Representing cyberspace with the Basic Formal Ontology*

Giacomo De Colle[1,2]

[1] University at Buffalo, Buffalo, NY, USA
[2] National Center for Ontological Research, Buffalo, NY, USA

Abstract
Building a comprehensive ontological representation of cyberspace allows for integration of cyberspace data with data coming from other sources. This would prove extremely valuable, for example by integrating cybersecurity data with other intelligence and security data. In this paper I briefly present a simple pattern for representing cyberspace entities using the Basic Formal Ontology. The pattern makes use of the three-fold distinction between the information bearer, the information content and the specifically dependent continuant concretizing the information content.

Keywords
Domain ontologies, cyberspace, Basic Formal Ontology, Common Core Ontologies

1. Introduction

The Basic Formal Ontology (BFO) has emerged as a standard top-level ontology architecture and is currently employed in multiple initiatives for the purposes of data sharing and interoperability [1, 2]. BFO has been extended by multiple ontology projects in different areas, including for example the biomedical field [3], occupations [4], documents and information [5], the military domain [6], intelligence [7], industry [8], and social entities [9]. Moreover, the Common Core Ontologies (CCO) have recently emerged as a mid-level architecture widely adopted in the defense and intelligence community [10].

The domain of information processing and computation nevertheless remains a complex domain to represent ontologically. More specifically cyberspace, provisionally understood as the aggregate composed of computing artifacts, the information they process and the connections between such artifacts, is the object of interest for this paper. Ontologically representing cyberspace allows for integration of data about cyberspace itself with data about other parts of reality – for example, watchlist data or geographical location data. This would allow for implementation of ontologies in different computer science fields, starting from cybersecurity, and be of crucial support for big data analysis in intelligence operations.

Foundational studies in the ontology of cyberspace were already introduced in the BFO community [11], while more recently the CCO community has developed the Cyber ontology, which is currently part of an IEEE initiative, to represent cyberspace and the entities inhabiting

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands.
gdecolle@buffalo.edu (G. De Colle)

it [12, 13]. The aim of this short paper is to introduce a general, high-level pattern for the representation of information sharing and modification in the BFO and CCO community.

2. Concretization relations in BFO

The representation of information processing through BFO stands on the threefold distinction of information, information bearer and the qualities or dispositions of the bearer which concretize the information content. Take first as an example the familiar case of a book and the information stored in it. The book itself is a material entity, which acts as information bearer, and is represented in CCO by the class "Information Bearing Entity" (IBE). The pattern of ink on the book is instead a quality of the book, which allows it to convey information. The pattern of ink is, in BFO terms, a specifically dependent continuant, given that it requires the continued existence of *this* specific instance of book in order to remain in existence. The corresponding CCO class is called Information Quality Entity (IQE).

Finally, the information itself is the content of the book. The latter entity is, in BFO terms, a generically dependent continuant: it can continue to exist as long at least one entity is structured in such a way that it carries the information in question. This means that different copies of the same book can carry the same, numerically identical, information content, despite the copies and the ink patterns being different from one another. The corresponding CCO class is called Information Content Entity (ICE).

Notice that the link between the generically dependent continuant (the information) and the material entity (the information bearer) is given by the presence of the qualities of the material entity. The relation between the generically dependent continuant and the qualities is called a relation of concretization. The generically dependent continuant in question can only be interacted with by modifying the qualities and dispositions of a certain material entity. For example, changing the pattern of ink on a piece of paper makes it such that the pattern of ink on the piece of paper is now concretizing a different information content.

3. Modelling cyberspace in BFO – a test case

We can provisionally consider cyberspace as the aggregate of multiple computing artifacts and the information stored in and exchanged by these computing artifacts. The processes taking place in cyberspace are then mostly processes of information processing and sharing. An ontological representation of cyberspace will then make use of material entities (computing artifacts), ICEs (the information stored in the devices) and the patterns of qualities and dispositions which concretize these ICEs.

Image 1 shows a simplified model of a cyber-attack known as active packet sniffing. In one such attack, a malicious actor intercepts a data packet which is being transmitted over a network, reads the information in the packet, and changes it by inserting malicious code in the data packet. The ontological model of this process represents the process of active packet sniffing as modifying the quality pattern that concretizes the information content of the data packet. This information content is non-malicious at t1, and is replaced by malicious information at t3, after the process of modification has taken place at t2. The quality is borne by a data packet bearer, which in this case could be a part of a disk of a server. The quality itself could be identified with the pattern of electromagnetic energy that is stored on the disk.

The precision at which the material entity and quality are represented nevertheless doesn't need to be as detailed as the one discussed above. Unless the ontology in question is used, say, in the realm of industry manufacturing of computer hardware, we can represent bearers and their qualities at a higher level of granularity. For example, we can consider the whole server as data packet bearer, and we can introduce corresponding qualities identified by the type of information they are concretizing. In the case shown in figure 1, the quality concretizing data packet information is simply called a data packet quality.

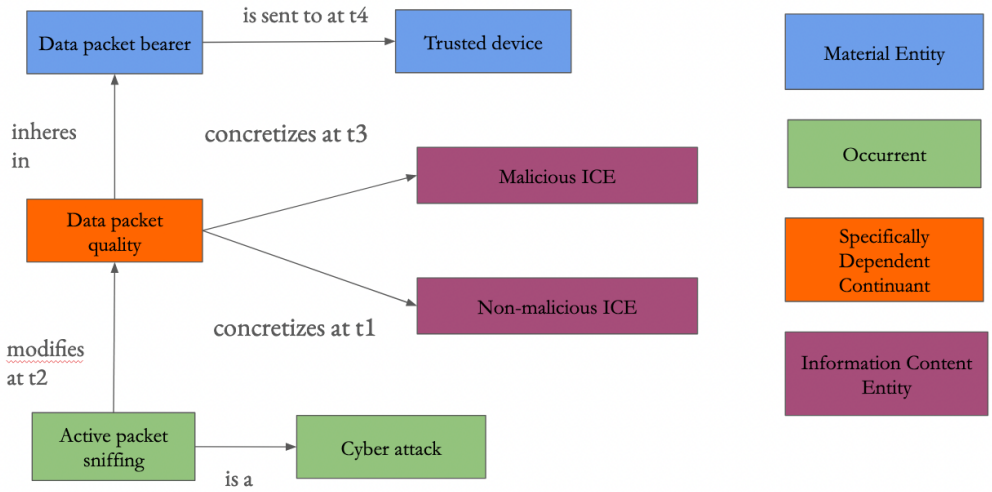

Figure 1: a simplified model of an active packet sniffing attack represented in BFO. The model makes use of the threefold distinction between information bearer, quality and content.

The threefold distinction introduced in BFO and extended in CCO has great representational power, insofar as it allows for the ontologist adopting it to represent different sides of the entities involved in information processing. Clearly, the full extent of the threefold distinction doesn't need to be employed in all situations. For example, in some cases it might be sufficient to say that a malicious actor is trying to access or modify the information (ICE) stored on a certain device. Nevertheless, BFO and CCO allow for a more fine-grained distinction which can be used to model tampering of code, direct modification of information stored on a specific hardware, as well as tracking detailed provenance of information, and information about the specific place and time where a certain cyber-attack has taken place.

4. Conclusion

BFO and CCO offer a well-developed and stable pattern which can be used to ontologically represent information processing, copying and sharing, as well as related operations in cyberspace. Adopting these top- and mid-level ontological layers provides a starting semantic layer which will be extremely useful in bridging data coming from different sources for interoperability of cybersecurity data.

Acknowledgements

The author wishes to acknowledge the insightful discussions with John Beverley, the members of the 2024 Ontology Engineering and Intelligence Analysis seminar at the University at Buffalo, and the members of the Cyber Ontology IEEE group.

References

[1] ISO/IEC 21838-1:2021. “Information Technology — Top-Level Ontologies (TLO) — Part 1: Requirements.” Accessed Feb 19, 2024.

[2] Arp R., Smith B., Spear A. Building Ontologies with Basic Formal Ontology, MIT Press, 2015.

[3] Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R. H., Shah, N., Whetzel, P. L., & Lewis, S. “The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration”. Nature Biotechnology, (2007), 25(11), 1251–1255.

[4] Beverley, J., Smith, S., Diller, M., Duncan, W.D., Zheng, J., Judkins, J.W., Hogan, W.R., McGill, R., Dooley, D.M., & He, Y. “The Occupation Ontology (OccO): Building a Bridge between Global Occupational Standards”. Proceedings International Workshop on Ontologies for Services and Society, July 17–20, (2023), Sherbrooke, Canada

[5] Smith Barry, Ceusters Werner. “Aboutness: Towards Foundations for the Information Artifact Ontology”. In Proceedings of the Sixth International Conference on Biomedical Ontology (ICBO), (2015).

[6] Morosoff, P., Rudnicki, R., Bryant, J., Farrell, R., & Smith, B. “Joint Doctrine Ontology: A Benchmark for Military Information Systems Interoperability”. Semantic Technologies for Intelligence, Defense, and Security, (2015).

[7] Mandrick, B., & Smith, B. “Philosophical foundations of intelligence collection and analysis: a defense of ontological realism”. Intelligence and National Security, 37, (2022): 809 - 819.

[8] Drobnjakovic, M., Kulvatunyou, B., Ameri, F. Will, C., and Smith, B., “The Industrial Ontologies Foundry (IOF) Core Ontology, Formal Ontologies Meet Industry (FOMI)”. Tarbes, FR, [online], 2022. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=935068 (Accessed February 17, 2024)

[9] Hicks A, Hanna J, Welch D, Brochhausen M, Hogan WR. “The ontology of medically related social entities: recent developments”. Journal of Biomed Semantics, (2016) Jul 12;7:47. doi: 10.1186/s13326-016-0087-8. PMID: 27406187; PMCID: PMC4942889.

[10] CUBRC. white Paper—"An Overview of the Common Core Ontologies”, 2019.

[11] Koepsell D. R., “The Ontology of Cyberspace”, 2000.

[12] Donohue B., Jensen M., Cox A. P., Rudnicki R., “A common core-based cyber ontology in support of cross-domain situational awareness.” Defense + Security (2018).

[13] IEEE Cyber Ontology Working Group, "Cyber Ontology Releases," IEEE Open Source. [Online]. Available: https://opensource.ieee.org/cyber-ontology-working-group/cyber-ontology-releases. [Accessed 25 February 2024].