088, pages 734–741. Springer-Verlag, 2006.

[22] S. Kopp, K. Bergmann, and K. Wachsmuth. Multimodal communication from multimodal thinking - towards an integrated model of speech and gesture production. *Semantic Computing*, 2(1):115–136, 2008.

[23] S. Kopp, L. Gesellensetter, N. Krämer, and I. Wachsmuth. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, LNAI 3661, pages 329–345. Springer-Verlag, 2005.

[24] N. Krämer. *Soziale Wirkung virtueller Helfer*. Kohlhammer, 2008.

[25] O. Lemon, K. Georgila, J. Henderson, and M. Stuttle. An ISU dialogue system exhibiting reinforcement learning of dialogue policies: generic slot-filling in the talk in-car system. In *Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics: Posters & Demonstrations (EACLÂt'06)*, pages 119–122. Association for Computational Linguistics, 2006.

[26] V. Lopez, M. Pasin, and E. Motta. Aqualog: An ontology-portable question answering system for the semantic web. In *Proceedings of the European Semantic Web Conference (ESWC)*, pages 546–562, 2005.

[27] V. Lopez, V.S. Uren, M. Sabou, and E. Motta. Cross ontology query answering on the Semantic Web: An initial evaluation. In Y. Gil and N. Noy, editors, *K-CAP*, pages 17–24. ACM, 2009.

[28] mc schraefel and D. Karger. The pathetic fallacy of rdf (position paper). In *Proceedings of the Semantic Web User Interaction Workshop (SWUI'06)*, 2006.

[29] K.R. McKeown. *Text generation: using discourse strategies and focus constraints to generate natural language text*. Cambridge University Press, New York, NY, USA, 1985.

[30] K. O'Hara, H. Alani, Y. Kalfoglou, and N. Shadbolt. Trust strategies for the semantic web. In *Proceedings of the ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*, 2004.

[31] J. Orkin and D. Roy. Automatic learning and generation of social behavior from collective human gameplay. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, pages 385–392, 2009.

[32] M. Richardson, R. Agrawal, and P. Domingos. Trust management for the semantic web. In *Proceedings of the International Semantic Web Conference (ISWC)*, pages 351–368, 2003.

[33] T. Tran, H. Wang, S. Rudolph, and P. Cimiano. Top-k exploration of query candidates for efficient keyword search on graph-shaped (RDF) data. In *Proceedings of the International Conference on Data Engineering (ICDE)*, pages 405–416, 2009.

[34] L. Wanner. On lexically biased discourse organization in text generation. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, pages 369–378, 1994.

[35] F. Xu, P. Adolphs, H. Uszkoreit, X. Cheng, and H. Li. Gossip galore: A conversational web agent for collecting and sharing pop trivia. In *Proceedings of the International Conference on Agents and Artificial Intelligence (ICAART)*, pages 115–122, 2009.