

Through the validation, the ontology can support typical search queries using SPARQL.

An essential item of future work is to incorporate graph-based analytics into ODabler, and, further, explore applications of this technology. Although the exact capabilities of a graph-based ODabler, and more generally of graph-based datacenter ODA, are largely unknown, this line of future work will bring evidence of whether our exploration in Section 4.3 is correct and could result in a new way of understanding datacenters.

Future work will also include extensions of the current ontology, such as more comprehensive relations between different entities. We will also add more individuals to the ontology and conduct graph-based experiments to see if the ontology can help better understand the datacenter operation. The current draft ontology is available online at <https://github.com/am-i-helpful/hpc-ontology-modeller>.

ACKNOWLEDGEMENT

We thank the Dutch National Supercomputing Center SURF for providing the data. We thank the China Scholarship Council (CSC) for supporting Xiaoyu Chu. We thank the support of Netherlands-funded projects NWO OffSense and GFP 6G FNS, and EU-funded projects MCSA-RISE Cloudstars and Horizon Graph-Massivizer.

REFERENCES

- [1] Gabriel G. Castañé, Huanhuan Xiong, Dapeng Dong, and John P. Morrison. 2018. An ontology for heterogeneous resources management interoperability and HPC in the cloud. *Future Gener. Comput. Syst.* 88 (2018), 373–384. <https://doi.org/10.1016/j.future.2018.05.086>
- [2] Óscar Corcho, David Chaves-Fraga, Jhon Toledo, Julián Arenas-Guerrero, Carlos Badenes-Olmedo, Mingxue Wang, Hu Peng, Nicholas Burrett, Jose Mora, and Puchao Zhang. 2021. A High-Level Ontology Network for ICT Infrastructures. In *The Semantic Web - ISWC 2021 - 20th International Semantic Web Conference, ISWC 2021, Virtual Event, October 24-28, 2021, Proceedings (Lecture Notes in Computer Science, Vol. 12922)*, Andreas Hotho, Eva Blomqvist, Stefan Dietze, Achille Fokoue, Ying Ding, Payam M. Barnaghi, Armin Haller, Mauro Dragoni, and Harith Alani (Eds.). Springer, 446–462. https://doi.org/10.1007/978-3-030-88361-4_26
- [3] Yu Deng, Ronnie Sarkar, HariGovind V. Ramasamy, Rafah Hosn, and Ruchi Mahindru. 2013. An Ontology-Based Framework for Model-Driven Analysis of Situations in Data Centers. In *2013 IEEE International Conference on Services Computing, Santa Clara, CA, USA, June 28 - July 3, 2013*. IEEE Computer Society, 288–295. <https://doi.org/10.1109/SCC.2013.98>
- [4] Cory Andrew Henson, Holger Neuhaus, Amit P Sheth, Krishnaprasad Thirunarayan, and Rajkumar Buyya. 2009. An ontological representation of time series observations on the semantic sensor web. (2009).
- [5] Fahim T Imam. 2016. Application of ontologies in cloud computing: The state-of-the-art. *arXiv preprint arXiv:1610.02333* (2016).
- [6] Alexandru Iosup, Fernando Kuipers, Ana Lucia Varbanescu, Paola Grosso, Animesh Trivedi, Jan S. Rellermeyer, Lin Wang, Alexandru Uta, and Francesco Regazzoni. 2022. Future Computer Systems and Networking Research in the Netherlands: A Manifesto. *CoRR abs/2206.03259* (2022). <https://doi.org/10.48550/ARXIV.2206.03259>
- [7] Guodong Jin, Xiyang Feng, Ziyi Chen, Chang Liu, and Semih Salihoglu. 2023. KÜZU Graph Database Management System. In *13th Conference on Innovative Data Systems Research, CIDR 2023, Amsterdam, The Netherlands, January 8-11, 2023*. www.cidrdb.org. <https://www.cidrdb.org/cidr2023/papers/p48-jin.pdf>
- [8] Jean-Baptiste Lamy. 2017. Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies. *Artificial intelligence in medicine* 80 (2017), 11–28.
- [9] Chunhua Liao, Pei-Hung Lin, Gaurav Verma, Tristan Vanderbruggen, Murali Emani, Zifan Nan, and Xipeng Shen. 2021. HPC Ontology: Towards a Unified Ontology for Managing Training Datasets and AI Models for High-Performance Computing. In *IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments, MLHPC@SC 2021, St. Louis, MO, USA, November 15, 2021*. IEEE, 69–80. <https://doi.org/10.1109/MLHPC54614.2021.00012>
- [10] Hongzi Mao, Malte Schwarzkopf, Shaileshh Bojja Venkatakrisnan, Zili Meng, and Mohammad Alizadeh. 2019. Learning scheduling algorithms for data processing clusters. In *Proceedings of the ACM Special Interest Group on Data Communication, SIGCOMM 2019, Beijing, China, August 19-23, 2019*, Jianping Wu and Wendy Hall (Eds.). ACM, 270–288. <https://doi.org/10.1145/3341302.3342080>
- [11] Fabian Mastenbroek, Georgios Andreadis, Soufiane Jounaid, Wenchen Lai, Jacob Burley, Jaro Bosch, Erwin Van Eyk, Laurens Versluis, Vincent Van Beek, and Alexandru Iosup. 2021. OpenDC 2.0: Convenient modeling and simulation of emerging technologies in cloud datacenters. In *2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. IEEE, 455–464.
- [12] Alessio Netti. 2022. *Holistic and Portable Operational Data Analytics on Production HPC Systems*. Ph. D. Dissertation. Technische Universität München.
- [13] Natalya F Noy, Deborah L McGuinness, et al. 2001. Ontology development 101: A guide to creating your first ontology.
- [14] Sherif Sakr, Angela Bonifati, Hannes Voigt, Alexandru Iosup, Khaled Ammar, Renzo Angles, Walid G. Aref, Marcelo Arenas, Maciej Besta, Peter A. Boncz, Khuzaima Daudjee, Emanuele Della Valle, Stefania Dumbrava, Olaf Hartig, Bernhard Hashlhofer, Tim Hegeman, Jan Hidders, Katja Hose, Adriana Iamnitchi, Vasiliki Kalavri, Hugo Kapp, Wim Martens, M. Tamer Özsu, Eric Peukert, Stefan Plantikow, Mohamed Ragab, Matei Ripeanu, Semih Salihoglu, Christian Schulz, Petra Selmer, Juan F. Sequeda, Joshua Shinavier, Gábor Szárnyas, Riccardo Tomasini, Antonino Tumeo, Alexandru Uta, Ana Lucia Varbanescu, Hsiang-Yun Wu, Nikolay Yakovets, Da Yan, and Eiko Yoneki. 2021. The future is big graphs: a community view on graph processing systems. *Commun. ACM* 64, 9 (2021), 62–71. <https://doi.org/10.1145/3434642>
- [15] Siddharth Samsi, Matthew L. Weiss, David Bestor, Baolin Li, Michael Jones, Albert Reuther, Daniel Edelman, William Arcand, Chansup Byun, John Holodnack, Matthew Hubbell, Jeremy Kepner, Anna Klein, Joseph McDonald, Adam Michaleas, Peter Michaleas, Lauren Milechin, Julia S. Mullen, Charles Yee, Benjamin Price, Andrew Prout, Antonio Rosa, Allan Vanterpool, Lindsey McEvoy, Anson Cheng, Devesh Tiwari, and Vijay Gadepally. 2021. The MIT Supercloud Dataset. In *2021 IEEE High Performance Extreme Computing Conference, HPEC 2021, Waltham, MA, USA, September 20-24, 2021*. IEEE, 1–8. <https://doi.org/10.1109/HPEC49654.2021.9622850>
- [16] Axel Tenschert. 2016. Ontology matching in a distributed environment. (2016).
- [17] Lamia Youseff, Maria Butrico, and Dilma Da Silva. 2008. Toward a unified ontology of cloud computing. In *2008 Grid Computing Environments Workshop*. IEEE, 1–10.
- [18] Y Zhao, C Liao, and X Shen. 2017. *An Infrastructure for HPC Knowledge Sharing and Reuse*. Technical Report. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States).
- [19] Aolong Zhou, Kaijun Ren, Xiaoyong Li, Wen Zhang, and Xiaoli Ren. 2019. Building quick resource index list using wordnet and high-performance computing resource ontology towards efficient resource discovery. In *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. IEEE, 885–892.