Ontologies for the Behavioural and Social Sciences: Opportunities and Challenges

Janna Hastings[1,2], Robert West[3], Susan Michie[2], Caitlin Notley[4] and Sharon Cox[3]

[1]*Department of Computer Science, Otto von Guericke Universität Magdeburg, Germany*
[2]*Department of Clinical, Educational and Health Psychology, University College London, UK*
[3]*Department of Behavioural Science and Health, University College London, UK*
[4]*Norwich Medical School, University of East Anglia, Norwich, UK*

Abstract

This report summarises the presentations and discussions that took place during the first Ontologies for Research in the Behavioural and Social Sciences (OntoBess) workshop, held in Bozen-Bolzano and virtually on the 18th September 2021. The workshop highlighted the relevance of ontologies to address overarching challenges in the behavioural and social sciences, including evidence synthesis in the face of conflicting findings, challenging definitions for key entities, and the need to incorporate and advance theory. It also highlighted the perspectives and resources that the behavioural and social sciences can bring to the ontology development community.

Keywords

ontologies, behavioural science, social sciences, ontobess

1. Introduction

Use of ontologies has become widespread in the biomedical sciences and in many other domains [1, 2]. However, they have not yet been widely adopted within the social and behavioural sciences. Recently, interest in the development and use of ontologies for advancing the behavioural and social sciences has been growing (e.g. [3, 4]. Recent applications of ontologies within the social and behavioural sciences include the Human Behaviour Change Project [5, 6] and the Addiction Ontology (AddictO; [7]). Relatedly, an ontology that addresses social aspects within the biomedical domain is the Ontology for Medically Related Social Entities (OMRSE; [8]).

Alongside the International Conference on Biomedical Ontologies and as part of the Bolzano Summer of Knowledge, a workshop was held addressing the use of ontologies to advance research for the behavioural and social sciences. The workshop aimed to provide a forum for the discussion of all aspects related to the use and development of ontologies for research in the social and behavioural sciences, and for the description of social, behavioural and economic variables in studies involving humans or in population studies. Specific topics that were of interest included the representation of population and personal attributes, behavioural and

socioeconomic attributes, measurement, research methods and formal descriptions of social research studies, relating social theory and ontologies, and the adoption of upper-level and mid-level ontologies for the behavioural and social sciences. An important additional objective of the workshop was to initiate the formation of a community of practice and exchange around the development of ontologies for the social and behavioural sciences.

The workshop presentations and accompanying discussions covered diverse aspects within this overarching theme. The next section discusses these presentations in the context of three overarching topics: (1) ontologies for evidence synthesis, (2) challenging definitions for the behavioural and social science domains, and (3) theories and ontologies.

2. Current Status and Topics

2.1. Ontologies for Evidence Synthesis

One of the longstanding challenges in the behavioural and social sciences is the need for robust, reliable and up-to-date synthesis of the evidence base in order to answer questions, guide future research, and provide guidance to policymakers and practitioners. This was also one of the central applications areas where ontologies are foreseen to make a difference to these disciplines, which was reflected in the topics presented and discussed at the workshop.

Professor Susan Michie gave the keynote for the workshop, during which she described the development of the Behaviour Change Intervention Ontology (BCIO, [4, 9]) and its use for the annotation of intervention literature in support of the automation of evidence synthesis for the behaviour change domain in the context of the Human Behaviour Change Project (HBCP). The HBCP is a proof-of-principle project showcasing the use of artificial intelligence technologies for the management of behavioural intervention evidence, incorporating both an automated information extraction part and an automated prediction part, both of which are organised around the ontology to provide the semantic structure for the domain. The BCIO consists of an upper level and several largely independent lower-level modules such as for behaviour, population attributes, settings [10] and modes of delivery [11], each of which consists of several or even a few hundred classes. Each class may be used in annotations if it is mentioned in the context of the text describing an intervention in the literature.

Professor Caitlin Notley spoke on the motivation for the development and use of ontologies arising from recent experiences performing systematic reviews of the evidence in the context of the e-cigarette literature. E-cigarettes and related products have been the subject of intense debate in recent years, with their relative risks, harms and benefits for reducing the burden of smoking-associated disease through supporting smoking cessation being subject to differing interpretations in the evidence base (e.g. [12]). Professor Notley illustrated how many challenges in synthesis in this evidence based related to vague or poorly defined groupings in the literature, and how even central and widely used classes such as "youth" as an age-defined group of individuals, were poorly and incompatibly defined across different studies, leading for example to challenges in synthesising evidence relating to youth e-cigarette use and the adoption of flavours [13].

Continuing the theme of ontologies for the purpose of evidence synthesis, Professor Robert West followed next by giving a detailed look at the specification of behavioural outcomes in

behavioural intervention descriptions, noting all the different factors that need to be detailed and offering a preliminary semantic formalisation of an outcome specification that would enable automated comparison of outcomes.

2.2. Challenging Definitions in the Behavioural and Social Science Domains

Professors Mark Fox and Kate Ruff presented an ontology for specifying and defining the social and environmental impacts of social purpose organisations, the Common Impact Data Standard (CIDS), targeting the application area of enabling standardised reporting of and comparisons between such impacts. This ontology is part of a wider effort to enable an infrastructure for maximising the impact of social organisations through better understanding the pathways from organisational activities to real-world outcomes. The CIDS aims to serve as an integrative framework that allows different ways of measuring impact to be interconverted and thus rendered comparable, in part through the consensus-driven definition of overarching dimensions of impact.

On a different topic but also with policy relevance, Dr Adrien Barton spoke about the ontological definition and classification of nudging, a well-known and yet in some ways controversial behaviour change technique that involves changing the organisation of the environment in order to influence people to behave in ways that are beneficial to their health. Dr Barton emphasised the need to distinguish between nudges themselves, which he defines as environmental dispositions, and the processes that give rise to those nudges and that activate them. These entities are all classified beneath Basic Formal Ontology (BFO; [1]), a top-level ontology widely used in the biomedical sciences among other domains. The classification of the entities in different ontologies beneath a common top-level ontology, or at least in terms of inter-convertible top-level ontologies, is helpful in enabling the ontologies to be used together. A discussion ensued about which top-level ontology might be best suited to representing entities in the behavioural and social sciences, and the criteria by which top-level ontologies might be selected, evaluated and compared – itself a fruitful direction of additional research at the interface between the behavioural and social sciences and the applied ontology community.

2.3. Theories and Ontologies

Professor Michie also discussed the relationship between theories, which are an integral component of the behavioural sciences, and ontologies [14]. Theory proposition, development and refinement is an integral part of the advancement of the behavioural and social sciences, and as such it makes sense to ask how ontologies and theories are related and how they differ in scope and purpose. Theories propose testable hypotheses and explanations for what the entities are in a given domain and how they are related. Their purpose is to advance the understanding of the domain. However, they may speak to only a subset of the entities in the domain or seek to explain just a partial or specific phenomenon. Ontologies, on the other hand, do not represent testable hypotheses as such, nor seek to provide explanations for phenomena. They aim to provide a shareable and re-usable representation of the entities in the domain in as theory-neutral a way as possible in order to enable consistent descriptions in data and text across many different contexts, as a kind of interlingua. Thus, it even makes sense to try to link theories to ontologies

in order to make semantic-driven theory integration and comparison possible, which is exactly what the Ontology-Based Modelling System [15, 16], also presented by Professor Michie, aims to enable just such a connection between theory constructs and ontology classes, with a view to systematically comparing theories about behaviour and behaviour change.

Continuing the theme of ontology to theory mappings, for the last presentation of the workshop, Irshad Ally and Professor Werner Ceusters discussed the findings arising from their efforts to map the constructs from motivational learning theories to BFO classes, with outcomes of the study presented both in terms of the learning exercise involved in coming to terms with the distinctions between upper-level entities in the BFO top-level ontology, and in terms of the challenges with unambiguously interpreting and defining the entities from such theories based on their available documentation.

3. Opportunities and Challenges

It became evident through a multitude of the discussions that took place during the session that there is a clear need for the advantages offered by ontologies in these domains: clarity of reporting and exchange, interpretation and integration of data and evidence, and clearer semantics for theories towards cumulative science. However, several challenges were also highlighted, including barriers to the understanding of the nature, purpose and scope of ontologies within the communities of research and practice in these domains, a shortage of available materials for wider education and dissemination, a need for tools and resources supporting the development of ontologies for these domains, and the lack of standardised infrastructures for sharing relevant content.

In particular, the behavioral and social science domains are characterized by a very large number of distinct semantic types – individuals, populations, interventions, the full range and scope of behaviors, population and personal attributes, location attributes, and moreover all the semantic types from biomedicine which may also have relevance in some intervention contexts. Added to this semantic complexity, there are very heterogeneous ways in which such semantic types are described in the literature, are theorised about, are defined and operationalised, and contexts in which they are applied. Moreover, the theoretical frameworks in which such research is conducted may introduce technical terminology from different theoretical backgrounds which may be difficult for a language model to disentangle, as there are a wide range of different theories and constructs in use within behavioral science more broadly, and as yet no systematic way to map their overlaps and commonalities.

A further challenge is that there is a lack of an organized repository for all literature or data relevant to the behavioral sciences. Relevant literature spans several different disciplines including psychology, the social sciences, and economics as well as the more medically-oriented branches e.g. behavioral medicine. These disciplines may have quite different publication practices, and in some of these even discovering the existence of the relevant literature may be subject to a "paywall". There is also limited publicly available data aside from text - there is significantly less publicly available structured and ontology-annotated data, and any research data involving humans is subject to privacy and ethical concerns that may mean that this situation will not change dramatically in the future.

Finally, a particular challenge for the development and adoption of ontologies for the behavioural and social sciences is, when viewed from a different perspective, also equally valid as an important argument for their adoption: the terminology that is used in scientific studies in the behavioral sciences overlaps with informal and colloquial usages of vocabulary and terminology and are therefore frequently polysemous. While in the biomedical sciences, technical vocabulary includes gene names, drug names and disease names, each of which is strongly semantically typed and forms a vocabulary that for the most part is distinct from the vocabularies used in other sciences and domains, in the behavioral and social sciences the terminology that is used reflects human activities and concerns, social structures and patterns of behaviour. These entities are also referred to frequently within other disciplines, as many studies that involve humans, even biomedical studies of drugs and diseases, will ultimately need to refer to population groups, activities and behaviors.

4. Conclusions and Next Steps

The workshop aimed to initiate conversations and exchanges amongst the broader community of persons developing and using ontologies for the behavioural and social sciences. The workshop underlined the importance of ontologies across a wide range of applications within the behavioural and social sciences, and emphasised that while the subject matter might be in some ways more challenging than the more narrowly focused subject matter of biomedical ontologies, the time is ripe to harness synergies and opportunities to advance this integrative field. In hybrid format (in-person and online), the workshop was well attended, with more than 30 participants overall, indicating that there is at least the beginnings of a robust community to draw from. It was clear, however, that the hybrid workshop format provided a restricted opportunity for discussion and sharing of experiences as compared to that offered by face-to-face workshops, and it was agreed that an important future step will be to try to hold an in-person follow-up workshop as soon as it is safe to do so given the global pandemic situation.

Acknowledgments

References

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

[2] M. A. Haendel, C. G. Chute, P. N. Robinson, Classification, Ontology, and Precision Medicine, New England Journal of Medicine 379 (2018) 1452–1462. URL: http://www.nejm.org/doi/10.1056/NEJMra1615014. doi:10.1056/NEJMra1615014.

[3] K. R. Larsen, S. Michie, E. B. Hekler, B. Gibson, D. Spruijt-Metz, D. Ahern, H. Cole-Lewis, R. J. B. Ellis, B. Hesse, R. P. Moser, J. Yi, Behavior change interventions: the potential of

ontologies for advancing science and practice, Journal of Behavioral Medicine 40 (2017) 6–22. doi:10.1007/s10865-016-9768-0.

[4] S. Michie, R. West, A. N. Finnerty, E. Norris, A. J. Wright, M. M. Marques, M. Johnston, M. P. Kelly, J. Thomas, J. Hastings, Representation of behaviour change interventions and their evaluation: Development of the Upper Level of the Behaviour Change Intervention Ontology, Wellcome Open Research 5 (2020) 123. URL: https://wellcomeopenresearch.org/articles/5-123/v1. doi:10.12688/wellcomeopenres.15902.1.

[5] S. Michie, J. Thomas, M. Johnston, P. M. Aonghusa, J. Shawe-Taylor, M. P. Kelly, L. A. Deleris, A. N. Finnerty, M. M. Marques, E. Norris, A. O'Mara-Eves, R. West, The Human Behaviour-Change Project: harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation, Implementation Science 12 (2017) 121. URL: https://doi.org/10.1186/s13012-017-0641-5. doi:10.1186/s13012-017-0641-5.

[6] S. Michie, J. Thomas, P. Mac Aonghusa, R. West, M. Johnston, M. P. Kelly, J. Shawe-Taylor, J. Hastings, F. Bonin, A. O'Mara-Eves, The Human Behaviour-Change Project: An artificial intelligence system to answer questions about changing behaviour, Wellcome Open Research 5 (2020) 122. URL: https://wellcomeopenresearch.org/articles/5-122/v1. doi:10.12688/wellcomeopenres.15900.1.

[7] J. Hastings, S. Cox, R. West, C. Notley, Addiction Ontology: Applying Basic Formal Ontology in the Addiction domain, Qeios (2020). URL: https://www.qeios.com/read/HZHJIP. doi:10.32388/HZHJIP.

[8] A. Hicks, J. Hanna, D. Welch, M. Brochhausen, W. R. Hogan, The ontology of medically related social entities: recent developments, Journal of Biomedical Semantics 7 (2016) 47. URL: https://doi.org/10.1186/s13326-016-0087-8. doi:10.1186/s13326-016-0087-8.

[9] A. J. Wright, E. Norris, A. N. Finnerty, M. M. Marques, M. Johnston, M. P. Kelly, J. Hastings, R. West, S. Michie, Ontologies relevant to behaviour change interventions: a method for their development, Wellcome Open Research 5 (2020) 126. URL: https://wellcomeopenresearch.org/articles/5-126/v2. doi:10.12688/wellcomeopenres.15908.2.

[10] E. Norris, M. M. Marques, A. N. Finnerty, A. J. Wright, R. West, J. Hastings, P. Williams, R. N. Carey, M. P. Kelly, M. Johnston, S. Michie, Development of an Intervention Setting Ontology for behaviour change: Specifying where interventions take place (2020). URL: http://bura.brunel.ac.uk/handle/2438/20971. doi:10.12688/wellcomeopenres.15904.1, accepted: 2020-06-12T20:54:13Z Publisher: F1000 Research Ltd.

[11] M. M. Marques, R. N. Carey, E. Norris, F. Evans, A. N. Finnerty, J. Hastings, E. Jenkins, M. Johnston, R. West, S. Michie, Delivering Behaviour Change Interventions: Development of a Mode of Delivery Ontology, Wellcome Open Research 5 (2020) 125. URL: https://wellcomeopenresearch.org/articles/5-125/v1. doi:10.12688/wellcomeopenres.15906.1.

[12] J. Hartmann-Boyce, H. McRobbie, A. R. Butler, N. Lindson, C. Bullen, R. Begh, A. Theodoulou, C. Notley, N. A. Rigotti, T. Turner, T. R. Fanshawe, P. Hajek, Electronic cigarettes for smoking cessation, Cochrane Database of Systematic Reviews (2021). URL: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010216.pub6/full. doi:10.1002/14651858.CD010216.pub6, publisher: John Wiley & Sons, Ltd.

[13] C. Notley, S. Gentry, S. Cox, M. Dockrell, M. Havill, A. S. Attwood, M. Smith, M. Munafo, Youth Use of E-Liquid Flavours – A systematic review exploring patterns of use of e liquid flavours and associations with continued vaping, tobacco smoking uptake,

or cessation., Addiction (2021). URL: https://research-portal.uea.ac.uk/en/publications/youth-use-of-e-liquid-flavours-a-systematic-review-exploring-patt, publisher: Wiley.

[14] J. Hastings, S. Michie, M. Johnston, Theory and ontology in behavioural science, Nature Human Behaviour 4 (2020) 226–226. URL: http://www.nature.com/articles/s41562-020-0826-9. doi:10.1038/s41562-020-0826-9.

[15] R. West, C. A. Godinho, L. C. Bohlen, R. N. Carey, J. Hastings, C. E. Lefevre, S. Michie, Development of a formal system for representing behaviour-change theories, Nature Human Behaviour 3 (2019) 526–536. URL: https://www.nature.com/articles/s41562-019-0561-2. doi:10.1038/s41562-019-0561-2, number: 5 Publisher: Nature Publishing Group.

[16] J. Hale, J. Hastings, R. West, C. E. Lefevre, A. Direito, L. C. Bohlen, C. Godinho, N. Anderson, S. Zink, H. Groarke, S. Michie, An ontology-based modelling system (OBMS) for representing behaviour change theories applied to 76 theories, Wellcome Open Research 5 (2020) 177. URL: https://wellcomeopenresearch.org/articles/5-177/v1. doi:10.12688/wellcomeopenres.16121.1.