Unrethinking the Semantic Web

Aidan Hogan[1[0000−0001−9482−1982]], Ora Lassila[2[0000−0002−0001−9137]], and James A. Hendler[3[0000−0003−3056−1960]]

[1] DCC, Universidad de Chile & IMFD, ahogan@dcc.uchile.cl
[2] Accenture, ora.y.lassila@accenture.com
[3] Rensselaer Polytechnic Institute, hendler@cs.rpi.edu

Abstract. We argue that the original vision articulated in the 2001 Scientific American article is as relevant and important today as it was 25 years ago. Enabling users to benefit more from the Web via increased automation is an idea that, if anything, has only become more compelling with time. Semantics, agents, trust, inference and (even) ontologies are once again in vogue, but this time it is the AI community that speaks about them with excitement, not us. Though the path to achieve the vision requires some rerouting from that envisaged in 2001, so long as the destination remains unchanged, it is closer than ever. In this (un)visionary paper, we argue that a quarter century further on, it is not the time to rethink the Semantic Web – it is the time to reclaim it.

Keywords: Semantic Web

1 Introduction

The year is 2001 and "*We Can Work It Out*" is playing on the stereo. An incoming call interrupts the chorus. In the short conversation that follows, Pete and Lucy discuss the health concerns of their mother and coordinate biweekly physical therapy sessions for her. With the help of their "agents", and something called the "Semantic Web", they select a clinic with availability at convenient times that is in-plan for her insurance, organise turns to bring her to the clinic, and reschedule other non-essential commitments conflicting with those appointments. The call ended, the Beatles fill the room again.

Later, when prompted by Pete, the Semantic Web explains how it found that clinic, and made the other arrangements, showing its workings.

So was the vision described by Berners-Lee et al. [6] in the Scientific American article entitled "The Semantic Web". The trust that Lucy and Pete had in their agents and this Semantic Web seems almost utopian in a modern context. But at that time, optimism was warranted. The vision was a call to arms: to build an even wider, wiser, worldlier Web. Many answered the call.

But despite multiple decades of innovation and toil, of experiments and proofs, of standards and specifications, of investment and deliverables, of development and research – of spirited debates on blank nodes and httpRange-14 and literals-as-subject and the one-right-way-to-reify and how-much-semantics-is-*the*-right-amount-of-semantics – the Petelucian Web envisaged at the dawn of the new millennium yet remains just out of reach.

Hence, after 25 years, there are calls to rethink the Semantic Web envisaged in 2001. But if we were to rethink it, *what would we change?*

This question is left as an exercise to the reader [6].

2 So Far

After a quarter century of investment, research and development, it is perhaps only fair to ask: *what has the Semantic Web ever done for us?*

Some might respond: KNOWLEDGE GRAPHS! Enterprise knowledge graphs can now be found in companies ranging from Amazon [35] to Zillow.[4] Open knowledge graphs, like DBpedia [23], Wikidata [33] and YAGO [32], continue to spur on research and development. Wikidata, in particular, describes more than one hundred million entities and is continuously edited by thousands of active editors in hundreds of languages; it supplies Wikipedia with infoboxes, interlanguage links, and rich references; it links the Web via an identifier scheme that includes over ten thousand distinct external identifiers; it drives applications for digital libraries, cultural heritage, biomedicine, among many other domains.

Others might respond: METADATA! As one of the earliest adopters, RDF metadata is still found in every Adobe document. JSON-LD [21] is now used to embed metadata on most websites.[5] The schema.org vocabulary – led by key members of the Semantic Web community[6] – powers features in major web search engines like Google's Rich Snippets and the Google Knowledge Graph.

Others still might respond: ONTOLOGIES! Ontologies play a central role in biomedicine, which now boasts hundreds of interoperable ontologies that provide standardised vocabularies for diseases [29], drugs [17], genes [2] and phenotypes [28], to name but a few examples. These ontologies support clinical trials [12], diagnosis [29], drug discovery [2], and (thus) better outcomes for patients. And beyond biomedicine, ontologies continue to support data interoperability in diverse domains, such as cultural heritage [10], digital libraries [24], finance [5], agriculture [30] and manufacturing [13], to name some examples.

Others yet might respond: the SYSTEMS! They might point to Jena [34], or Sesame/RDF4j [9], or Pellet [31], or Hermit [16], or Protégé [27], or Virtuoso [14], or Qlever [3] used to host the DBLP knowledge graph [1], or Blazegraph used to host Wikidata and evolving into Amazon Neptune [4], or RDFox [25] powering personal knowledge graphs on every Samsung Galaxy S25 phone [11].

[4] https://www.zillow.com/news/leveraging-knowledge-graphs-in-real-estate-search/; retr. 2026-08-03

[5] https://w3techs.com/technologies/details/da-jsonld; retr. 2026-08-03; specifically JSON-LD is found in most websites *surveyed*.

[6] https://schema.org/docs/about.html; retr. 2026-08-03

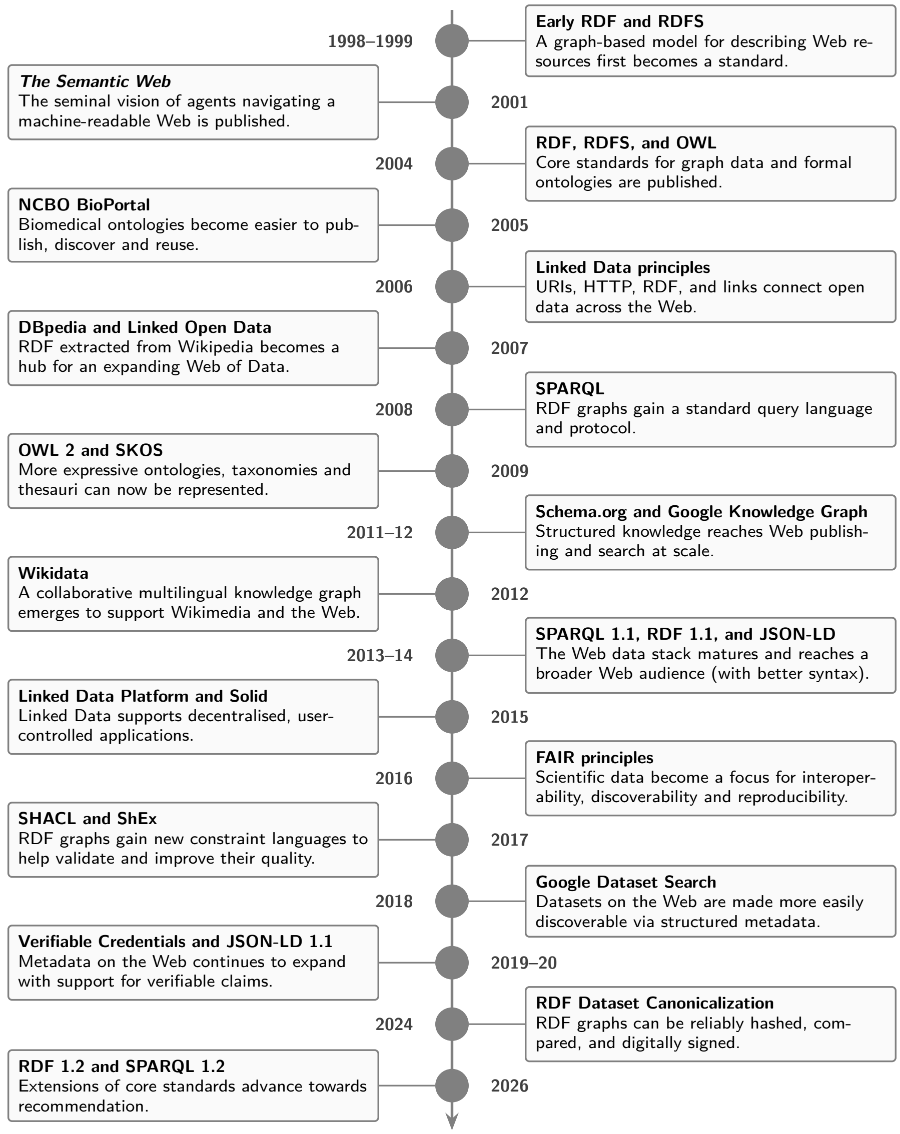

Fig. 1. A selection of milestones for the Semantic Web (et al.) so far

Others yet still might respond: Linked Data! or FAIR Principles! or The standards! or The theory! or The start-ups! The research! The literature! The PhD graduates! The know-how! The policies! The ...!

We give a small selection of milestones for the Semantic Web in Figure 1.

Charles Ivie, in praise of the Semantic Web, astutely describes it as a "*catastrophically successful failure*". To this, we would simply append: "... *so far*".

3 And Even Further

This year, we commemorate two anniversaries. The first marks 25 years since 2001, when Scientific American published "The Semantic Web" paper [6]. The second marks 10 years since 2016, when the International Semantic Web Conference last published a research track paper with "Semantic Web" in its title [8].[7]

A quarter century ago, the Semantic Web community was inspired by autonomous agents that could make Pete and Lucy's lives easier. Jump forward to today and it is ostensibly the Artificial Intelligence community, not the Semantic Web community, leading the way towards making these agents a reality. The trending adjective for AI – classified by Merriam–Webster, at least for now, as slang[8] – is "*agentic*". Semantics, ontologies and graphs are in vogue again, but this time to make generative AI more reliable, explainable, etc.

What is currently missing in order to realise these autonomous agents? An agent should have the ability to make decisions and act independently on someone else's behalf. An autonomous agent should do so autonomously: it should have agency. The new "agentic" systems are not really exhibiting these qualities.

To build true autonomous agents, first, Pete and Lucy would need unfettered access to their own personal data. Second, they would need control over their personal data: what their agents can access and use under what policies. Third, they would need agents that can leverage these personal data, in compliance with the policies, in order to make decisions and act independently on their behalf via the Web. Fourth, they would need agents that can likewise make decisions and act independently on their user's behalf when interacting with other agents. Fifth, they would need agents that engender trust in their users: that are reliable, effective, aligned with their users, and can verify and explain their workings.

Those who worked on Semantic Web Services [15] in the decade following 2001 – arguably the most "agentic" approach to the Semantic Web – might further interject that this alone is not enough. The problems of discovery, invocation, composition, interoperation, etc., will inevitably re-emerge in this new era.

The possible paths towards realising such agents recall the old AI distinction of "neats" and "scruffies" [26]. The neats favoured elegant, general, formally grounded methods; the scruffies favoured pragmatic combinations of techniques and domain knowledge that *worked*, even if without a single unifying theory.

The 2001 Semantic Web article falls mostly into the "neats" tradition, suggesting that Pete and Lucy's agents will "*know*", purely from semantics encoded into the Web page, that "*Dr. Hartman works at this clinic on Mondays* [...] *without needing artificial intelligence on the scale of 2001's Hal or Star Wars's C-3PO*" [6]. But now that modern AI more closely resembles Hal and C-3PO than the AI of 2001, it seems reasonable to ask how it might play a role.[9]

---

[7] The streak continues among 51 papers accepted to the research track of ISWC 2026. The authors now regret not doing their part to break this streak.

[8] See https://www.merriam-webster.com/slang/agentic; retr. 2026-08-31

[9] When the article was written society had not yet really seen the "miraculous success" of statistical (non-symbolic) AI, machine learning, and all that.

There are then three avenues – the first two neat and the last one scruffy – to build agents that can help us to automate complex tasks on the Web: build them symbolically, build them neurally, and build them neurosymbolically. In 2001, the premise was that we could knowledge-represent such agents into existence, but the idea of constructing them (purely) symbolically appears more optimistic than once thought. On the other hand, constructing agents (purely) neurally, and then trusting them to act on our behalf – to entrust them with key healthcare decisions, for example – likewise appears too optimistic. The best prospect to build these agents appears increasingly scruffy, and neurosymbolic.

Many of the pieces needed to complete this puzzle are available (Solid, access control, verifiable credentials, "agentic" AI, GraphRAG, etc. [7]). What is missing is to put these pieces together, and to start to *build* such agents.

4 Reclaiming the Semantic Web

We have seen lots of exciting work on the Semantic Web following (and even preceding) the publication of the Scientific American article. It is obvious that the greatest contribution of that article was to inspire an entire community to "chase a dream". The vision, as we have described, is as valid and relevant as ever, and we are now in a much better position to fully realise it [22, 19, 7].

Rather than rethink the Semantic Web, we contend that it is time to reclaim it. A good starting point for such a reclamation is the German adage from 1512: "*das Kind mit dem Bade ausschütten*". But in the past 25 years of Semantic Web endeavour, what is the bathwater, and wherein lies the baby?

The baby remains that envisaged a quarter century ago to help Pete and Lucy. A Web with personal agents that automate tasks for their users, that yield better outcomes for them, that empower them and make their lives better. A Web with agents that can explain themselves, that are trusted by their users, that trust the agents of users whom their users trust, that don't trust the agents of users whom their users don't trust. A Web where users and their agents control, protect and leverage personal data in a decentralised way. A Web with agents that grant agency to their users. A Web that grants agency to such agents.

The bathwater is the murkier, and the more controversial. One could point to a lack of good tooling, a lack of usability, or that our community has been "too academic". But this is all part of the maturation process: tooling has improved usability, users have improved tooling, theory has informed practice, and practice has informed theory. This has given us many success stories.

Rather than its growing pains, the bathwater is the antithesis of the baby. It is a Web with agents that yield worse outcomes for its users, that disempower them and make their lives no better. A Web with agents that cannot explain themselves, that are not trusted by their users, that trust the agents of users whom their users don't trust, and don't trust the agents of users whom their users trust. A Web where users and their agents don't control, cannot protect and fail to leverage personal data in a decentralised way. A Web with agents that rob users of their agency. A Web that grants agency to such agents.

This bathwater may sound vaguely familiar. The Web of 2026 is not the same as the Web of 2001. Much of the original vision of decentralisation and "users in control" was lost to the commercial push to centralise social publishing. Perhaps reclaiming the Semantic Web vision will require us to reclaim the Web in a broader sense, and use agents (the "good" kind) to do that.

The concept of "agents" is not new, and dates back to at least 1750 BCE, and the Code of Hammurabi. Many cultures – Chinese Confucianism, early Jewish Rabbinical tradition, early Hinduism and Roman rulers – further found the concept of an "agent" useful for their societies. The agents we speak of are less fleshy than these precedents, but the core principles of a good agent are similar [20]: delegation, trust, alignment, autonomy. Most importantly, all of these various traditions share a commonality: that some of the responsibility lay with the agent, and some with the entity that empowered the agent.

In modern terms (only a century old), the notion of formalising this empowerment as "deontic logic" emerged: to define what the agent should or must do and should not or must not do, potentially paving the way for web users to create agents where they keep control over these aspects, where agents interacting with other agents do not violate the rules, and where websites are able to recognize and comply with the user's wishes. Previous work has tried to explain these ideas in simple terms [18], but contemporary "agentic" AI remains focused narrowly and primarily on the automation of simple tasks, not the complex interaction of agents and websites as put forth in the original Semantic Web vision.

The original vision of the Semantic Web – seen as about intelligent, autonomous agents and the data needed to "feed" these agents – was hurt by the timing of its publication: in 2001, we were still at the tail-end of the second "AI Winter", and the agent part of the vision was largely ignored or forgotten. The "new AI" – generative AI, LLMs, and "agentic" systems – has reinvigorated the discussion and consideration of intelligent, autonomous agents. Much of the contemporary work on "agentic" systems may not be exactly what the Semantic Web vision suggested, but now is the perfect moment for us to reclaim that idea. The work on the Semantic Web and its underpinnings has yielded many useful components for this endeavour (see, for example, Figure 1).

Thus, consider this a renewed "call to action" for our community and the broader AI community to finish what was started a quarter century ago: to build an even wider, wiser, worldlier Web, powered by the right kind of agents.

Because whatever the challenge, we can work it out.

Acknowledgements. We thank Tim Berners-Lee for encouragement and suggestions regarding the paper. Hogan was supported by ANID – Millennium Science Initiative Program – Code ICN17 002, Chile.

Use of Generative AI

ChatGPT was used to help generate the structure of Figure 1, and to check grammar before submitting.

References

1. Ackermann, M.R., Bast, H., Beckermann, B.M., Kalmbach, J., Neises, P., Ollinger, S.: The dblp Knowledge Graph and SPARQL Endpoint. Transactions on Graph Data and Knowledge **2**(2), 3:1–3:23 (2024). https://doi.org/10.4230/TGDK.2.2.3
2. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al.: Gene ontology: tool for the unification of biology. Nature genetics **25**(1), 25–29 (2000)
3. Bast, H., Buchhold, B.: Qlever: A Query Engine for Efficient SPARQL+ Text Search. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. pp. 647–656 (2017). https://doi.org/10.1145/3132847.313292
4. Bebee, B.R., Choi, D., Gupta, A., Gutmans, A., Khandelwal, A., Kiran, Y., Mallidi, S., McGaughy, B., Personick, M., Rajan, K., Rondelli, S., Ryazanov, A., Schmidt, M., Sengupta, K., Thompson, B., Vaidya, D., Wang, S.: Amazon Neptune: Graph Data Management in the Cloud. In: ISWC (P&D/Industry/BlueSky) (2018)
5. Bennett, M.: The financial industry business ontology: Best practice for big data. Journal of Banking Regulation **14**(3), 255–268 (2013)
6. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American **284**(5), 34–43 (2001)
7. Bonatti, P.A., Domingue, J., Gentile, A.L., Harth, A., Hartig, O., Hogan, A., Hose, K., Jimenez-Ruiz, E., McGuinness, D.L., Sun, C., Verborgh, R., Wright, J.: Towards Computer-Using Personal Agents (2025), https://arxiv.org/abs/2503.15515
8. Brasileiro, F., Almeida, J.P.A., Carvalho, V.A., Guizzardi, G.: Expressive multi-level modeling for the semantic web. In: International Semantic Web Conference. pp. 53–69. Springer (2016)
9. Broekstra, J., Kampman, A., van Harmelen, F.: Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. In: International Semantic Web Conference. pp. 54–68 (2002)
10. Bruseker, G., Carboni, N., Guillem, A.: Cultural heritage data management: The role of formal ontology and CIDOC CRM. Heritage and archaeology in the digital age: acquisition, curation, and dissemination of spatial cultural heritage data pp. 93–131 (2017)
11. Crocker, P., Horrocks, I., Nenov, Y.: Semantic Technology in Your Pocket. In: ISWC (Industry/Doctoral Consortium/Posters/Demos). pp. 17–19 (2025)
12. Donnelly, K.: SNOMED-CT: The advanced terminology and coding system for eHealth. Stud Health Technol Inform **121**, 279–290 (2006)
13. Drobnjakovic, M., Kulvatunyou, B.S., Ameri, F., Will, C., Smith, B., Jones, A.T.: The Industrial Ontologies Foundry (IOF) Core Ontology. In: International Workshop on Formal Ontologies meet Industry (FOMI 2022). CEUR Workshop Proceedings, vol. 3240. CEUR-WS.org (2022)
14. Erling, O.: Virtuoso, a Hybrid RDBMS/Graph Column Store. IEEE Data Eng. Bull. **35**(1), 3–8 (2012)
15. Fensel, D., Facca, F.M., Simperl, E., Toma, I.: Semantic Web Services, vol. 357. Springer (2011)
16. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z.: HermiT: An OWL 2 Reasoner. J. Autom. Reasoning **53**(3), 245–269 (2014). https://doi.org/10.1007/s10817-014-9305-1
17. Hanna, J., Joseph, E., Brochhausen, M., Hogan, W.R.: Building a drug ontology based on RxNorm and other sources. Journal of biomedical semantics **4**(1), 44 (2013)

18. Hendler, J.: Is there an intelligent agent in your future? Nature (1999). https://doi.org/10.1038/nature28027
19. Hendler, J., Lassila, O.: Is the Semantic Web vision still relevant? KGC Keynote (2026), https://www.lassila.org/publications/2026/hendler+lassila-kgc2026.pdf
20. Ibáñez, L.D., Domingue, J., Kirrane, S., Seneviratne, O., Third, A., Vidal, M.E.: Trust, Accountability, and Autonomy in Knowledge Graph-Based AI for Self-Determination. Transactions on Graph Data and Knowledge **1**(1), 9:1–9:32 (2023). https://doi.org/10.4230/TGDK.1.1.9
21. Kellogg, G., Champin, P.A., Longley, D., Sporny, M., Lanthaler, M., Lindström, N.: JSON-LD 1.1 – A JSON-based Serialization for Linked Data. W3C Recommendation (Jul 2020), https://www.w3.org/TR/json-ld11/
22. Lassila, O.: Semantic Web and AI: Can we finally realize the vision? ISWC Keynote (2024), https://www.lassila.org/publications/2024/lassila-iswc2024-keynote.pdf
23. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P.N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., Bizer, C.: DBpedia – A large-scale, multilingual knowledge base extracted from Wikipedia. Semantic Web **6**(2), 167–195 (2015). https://doi.org/10.3233/SW-140134
24. Miller, E., Ogbuji, U., Mueller, V., MacDougall, K.: Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. US Library of Congress Report (Nov 2012), https://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf
25. Nenov, Y., Piro, R., Motik, B., Horrocks, I., Wu, Z., Banerjee, J.: RDFox: A highly-scalable RDF store. In: International Semantic Web Conference. pp. 3–20. Springer (2015)
26. Nilsson, N.J.: Artificial intelligence prepares for 2001. AI Magazine **4**(4) (1983)
27. Noy, N.F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R.W., Musen, M.A.: Creating Semantic Web Contents with Protégé-2000. IEEE Intell. Syst. **16**(2), 60–71 (2001). https://doi.org/10.1109/5254.920601
28. Robinson, P.N., Köhler, S., Bauer, S., Seelow, D., Horn, D., Mundlos, S.: The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. The American Journal of Human Genetics **83**(5), 610–615 (2008)
29. Schriml, L.M., Arze, C., Nadendla, S., Chang, Y.W.W., Mazaitis, M., Felix, V., Feng, G., Kibbe, W.A.: Disease ontology: a backbone for disease semantic integration. Nucleic acids research **40**(D1), D940–D946 (2012)
30. Shrestha, R., Matteis, L., Skofic, M., Portugal, A., McLaren, G., Hyman, G., Arnaud, E.: Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice. Frontiers in Physiology **3**, 326 (2012)
31. Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y.: Pellet: A practical OWL-DL reasoner. J. Web Sem. **5**(2), 51–53 (2007)
32. Suchanek, F.M., Kasneci, G., Weikum, G.: YAGO: A Large Ontology from Wikipedia and WordNet. J. Web Sem. **6**(3), 203–217 (2008)
33. Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. Commun. ACM **57**(10), 78–85 (2014). https://doi.org/10.1145/2629489
34. Wilkinson, K., Sayers, C., Kuno, H.A., Reynolds, D., Ding, L.: Supporting scalable, persistent Semantic Web applications. IEEE Data Eng. Bull. **26**(4), 33–39 (2003)
35. Yu, C., Liu, X., Maia, J., Li, Y., Cao, T., Gao, Y., Song, Y., Goutam, R., Zhang, H., Yin, B., Li, Z.: COSMO: A Large-Scale E-commerce Common Sense Knowledge Generation and Serving System at Amazon. In: Companion of the 2024 International Conference on Management of Data (SIGMOD/PODS). pp. 148–160. ACM (2024). https://doi.org/10.1145/3626246.3653398