

Operationalizing Ontologies in the Crime Domain

Rachel A. Mavrovich^{1,2}

¹ University at Buffalo, Buffalo, NY, USA

² National Center of Ontological Research, Buffalo, NY, USA

Abstract

Crime databases, especially between different organizations, are disparate. Gaps in crime databases hinder the accurate comprehension and integration of information for solving crimes. This project proposes the development of a general, domain-level **Crime Ontology**, that is grounded in the **Basic Formal Ontology (BFO)** and **Common Core Ontologies (CCO)**, to formalize and integrate crime related data. Unlike previous efforts that are either narrowly scoped or legally focused, this project highlights the distinct nature of criminal events. Requiring an independent ontological framework from that of legal ontologies. I discuss two important methodological approaches: top-down modelling and bottom-up analysis. I present two initial case studies, that of the kidnapping and murder of Italian Prime Minister Aldo Moro, which were used to test the use of design patterns in representing crime events and psychological assessments. The project now aims to develop a general structure akin to the Open Biomedical Ontologies (OBO) Foundry, but specifically for crime. A type of reference ontology that supports varied sub-domains like human trafficking and kidnapping.

Keywords

Crime Ontology, Domain Ontology, Basic Formal Ontology, Common Core Ontologies

1. Introduction and Motivations

Information in crime databases have gaps, especially between different crime solving organizations [1]. The problem with these gaps is that they prevent relevant users from getting the most accurate data to answer their questions [1]. One solution to this problem is a framework that provides structure and integration to large amounts of data. Applied Ontology utilizes logical and metaphysical principles along with a well-defined and controlled vocabulary of terms and relations to better facilitate interoperability between information systems. My work aims to give an applied ontological approach that will help formalize the information in these databases. I will be using a top-down method to start, where I begin with the broader concepts and make our way to narrower concepts. I will then supplement this with a bottom-up approach, meaning I then use instance data to see how that matches (or conflicts) with the top-down results. This will result in a unified approach that is flexible enough to serve different users and interests while still achieving semantic interoperability between data sets.

Various attempts have been made toward creating a crime ontology. On GitHub there are projects dedicated to specific criminal acts [2], crime analysis with a legal slant [3, 4], and cybercrime [5, 6, 7]. All these projects aim at trying to organize data related to crime for one purpose or the other, but none of them conforms to a known ISO standard, except for a list of criminal acts terms on GitHub [2]. Moreover, these GitHub projects either focus on legal users or have a very narrow scope. This leads to siloed information, where the projects serve only specific use cases without being interoperable lessening the overall effectiveness of each project. Therefore, it is important that I use a top-level ontology like BFO for two reasons. First, it is a recognized standard which ensures consistency and transparency of open-source data for relevant users. This matters because families and relevant users should be able to see how and where their data is being

Proceedings of FOIS 2025 Satellite events co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 10-12, 2025, Catania, Italy

✉ rmavrovi@buffalo.edu (R. A. Mavrovich)

ORCID 0009-0003-1971-7458 (R. A. Mavrovich)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

used. Second, it allows for the explicit modelling of top-level categories, ensuring semantic interoperability between systems [8]. Moreover, these GitHub projects either focus on legal users or have a very specific scope.

While law-focused users are important, these approaches lack a clear top-down framework. An ontology with a wider scope allows for capturing different types of relations [9]. This would allow for more comprehensive and wide ranging analysis to be made from the structured data within criminal activities and databases. The Basic Formal Ontology (BFO) is a top-level ontology architecture used in information sharing and system interoperability. It is currently being applied in various fields and subjects, including but not limited to biomedical sciences [9] and national defense and security [10,11]. Additionally, the Common Core Ontologies, the most widely used extension of BFO, is made up of 12 mid-level ontologies which are also being applied to these domains [10]. Extensions from BFO and CCO include domain level ontologies, which are about a specific subject but do not fall under the high-level or mid-level umbrella because their content is not as encompassing or broad as the former. The goal of this project is to create and apply a domain level ontology for crime that extends from BFO and CCO. There have been some successful domain level ontologies that none-the-less are overarching enough to capture many different use cases. Take the Open Biomedical Ontologies (OBO) Foundry, a unified open access collection of ontologies that is meant to capture data integration in the biological sciences [9]. It acts as an overarching structure for medical and biological investigation which allows other more specific ontologies to fit under it. Other examples include the Gene Ontology (GO) [12] and the Industrial Ontologies Foundry (IOF) Core Ontology [13]. The goal for the Crime Ontology is to follow this design to create a solid and general foundation where particular use cases can fit underneath.

With all that being said a question that might pop up is why not just represent these matters through a law ontology? Why a separate one for crime? Just like with the crime ontology, there have been efforts and ideas for a robust law ontology [14]. I will now briefly discuss why developing a crime ontology separately is important and that both law and crime should not be reduced into each other, no matter their close relationship. To start off, law deals with criminal activity from an abstract view and describes it as a history, not a present/ongoing moment. Take the term ‘perpetrator’ as an example. A law ontology would describe the role of perpetrator as being a part of an individual’s past relative to a crime, while a crime ontology would treat the role of perpetrator as active in relation to the crime. This is an important distinction as the use of one or the other dictates how data is interpreted and applied. This means that a law ontology may not be enough to capture pragmatic issues with crime data. It is also important to note that not all laws include crime, such as a legislative directive. I would not want my research question about crime to accidentally include non-criminal violations. For these reasons a comprehensive, overarching ontology dedicated to the domain of crime is much more useful to tackling criminal activity than solely relying on a law ontology. This is not to exclude the law and crime ontologies from informing each other or overlapping. That will necessarily happen; interoperability is one of the main strengths of using ontologies. Additionally, the actual instances which generate the data will likely overlap, such as with cold cases, cases that are re-opened and cases whose location is in legal environments.

2. The Project

The initial stages of this project started with a dry run to test whether a BFO/CCO conformant ontology could be implemented in a specific use case: the kidnapping and execution of ex-prime minister of Italy, Aldo Moro. A psychological profile was created by a subject matter expert [15] and this profile was then represented through an ontology. The first run mapped the profile as an Information Content Entity, which is a type of entity that is about something [16]. This way I could represent documents (in this case, the psychological profile) without having to model behavior and personality themselves. Modelling assessments will be an important method that I intend to continually use in the future when it comes to the tagging of cases and their artefacts. I presented

the first run of this project at the Cognitive Science Emerging Scholars Symposium at SUNY at Buffalo, 2025 in poster format.

The next run of the project included modelling a timeline of Aldo Moro's case. The timeline is from the known hour of his kidnapping to the known date when his body was located. Unlike a psychological profile the timeline is strictly data points about factual events, there was no modelling of the psychological assessments of a person. For this run, the Aldo Moro kidnapping Wikipedia [17] was used as a baseline. This run was to test a particular use case, not of an assessment document like for the psychological profile, but to see if I could capture a chronological arrangement of data points for a crime that spans more than one day. I presented the second run of this project at SUNY at Buffalo's Student Showcase for Academic Excellence, 2025.

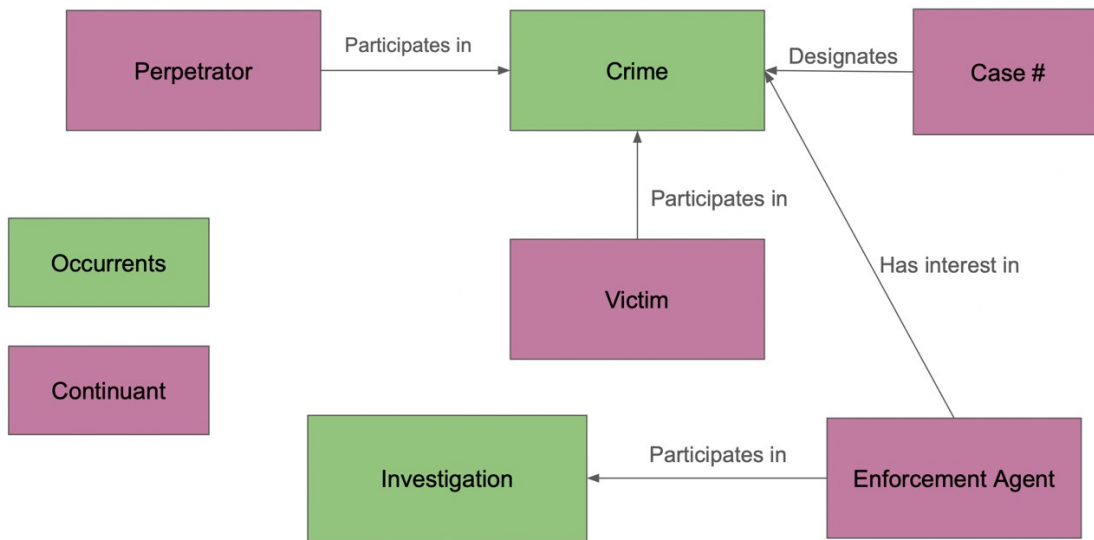


Figure 1: Design pattern of a general crime representation.

The crime ontology project has now moved to the wider scope than addressed in the introduction. Instead of sticking to an instance, I am tackling the general domain level ontology I originally planned for. Figure 1 shows a design pattern of general crime representation. I, in collaboration with others, am developing and refining high level terms, processes, and relations and then connecting this through design patterns. Design patterns are used to represent condensed elements of an ontology. Their purpose is for subject matter experts and relevant audiences to quickly grasp what the ontology and data is doing. In terms of collecting data, I am currently working on doing data ingestion for gathering and applying sources. Lastly, a current subject matter expert (SME) list includes ontologists, psychologists and data ingestion experts. An expanded SME list will be provided in the Target State section of this paper.

3. Target State

The goals of the crime ontology for the next 6 months are as follows: (1) conduct analysis on compiled data with Subject Matter Experts (SME) and relevant researchers; (2) analyze information provided by SMEs; (3) generate the rudimentary structure of general domain level crime ontology. Future SME'S will include law enforcement officials, prisoners, survivors, data analysts, and relevant researchers. It is to be noted that relevant researchers who are subject matter experts on crime will be part of societies like the National Centre for Missing and Exploited Children (NCMEC), and researchers at the Cold Case Analysis Center (CCAC) at University at Albany. NCMEC is a non-profit 501(c)(3) corporation that helps find missing children and helps in

preventing child exploitation [18]. CCAC is a part of the School of Criminal Justice at University at Albany, they aid families and agencies in case investigation [19]. Relevant researchers who are analyzing the information provided by the SMEs will be CITI and facilitator certified. This is to ensure all narrative analyses and meta-analyses are not only done by qualified researchers but are done respectfully and ethically. This is important because I am dealing with sensitive human source topics, and I wish to take seriously the difficulties and weight of these topics. When tackling the domain of crime, I think it is of importance to get various perspectives, from those who are meant to enforce the law and those who are said to have violated it.

The ontological methods that will be applied in the process of this project will be as follows. I will first gather data from our sources using, but not limited to, interview and data ingestion techniques. Meta-analysis and narrative analysis will be conducted on the data to synthesize and identify any gaps. Design patterns and high-level terms will then be produced to test the data. The crime ontology will extend from BFO and CCO frameworks, and I will work from a top-down approach.

Throughout this paper I have made clear this project's intention is to develop a general ontology of crime, one that can be reused and applied generally to different instances, the goal is to eventually have something like the OBO Foundry albeit for crime. For example, a few ontologies that could be developed are those for human trafficking and kidnapping. Where time is of the essence, so large and fast changing amounts of data need to be accurately and swiftly organized. Being able to pull information quickly from a reliable structured source ensures accurate evaluation and good response to these types of criminal activity. This is important because, while the structure of the ontology may be good, it is important to have a pragmatic goal of where this ontological structure may be applied.

4. Conclusion

In an increasingly interconnected world, there is a need for structured, reliable and interoperable crime data systems. This paper has outlined the motivation, current progress, and future endeavors for the development of a general domain-level crime ontology, which will be grounded in the Basic Formal Ontology (BFO) and Common Core Ontologies (CCO). This project aims to address the gaps in existing crime databases and support better data integration, transparency, and analysis for different stakeholders. This paper addressed the important philosophical and practical distinctions between that of abstract legal processes and the active concrete processes of criminal events, which showed the need for a standalone crime ontology. Building a general and extensive ontology lays the foundation for more specific ontologies, as seen in the example of the successful OBO Foundry. Utilizing the insights and expertise from a wide range of subject matter experts—from law enforcement officials to survivors to researchers—this project strives not just for thoroughness but for ethical responsibility in representing complex human experiences. This is especially important in sensitive and rapidly evolving cases such as kidnapping and human trafficking. The crime ontology project aims, through better data management and interoperability, to become a shared option for researchers, analysts, and users who are working to understand, prevent, and respond to criminal activity.

Acknowledgments

I would like to give thanks and acknowledgement to the following people: John Beverley for his continued mentorship, support and discussions; Giacomo De Colle for his major support both emotionally and in his insightful discussions and comments; Noor Al-Khaouli for providing background and being my subject matter expert in the first run of this project; Jeremy Ravenel in helping develop the draft list of I processes for the crime ontology and helping with data ingestion; and Toby Friend, Matthew Jones and Sean Kindya for their helpful comments in the early drafting of this paper.

Declaration on Generative AI

The author has not employed any generative AI tools.

References

- [1] Centre for Public Data. (n.d.). Justice data gaps report. Centre for Public Data. <https://static1.squarespace.com/static/5ee7a7d964aee7e5c507900/t/64230b79130cdc7e4b83930f/1680018298114/CFPD+justice+data+gaps+report.pdf>
- [2] Sculley, A. (n.d.). Crime ontology. GitHub. <https://github.com/avsculley/CrimeOntology>
- [3] LiDaKrA. (n.d.). Ontology. GitHub. <https://github.com/LiDaKrA/Ontology>
- [4] Tripathi, A.(n.d.). Knowledge graph. GitHub. https://github.com/AnjaneyaTripathi/knowledge_graph
- [5] Naqvi, S.A.M.(n.d.). CyberCrimeAttackTypesOntology. GitHub. <https://github.com/syedamoaddatnaqvi-141/CyberCrimeAttackTypesOntology>
- [6] Naqvi, S. A. M. (n.d.). CyberCrime Ontology. GitHub. <https://github.com/syedamoaddatnaqvi-141/CyberCrimeOntology>
- [7] Razia.(n.d.). Cybercrime-Attacks-Ontology. GitHub. <https://github.com/anam236razaq/CyberCrime-Attacks-Ontology>
- [8] International Organization for Standardization. (2021). Information technology – Top-level ontologies (TLO) – Part 1: Requirements (ISO/IEC 21838-1:2021).
- [9] Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., OBI Consortium, Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.A., Scheuermann, R. H., Shah, N., Whetzel, P. L., & Lewis, S. (2007). The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25(11), 1251–1255. <https://doi.org/10.1038/nbt1346>
- [10] Jensen, M., Oliveira, R., & Guizzardi, G. (2024). The Common Core Ontologies [Conference paper]. FOIS 2024. <https://www.utwente.nl/en/eemcs/fois2024/resources/papers/jensen-et-al-the-common-core-ontologies.pdf>
- [11] Mandrick, B., & Smith, B. (2022). Philosophical foundations of intelligence collection and analysis: A defense of ontological realism. *Intelligence and National Security*. <https://doi.org/10.1080/02684527.2022.2076330>
- [12] 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., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, 25(1), 25–29. <https://doi.org/10.1038/75556>
- [13] Industrial Ontology Foundry (IOF). (n.d.). Publication. National Institute of Standards and Technology. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=935068
- [14] Oliveira, R., Almeida, J., & Guizzardi, G. (2024). LEGION-IDS: Legal Interoperability Ontology for International Data Spaces [Conference paper]. FOIS2024. <https://www.utwente.nl/en/eemcs/fois2024/resources/papers/oliveira-et-al-legion-ids-legal-interoperability-ontology-for-international-data-spaces.pdf>
- [15] Al-Khaouli, N. (2024, December 15). Psychological profile of Aldo Moro (PHI 489 course paper). [Unpublished manuscript].
- [16] Arp, R., Smith, B., & Spear, A. D. (2015). Building ontologies with Basic Formal Ontology. MIT Press.
- [17] Wikipedia contributors. (n.d.). Kidnapping and murder of Aldo Moro. Wikipedia. Retrieved April 27, 2025, from https://en.wikipedia.org/wiki/Kidnapping_and_murder_of_Aldo_Moro
- [18] National Center for Missing & Exploited Children. (n.d.). Home. <https://www.missingkids.org/home>

[19] University at Albany, School of Criminal Justice. (n.d.). Cold Case Analysis Center.
<https://www.albany.edu/scj/cold-case-analysis-center>