

Онтологический Саммит 2021.

Коммюнике: генерация и гармонизация онтологий

Ontology Summit 2021. Communique: «Ontology Generation and Harmonization»¹

Ноябрь 2021

Кен Баклавски, Северо-Восточный университет, Бостон, Массачусетс, США (kenbaclawski@gmail.com)

Майкл Беннет, Hypercube Limited, Лондон, Великобритания (mbennett@hypercube.co.uk)

Гэри Берг-Кросс, соруководитель ESIP по семантической гармонизации (gbergcross@gmail.com)

Лея Дикерсон, Счётная палата правительства, Вашингтон, округ Колумбия, США (ldickers@gmail.com);

Тодд Шнайдер, инженерная семантика, Фэрфакс, Вирджиния, США (tschneider@engineeringsemantics.com);

Селья Сеппала, Университетский колледж Корка, Ирландия (selja.seppala@ucc.ie);

Рави Шарма, старший архитектор предприятия, Elk Grove, Калифорния, США (dravisharma@gmail.com)

Рам Д. Шрирам, Национальный институт стандартов и технологий, Мэриленд, США (sriram@nist.gov)

Андреа Вестеринен, OntoInsights LLC, Элтон, Мэриленд, США (arwesterinen@gmail.com)

Аннотация

Достижения в области машинного обучения и разработки больших графов знаний сопровождали распространение онтологий многих типов и для многих целей. Эти онтологии обычно разрабатываются независимо, и в результате может быть трудно обмениваться информацией о них и между ними. Для решения этой проблемы коммуникации онтологи должны договориться о том, как их терминология и формализация соотносятся друг с другом. Процесс достижения договорённости и согласия называется «гармонизацией». На Саммите Онтологий 2021 был рассмотрен общий ландшафт онтологий, многие виды генерации и гармонизации онтологий, а также устойчивость онтологий. В коммюнике обобщаются и резюмируются выводы саммита, а также предыдущих саммитов по смежным вопросам. Одним из основных препятствий на пути к гармонизации является относительно низкое качество определений на естественном языке во многих онтологиях. На саммите было рассмотрено состояние дел в разработке определений естественного языка, основанное на лексикографических принципах, а также примеры текущих проектов, которые связаны с гармонизацией и устойчивостью.

Ключевые слова: онтология, машинное обучение, определения, устойчивость, гармонизация.

Цитирование: Баклавски, К. Онтологический Саммит 2021. Коммюнике: генерация и гармонизация онтологий / К. Баклавски, М. Беннет, Г. Берг-Кросс, Т. Шнайдер, С. Сеппала, Р. Шарма, Р. Д. Шрирам, А. Вестеринен. Перевод с англ. Д. Боргест // Онтология проектирования. – 2021. – Т.11, №4(42). - С.533-548. – DOI: 10.18287/2223-9537-2021-11-4-533-548.

1 Введение

Онтологии быстро распространяются, создавая сложный ландшафт многих типов, ролей и использования для многих целей [Hitzler, 2021b], таких как интеграция данных, приложения семантической сети, бизнес-отчёты и искусственный интеллект (ИИ). Онтологии можно извлекать, изучать, модулировать, связывать, преобразовывать, анализировать и согласовывать, а также разрабатывать в рамках формального процесса, который может быть ручным или автоматическим.

В настоящее время существует множество способов взаимодействия онтологий с другими технологиями, включая, помимо прочего, статистические и лингвистические методы, создание инструментов машинного обучения, служащих основой для машинного обучения

¹ Перевод финальной версии Коммюнике Онтологического Саммита за 2021 год, опубликованного в ноябре 2021 года. - <https://ontologforum.s3.amazonaws.com/OntologySummit2021/Communique/Ontology+Summit+2021+Communique.pdf>.

Предварительная краткая информация о Саммите представлена во 2-ом номере 11-го тома журнала. - [https://www.ontology-of-designing.ru/article/2021_2\(40\)/Ontology_Of_Designing_2_2021_12_Ontology_Summit_2021_251-252.pdf](https://www.ontology-of-designing.ru/article/2021_2(40)/Ontology_Of_Designing_2_2021_12_Ontology_Summit_2021_251-252.pdf).

(МО) с целью повышения качества и объяснимости его результатов, и интеграция в архитектуры МО [Gaur, Faldu, and Sheth, 2021]. К сожалению, обычно онтологии строятся независимо. Хотя такие онтологии могут очень хорошо служить своим целям, очень сложно согласовать онтологии, которые были разработаны для перекрывающихся предметных областей (ПрО). Более того, онтологии, созданные с использованием автоматизированных методов, людям трудно понять. Растёт понимание необходимости гармонизации и лучшего определения терминов в онтологиях, а также передовых практик для их обеспечения.

На Саммите Онтологий 2021 года был изучен ландшафт генерации и гармонизации онтологий с использованием серии виртуальных презентаций и сессий, проведённых с февраля по май 2021 года. В этом коммюнике обобщаются и резюмируются результаты этой серии встреч. Основываясь на интересах сообщества, Саммит Онтологий 2021 был разбит на четыре трека, и в Коммюнике отражены подтемы этих треков. В треке «Онтологический ландшафт» (раздел 2) представлена общая основа для различных видов онтологии и для остальных треков саммита. В разделе 2.2 представлена одна из целей онтологий - коммуникация и показано, что могут возникнуть неправильные интерпретации, если термины не определены должным образом. Во избежание неправильного толкования в контексте данного Коммюнике термин следует понимать как имеющий определение на естественном языке, если только термин не имеет префикса пространства имён. Если важный термин естественного языка неоднозначен, то предполагаемое определение или определения будут уточнены. Например, в разделе 3.1 используется слово «аннотация» и указывается предполагаемое определение из словаря. Именно это понятие аннотации используется в Коммюнике, а не, например, *owl: AnnotationProperty*. В качестве другого примера в Коммюнике используется понятие «определение». Различные понятия определения приведены в начале раздела 3. В Коммюнике используются эти понятия определения, а не *rdfs: isDefinedBy*.

Определения естественного языка, особенно определения ПрО, являются важной частью онтологий; но с онтологической точки зрения они часто плохо написаны либо из-за недостатка опыта написания определений, нехватки ресурсов, либо из-за отсутствия акцента на правильном определении терминов в процессе разработки онтологии. Подробности о понятии семантически адекватного определения естественного языка и его роли в развитии и гармонизации онтологии представлены в разделе 3.

Недавние разработки более эффективных методов МО позволяют извлекать онтологическую информацию из исходных документов. Хотя такие методы эффективны для конкретных задач, в настоящее время они, как правило, непрозрачны и трудно поддаются пониманию и объяснению. Проблема объяснимости рассматривалась на предыдущем саммите онтологов [Waclawski et al., 2020a]. Один из способов сделать методы МО более объяснимыми - это интегрировать МО с онтологиями [Waclawski et al., 2018b; Гаур, Фалду и Шет, 2021].

Архитектуры, которые были разработаны для интеграции онтологий с МО и, в более общем плане, интеграции символьных (включая графы знаний и онтологии) и субсимвольных (таких как МО) методов, представлены в разделе 4.

Недостаточно создавать онтологии с хорошо разработанными определениями и документацией, они должны поддерживаться в течение долгого времени. Устойчивость онтологий - это гораздо больше, чем просто обеспечение постоянным финансированием. Результаты трека «Устойчивость онтологий» представлены в разделе 5.

Раздел 6 – выводы и заключение по проблемам, поднятым Саммитом Онтологий.

2 Онтологический ландшафт

Онтологии имеют множество аспектов и целей. В результате растущая классификация онтологий многомерна и образует богатый ландшафт, а не простой линейный «спектр».

Различия между типами онтологий изучались и анализировались на первых двух Саммитах Онтологий в 2006 и 2007 годах. Первым был саммит верхних онтологий [Cassidy, Obrst, Ray, Soergel, and Yim, 2006], а вторым - «Онтология, Таксономия, Фольксномия: понимание различий» [Obrst et al., 2007]. На этих двух встречах на высшем уровне был определён ряд онтологий, и некоторые из наиболее известных онтологий того времени были классифицированы [Baclawski 2007a, 2007b; Baclawski & Duggar, 2007]. Наиболее распространенная классификация онтологий - по уровню их общности, а в разделе 2.1 обсуждаются типы онтологий, различающиеся по степени их общего применения.

Одна из основных целей онтологий - коммуникация, и в разделе 2.2 рассматриваются типы онтологий с точки зрения различных коммуникационных требований. В разделе 2.3 обсуждаются различные подходы к тому, что должны представлять объекты онтологий. В разделах 2.4 и 2.5 рассматриваются отношение к реализму и представление неопределённости. Раздел 2 заканчивается кратким обсуждением результатов.

2.1 Общность

Онтологии чаще всего классифицируют по уровню их общности. Наиболее общие или абстрактные онтологии называются базовыми онтологиями, общими онтологиями и онтологиями верхнего уровня [Ontology Term List, 2020]. Наиболее специфические онтологии называются прикладными онтологиями, потому что они обычно связаны с конкретным приложением или узким диапазоном приложений. Между этими двумя крайностями находятся справочные онтологии и онтологии ПрО. Справочные онтологии более специфичны, чем базовые онтологии, но не ограничиваются определённой областью. Онтологии ПрО ограничены одной ПрО, но ПрО могут образовывать иерархии с множеством уровней общности, а онтологии ПрО также могут иметь много уровней общности [Schneider, 2021]. Были предложены другие понятия общности онтологий, такие как эталонные онтологии верхнего уровня, которые обеспечивают спецификации требований, функций, дизайна или стандартов для конкретного приложения [Chen, Ludwig, Ma, and Walther, 2019].

2.2 Коммуникация

Одна из целей онтологии - улучшить коммуникацию. Общение может происходить между людьми, между людьми и машинами, а также между машинами. Машины работают и взаимодействуют лучше всего, когда элементы онтологии определены точно и логически, так что они могут быть обработаны (например, выводы могут быть вычислены) и их неоднозначность сведена к минимуму. Человеческий язык, для сравнения, гораздо богаче большим количеством типов речи и присущими ему двусмысленностями [Hanks, Jezek, 2008; Baclawski, 2021]. Это богатство далеко не недостаток человеческого языка, а его сила. Люди справляются с двусмысленностью языка через контекст диалога, манеру общения (например, интонацию и жесты) и способность просить разъяснений [Sowa, 2021].

Онтология и её документация должны распознавать различия между потребностями людей и машин. Одним из средств достижения этого является наличие «языкового интерфейса», который является посредником между человеческой и машинной терминологией. Языковой интерфейс - важная особенность современных процессов разработки программного обеспечения, и разработка онтологий также может извлечь выгоду из этих практик [Bennett, 2021; Woods & Low, 2021].

Одной из функций определения является уточнение умозаключений читателей или систем, то есть того, что читатели или системы делают, когда сталкиваются с определённым термином [Seppala, Ruttenberg, Schreiber, and Smith, 2016]. Теоретическое объяснение функций определений естественного языка в онтологиях, подкрепленное эмпирическими нейропсихологическими исследованиями, см. в [Seppala, Ruttenberg, and Smith, 2016].

Другим эффективным средством коммуникации является построение онтологии на событиях и состояниях, представляющих интерес в ПрО, а также на аргументах (субъекты, объекты и т.д.) этих событий (согласно нео-Дэвидсоновской семантике). Основание онтологии на событиях и состояниях позволяет онтологиям быть глубоко информированным нарративом и лингвистической теорией событий [Westerinen, 2021].

Для онтологии, использующей OWL, Common Logic или другие языки представления онтологий на основе логики, идентификаторы (для элементов онтологии) фактически являются символами логического языка. Однако люди уже знакомые с такими терминами или фразами естественного языка, иногда воспринимают их эмоционально и предвзято. В результате люди могут легко забыть или проигнорировать, что идентификаторы должны рассматриваться как символы, интерпретация которых продиктована формализацией, используемой для определения идентификаторов. Неправильная интерпретация ещё более вероятна, когда элемент онтологии не определён должным образом, чтобы соответствовать предполагаемым интерпретациям. Неадекватные определения возникают по многим причинам. Одна из причин, по которой определение естественного языка может быть неадекватным, - это неявная уверенность разработчиков-онтологов в том, что пользователь элемента «поймёт», как этот элемент должен быть интерпретирован. Разработчики онтологий могут иметь определённое понимание термина или фразы на естественном языке, не осознавая, что некоторые пользователи онтологии имеют другое понимание [Schneider, 2021].

Чётко определённые концепции являются важным компонентом коммуникации и играют центральную роль в семантическом взаимодействии. Как коммуникация, так и функциональная совместимость основаны на общем понимании концепций, услуг, информации и данных. На Саммите Онтологий 2007 обсуждался спектр полезных семантических артефактов, начиная с неявного предпочтения строгих формальностей в онтологиях [Obrst et al, 2008]. Существует ряд ресурсов, таких как фолксонии (например, простые определяемые пользователем списки ключевых слов, полезные для аннотирования ресурсов в Интернете), а также таксономии, концептуальные модели и контролируемые словари (например, медицинские предметные рубрики, *MeSH*). Контролируемый словарь отражает ассоциацию терминов и является частью использования языка в ПрО, которая, подобно модели, налагает некоторую упрощающую, ограничивающую и организующую форму на изменяющуюся ПрО. Стимулом контролируемых словарей является формализация понятий на логическом языке, позволяющая в некоторой степени автоматизировать обработку данных. На Саммите Онтологий 2007 было меньше внимания уделено стандартизации других форм семантических ресурсов на этапах преформализации. В последнее время стали широко признаваться и цениться как онтологии, так и стандартные словари для поддержки коммуникации, совместного использования данных и взаимодействия.

2.3 Различия онтологий

Онтологии могут использоваться для ряда целей, таких как интеграция данных, приложения семантической паутины, бизнес-отчёты и ИИ. Различные варианты использования лучше всего поддерживаются разными, а иногда и противоположными стилями моделирования. Ключевым отличием онтологий, как и любой модели, является вопрос о том, какого рода вещи представлены элементами этой модели (например, классы, отношения, свойства и т.д.). Онтология - это модель, которая представляет вещи и отношения между ними, актуальные для определённой цели. Многие онтологии предназначены для контекстуализации данных конкретной ПрО.

Онтологии определяют классы с точки зрения необходимых и достаточных характеристик, которые делают что-то членом этого класса. Большинство этих различий не выражаются в виде данных. Например, юридический или финансовый инструмент будет определяться с точки зрения юридических или финансовых характеристик [«Social Constructs» in Searle, 2010]; недвижимое имущество, улицы и острова - это физические вещи в мире; и т.д. Между тем, приложения будут использовать данные, которые были определены как наиболее полно отражающие эти вещи в мире. Например, ваша дата рождения может быть представлена как тип «Событие» с датой (рождения), в которой произошло событие, и эта дата может быть представлена данными под названием «*dateOfBirth*» с типом данных «*date*».

Один из видов онтологии, часто называемый онтологией «концепций», будет содержать логические утверждения, сформулированные в терминах атрибутов ПрО и того, как они определяют значения вещей в этой ПрО [Bennett, 2021]. Эти логические утверждения могут использоваться людьми для понимания вещей в ПрО, например, в деловых целях. Другой вид онтологии используется для моделирования данных, связанных с предметами в ПрО.

Выбор данных для представления этих вещей является проектным решением, даже если во многих случаях выбор решения очевиден. Их часто называют «операционными» или «прикладными» онтологиями.

Дальнейший анализ показывает, что могут существовать разные стили онтологий, которые имеют дело с операционными данными. Онтология для интеграции нескольких источников данных может нуждаться в семантически более детальных различиях, чтобы иметь дело с различными способами, которыми эти источники данных отражают концептуализацию мира. В то время как онтология для обоснования данными (например, в графе знаний) была бы проще.

Другие различия отражают выбор операционной онтологии. Например, операционная онтология не обязательно должна использовать полную базовую онтологию для разделения своего мира. Как правило, у него будет меньше связей с минимальным использованием ограничений, таких как области свойств и диапазоны.

Напротив, онтология понятий будет иметь более богатый набор отношений, например, для различных типов отношений целого и частичного, и будет отражать множество ограничений, которые применяются к отношениям и свойствам.

При выводе операционной онтологии из онтологии понятий необходимо несколько этапов проектирования: выравнивание иерархии классов, извлечение понятий, относящихся к варианту использования приложения, выборочное удаление классов, которые отражают некоторые разделы онтологии верхнего уровня, такие как «вещи в ролях» и сокращение соответствующих путей собственности, так что, например, концептуальное различие между ссудой - заёмщиком (как сторона в роли) - лицом, становится простым отношением «ссуда - лицо». Каждый вариант использования приложения, будучи отдельным бизнес-контекстом, повлечёт за собой извлечение различного материала из онтологии понятий. В этом отношении проектирование операционных онтологий следует по пути, аналогичному развитию любой другой технологии, с концепцией онтологии, играющей роль «вычислительно независимой модели» в процессе разработки.

Чтобы онтология, особенно очень большая, была непротиворечивой в более широком смысле этого термина, она должна быть не только логически и концептуально последовательной, но и иметь единый стиль повсюду. Последовательность в этом более широком смысле требует соответствия руководящим принципам стиля, таким как соглашения об именах и процедуры управления изменениями [Uschold, 2021]. Последовательность также является необходимой частью устойчивости, что более подробно обсуждается в разделе 5.

2.4 Реализм

Другая сторона онтологии - степень «реализма». Философский реализм - это тезис о том, что существует реальность, существующая независимо от людей. Известно много понятий реализма, которые были предложены и изучены в философии. Онтология является реалистической по отношению к определённому понятию реализма, если она основана на тезисе философского реализма. Ни философские, ни вычислительные онтологии не требуют реализма; действительно, компьютерные учёные обычно более плюралистичны в отношении того, какие объекты могут иметь значение, являются ли они онтологическими (в философском смысле), концептуальными, когнитивными, гипотетическими и т.д. Более плюралистический подход имеет преимущества при работе с естественным языком, здравым смыслом и другими человеческими способностями. Ещё одним следствием плюралистического отношения является то, что оно допускает понятия «истины», которые лежат в основе различных механизмов рассуждения (то есть логики) и зависят от контекста и ситуации [Masolo, 2021]. Различные формы логики являются частой темой более ранних онтологических саммитов, в том числе в 2007 г. [Obrst et al, 2008], в 2010 г. «Создание онтологов будущего» [Neuhaus et al, 2010] и в 2017 г. «ИИ, обучение, рассуждение и онтологии» [Baclawski et al, 2018b].

2.5 Неопределённость

Человеческое понимание обычно вероятностное. Чтобы отразить лучшее понимание мира, онтологии для некоторых ситуаций или приложений должны поддерживать спецификацию неопределённости и рассуждения. Однако не существует общепринятого способа определения неопределённости в онтологии. Например, вероятность того, что конкретный оператор *RDF* будет истинным, может быть определён просто как добавление предиката: вероятность в качестве границы *RDF** или как повторение оператора *RDF* [Westerinen, 2021]; или можно было бы аннотировать утверждение с помощью предиката *strength* [Sharma, 2021]. Некоторые из предложений по учёту неопределённости в онтологии включают *PR-OWL* [Carvalho, Laskey, and Costa, 2017] и байесовский язык онтологий BEL [Ceylon and Penaloza, 2017]. Но, возможно, наличие гибкости в определении онтологических вероятностей является преимуществом. Возможно, понятие неопределённости, как и другие аспекты онтологии, будет зависеть от Про, сообщества и цели. Вместо того, чтобы пытаться создать единую эталонную онтологию неопределённости, должна быть мета-теория неопределённости, которая была бы модульной и настраиваемой для конкретной задачи и цели [Breiner, 2021].

2.6 Онтологические обязательства

Разработка онтологий и их устойчивость - это процессы, в ходе которых необходимо принять множество проектных решений. Эти решения различаются в зависимости от того, какая часть онтологии будет затронута. Более того, разные методологии разработки онтологий различаются в зависимости от того, какие решения принимаются, кто принимает решения, а также когда и как принимаются решения (т.е. от управления). Хотя модульность может помочь ограничить объём решения, тем не менее будут решения, имеющие серьёзные последствия для процесса и разрабатываемых артефактов. Проектные решения также влияют на то, насколько хорошо онтология может адаптироваться к будущим требованиям. Решения, принятые во время разработки онтологии,

обычно называют «онтологическими обязательствами». При использовании этого термина в контексте развития онтологии следует быть осторожным, чтобы не путать его с философским понятием онтологической приверженности, таким как критерий Куайна (Стэнфордская энциклопедия философии).

2.7 Понимание и анализ

Понимание и определение понятий, которые должны быть представлены в онтологии, являются фундаментальными проблемами в разработке онтологии, которые отличают её от разработки программного обеспечения или даже систем. В разработке онтологии есть как минимум две фазы для представления понятия или объекта. Во-первых, это понимание того, что такое понятие или сущность. Во-вторых, это этап анализа, на котором понятие или сущность представляется или моделируется с использованием выбранных или доступных конструкций. Например, кто-то может выбрать *OWL* в качестве языка онтологии или работать в контексте базовой, справочной или предметной онтологии.

На этапе «понимания» разработчику необходимо изучить соответствующие ссылки, начиная с «общего» словаря (например, Оксфордского или Кембриджского) и переходя к более конкретным ссылкам на ПрО (например, спецификациям *ISO*), если они существуют. Использование диапазона ссылок может обеспечить контраст между различными способами понимания понятия или объекта, а также контекстами, в которых возникает понятие или объект. Эти контрасты в значительной степени помогут в «понимании» понятия или объекта и того, что может потребоваться для представления удовлетворения заявленных потребностей разрабатываемой онтологии. В идеале фаза понимания должна привести к полезному определению на естественном языке, критическому элементу в представлении понятия или объекта для интерпретации человеком.

Следующий этап - процесс онтологического анализа. Этапы понимания и анализа обычно выполняются параллельно, а не последовательно. Подспорьем в онтологическом анализе может быть хорошо сконструированная базовая онтология (например, *DOLCE*, *BFO*, *UFO*, *GFO*), поскольку такая онтология уже включала онтологический анализ во время своего создания. Одна из распространенных ошибок на этапе анализа - это «соблазнение таксономии», которое представляет собой тенденцию помещать объект в таксономию без завершения понимания или анализа. Преждевременная «классификация» понятия или объекта может исказить все усилия по разработке и/или потребовать доработки [Schneider, 2021].

3 Определения и гармонизация

Есть много понятий определения. Здесь даётся краткое изложение некоторых из них. Для более полного обсуждения различных понятий определения см. Стэнфордскую энциклопедию философии. Одно из основных различий связано с тем, насколько точно и полно определение раскрывает термин.

Самая сильная форма - это *преднамеренное (intentional) определение*, которое даёт необходимые и достаточные условия для членства в расширении определённого термина. Более слабая форма - это *расширенное (экстенциональное) определение*, в котором явно перечислены все примеры. Расширенное определение может оказаться невозможным или выполнимым, и могут быть другие примеры, которые впоследствии необходимо будет добавить. Самая слабая форма - это *явное (ostensive) определение*, которое часто невербально и сопровождается жестом, указывающим на пример.

Другое различие связано с соответствием реальности. *Настоящее определение* пытается соответствовать действительности, в то время как *номинальное определение* касается только использования термина на практике. Номинальные определения различаются в зависимости от того, насколько точно они отражают существующее использование: *описательное определение* нацелено на совместимость со всеми существующими обычаями, в то время как *условное определение* вводится либо на временной основе, либо в конкретном контексте и не обязательно должно быть совместимо с каким-либо другим употреблением этого термина.

Один из аспектов определения состоит в том, что должны быть термины, такие как термин «объект» (или «сущность»), которые явно не определены. В противном случае определения неизбежно будут циклическими.

Определения могут служить связями между людьми, между человеческими сообществами, между людьми и онтологиями, а также между различными онтологиями. Исторически пытаются стандартизировать термины, включая создание основных моделей метаданных и общих концептуальных моделей для объединения данных в единое представление [Silva, Perez, and Kofuji, 2019]. Например, простая система организации знаний (*Simple Knowledge Organization System, SKOS*) является основным стандартом модели метаданных для Интернета [SKOS, 2009]. Однако эти попытки стандартизации в значительной степени не были приняты из-за ошибочной концептуализации, отсутствия согласия сообщества и неадекватного представительства; и, таким образом, привели к разобщённости.

В последнее время достигнут значительный прогресс в использовании передовых практик, включая использование онтологического анализа и проектирования. На Саммите Онтологий 2021 были рассмотрены различные понятия и уровни формальности определений с акцентом на практические методы согласования различных семантических ресурсов, и в этом разделе приводится краткое изложение этого обзора.

Словари ПрО значительно различаются по качеству и объёму, часто с альтернативными определениями одного и того же термина и определениями, имеющими разную степень формальности. Эта проблема признана давно. Когда Конфуция спросили, что он сделал бы, будь он правителем, он ответил, что «исправит имена», чтобы слова соответствовали действительности. Стандартизация значений терминов остаётся сложной задачей, поскольку существует множество противоречивых и пересекающихся глоссариев и несовместимых моделей данных, которые определяют термины ПрО своеобразными, специфичными для ПрО или приложения способами. Полное «исправление имён», как попытка разработать единую онтологию для всего, может быть слишком амбициозной целью, но гармонизация может быть достигнута, хотя и с некоторыми усилиями. Гармонизация терминологии ведётся в некоторых областях, таких как криосфера, которая представляет собой замороженную водную часть системы Земли, связанную с ледяными полями и ледниками [ESIP, 2021], а также в областях, связанных с правительством, таких как Национальная модель обмена информацией [NIEM, 2017].

Написать адекватное определение, будь то с точки зрения человека или компьютера, не просто. Одна из основных проблем при разработке определений на естественном языке для онтологий - это предположение, что каждый может писать определения и тогда они будут гармоничными. Опыт проектирования онтологий, а также эксперименты показали, что даже для хорошо изученных ПрО результаты ручной классификации, выполняемой экспертами в ПрО, очень противоречивы [Westerinen, 2021]. Итак, первый шаг в написании определений - признать, что это трудная задача. Существует множество принципов написания определений и руководств по лексикографии, терминологии и логике, которые можно использовать для написания определений [Seppala, Ruttenberg, and Smith, 2017].

3.1 Определения и гармонизация в экологических науках

В этом разделе подробно представлен конкретный пример ПрО, который направлен на определение и семантическую гармонизацию. ПрО - науки об окружающей среде, а онтология - онтология окружающей среды (*EnvO, EnvO Ontology*). *EnvO* - это семантический ресурс для семантически контролируемых описаний объектов окружающей среды. Например, глоссарий *Darwin Core* использует *EnvO* в своих описаниях среды обитания и был разработан путём применения методов интеллектуального анализа текста для извлечения информации о среде обитания из энциклопедии жизни и автоматического создания экспериментальных классов среды обитания в *EnvO*.

Пример определения на естественном языке *EnvO*: среда обитания - это «экологическая система, которая может поддерживать и обеспечивать рост экологической популяции».

Первоначально *EnvO* предназначена для представления биомов, характеристик окружающей среды и материалов окружающей среды, а первоначальной целью были исследования, связанные с геномом и микробиомом. Однако потребность в семантике среды присуща множеству полей, и объём *EnvO* неуклонно растёт с момента его первоначального описания. По мере расширения области применения онтология была расширена и обобщена для поддержки её всё более разнообразных приложений (таких как Кριο (*Cryo*, ледники и ледяные поля) и Морской (океанический) мир), которые теперь рассматриваются более подробно.

Глобальная служба криосферы (*Global Cryosphere Watch, GCW*) спонсировала усилия по согласованию определений терминов, которые уже используются в области Кριο. Иногда один термин может иметь десятки различных определений, например, термин «снежный покров». Согласованное определение *GCW* выглядит следующим образом: «Плотность площади, присущая снегу, распределённому по площади суши или другого субстрата».

Дополнительная информация об этом термине содержится в аннотациях. Аннотация - это пояснение или комментарий, добавленный к данным или метаданным. Определения - это аннотации, но есть и другие виды аннотаций, которые существенно влияют на понимание как людьми, так и машинами. Например, аннотации могут использоваться для объяснений [Baclawski et al, 2020a] и для определения контекста и происхождения информации [Baclawski et al, 2018a]. Пример аннотации для *GCW*, которая не является частью определения термина, утверждает, что, в общем, снежный покров - это слой снега на поверхности земли, и его можно сравнить с соответствующими терминами снежное поле и снежное покрытие.

Определения естественного языка *EnvO* соответствуют руководящим принципам «Минимум информации для отчётности онтологии» (*MIRO*) [Matentzoglou, Malone, Mungall et al, 2018].

Усилия по гармонизации *EnvO* привели к ряду достижений [Berg-Cross, 2021].

- Результаты анализа глоссария *GCW* были согласованы с *EnvO* и согласованы с соответствующими терминами в онтологии Семантической паутины для технологий Земли и окружающей среды (*Semantic Web for Earth and Environment Technology, SWEET*).
- Согласование со *SWEET* улучшило определения терминов *SWEET*.
- Термины *EnvO* и *SWEET* согласованы с другими терминами онтологии *OBO Foundry* [Ontology Tools and Resources, 2021].
- Специальное *envoPolar* подмножество *EnvO* было создано с соответствующими терминами и аксиомами.
- Приписывание всех обновлений было документально подтверждено людьми или группами, ответственными за эти изменения.
- Были внедрены передовые методы для документации, включая аннотирование времени добавления определения, *orcid.org* человека, создавшего обновление, и информацию о происхождении определений от *EnvO* до *SWEET*.
- Для ускорения создания классов использовались методы создания шаблонов, а для обновления онтологий с помощью определений использовались электронные таблицы. В частности, инструмент *OBO Robot* использовался для облегчения сотрудничества с *EnvO*. Инструмент *Robot* направляет пользователей в процессе создания новых терминов и предназначен для использования не онтологами. Робот организует запросы новых терминов в стандартизированном шаблоне таблицы *Google*, и пользователи могут следовать пошаговому процессу, чтобы заполнить соответствующие столбцы таблицы.

3.2 Извлечённые уроки и передовой опыт

В этом разделе представлены некоторые из лучших практик по гармонизации определений. Общие советы по разработке определений естественного языка см. в Руководстве по написанию определений в онтологиях [Seppala et al, 2017]. Рекомендации включают в себя такие советы, как: будьте краткими, согласовывайте, повторно используйте, расширяйте и пересматривайте семантические ресурсы. Следующие пункты объединяют и расширяют рекомендации, добавляя передовой опыт и советы [Berg-Cross, 2021b].

- Организуйте термины в систему понятий (сеть понятий) и укажите положение каждого термина в системе понятий. В некоторой степени неформальные концептуальные модели могут помочь организовать сети на ранних этапах разработки онтологии.
- Чтобы определения были краткими, поместите другую, более энциклопедическую информацию в примечания к аннотациям.
- Для структурирования таксономий используйте ближайшую общую концепцию высшего уровня, добавив одну или несколько ограничивающих характеристик.
- Рассмотрите возможность согласования роли лексических модификаторов, а также существительных и глаголов. Например, такие слова, как «альпийский» или «жидкий».
- Повторно используйте существующие словари, такие как *Schema.org*, *DCAT2*, *VIVO*, *DDI* и т.д. Однако это обычно требует специализации более общих терминов.
- Уменьшите (концептуальную) двусмысленность, явно показывая отношения между терминами. Неоднозначность уменьшается по семантическому спектру несколькими способами. Для людей неоднозначность термина уменьшается за счёт хорошего определения или включения в глоссарий связанных терминов. Дополнительная структура добавляется путём определения принадлежности к классу отношений подтипов. Использование логического языка для представления определений может уменьшить двусмысленность автоматизированных цифровых систем. Некоторые термины, описанные как «похожие» друг на друга, несут некоторую двусмысленность в том, насколько они похожи. Это можно уменьшить, указав более точные отношения. Например, вместо описания ледника как подобного ледяной массе можно указать, что «ледник - это тип ледяной массы».
- Будьте внимательны к вопросам детализации. Метонимия - это наименование вещи чем-то связанным с ней. Например, использование имени целого для части или наоборот. Такие речевые образы, как метонимия, настолько обычны, что можно вообще не осознавать, что они используют речевые образы.
- Избегайте расплывчатых сравнений, например, использования слова «похожий» в определении. Обычно существует много похожих терминов и терминов с перекрывающимися значениями. Выполните семантический анализ, чтобы различать похожие термины. Семантический анализ включает в себя установление отношений между терминами там, где это уместно, таких как подтипы, отношения части и целого, роли, влияния, отношения производства (выхода) и т.д.
- Интерфейс между деловым (естественным) языком и техническим языком (например, логикой описания или логикой первого порядка) помогает человеко-машинному общению, делая язык более точным и менее двусмысленным [Seppala, 2021; Woods & Low, 2021]. Влияние различий между текстовыми и логическими определениями в онтологиях исследовалось в [Seppala et al, 2016]. Однако, общаясь с естественным языком, нужно знать, как люди классифицируют мир. В отличие от классов и свойств онтологий, человеческие категории «мерцают» [Hanks, Jezek, 2008; Waclawski, 2021].
- Термины могут быть стандартизированы, но значения или их ожидаемые толкования должны быть такими же.

- Определения будут варьироваться в зависимости от контекста использования и целевой аудитории.

Эти уроки сообщества *EnvO* могут быть полезны другим сообществам.

4 Нейросимвольные онтологии обучения

Символьное мышление имеет долгую историю и продолжает оставаться активной областью исследований. МО, известное как субсимвольные методы, также является очень активной областью исследований. Хотя оба метода являются частью ИИ, эти две области были разработаны на разных технических основах и отдельными исследовательскими сообществами. Эти две области имеют взаимодополняющие сильные и слабые стороны. В результате поиск способов преодоления разрыва между символьным и субсимвольным подходами к ИИ является давней нерешённой проблемой, и интеграция этих двух областей в настоящее время является предметом растущего исследовательского интереса к ИИ. Устранение этого разрыва было рассмотрено на Саммите Онтологий 2017 [Baclawski et al, 2018b], но с тех пор были разработаны новые методы ИИ, особенно в МО, поэтому пересмотр этой темы, безусловно, своевременен.

Нейросимвольное обучение направлено на интеграцию нейронного обучения с символьными подходами, обычно используемыми в вычислительной логике и представлении знаний в ИИ. Одним из преимуществ такой интеграции является разработка эффективных методов извлечения знаний для объяснимого ИИ [Gaur, Faldu, and Sheth, 2021; Lamb, 2021; Sheth, 2021], но есть много других преимуществ [Sriram, 2021]. Хотя есть значительные преимущества более тесной интеграции нейронной и символьной парадигм, неизвестно, как лучше всего их интегрировать, и было предложено множество архитектур интеграции. Символьные модели могут быть результатом или основой различных стадий нейронного процесса. Некоторые архитектурные приёмы, объединяющие символьные и субсимвольные методы, приведены в [Kautz, 2021]:

1. Самая простая и наиболее распространённая архитектура - это архитектура, в которой символьные данные (например, документы) обрабатываются с помощью символьных методов для создания векторов, которые вводятся в подсимвольный модуль (например, нейронную сеть). Затем векторный вывод подсимвольного модуля интерпретируется в символьной форме с использованием символьных методов.

2. Другая архитектура - это символьная система, которая может вызывать субсимвольные подмодули. Что касается символьной системы, то подмодули - это просто подпрограммы, как и любые другие. В самоходном транспорте обычно используют эту архитектуру.

3. В принципе, можно было бы поменять местами символьное и субсимвольное в вышеупомянутой архитектуре (2), чтобы получить архитектуру, в которой именно субсимвольная система вызывает символьную. Преимущество инверсии ролей символьного и субсимвольного в том, что это позволяет принимать очень сложные решения, поскольку символьные рассуждения могут выполнять комбинаторные рассуждения гораздо более масштабно и эффективно, чем субсимвольные системы.

4. Архитектуры (1), (2) и (3) не включают обучение во время нормальной обработки. Символьный и подсимвольный модули уже запрограммированы и обучены соответственно. Некоторые недавно разработанные архитектуры включают символьное мышление в подсимвольную систему, организуя подсимвольную систему в соответствии с символьными правилами. Примерами таких архитектур являются представления тензорных произведений и логические тензорные сети, которые могут находить обобщения и иерархии части-целого. Другими словами, эти недавно разработанные архитектуры могут генерировать онтологии или,

по крайней мере, некоторые аспекты онтологий. Формы рассуждений включают временную логику, логику описания и логику предикатов первого порядка [Hitzler, 2021].

5. Другая архитектура обучения - ввод экземпляров логических выводов, выраженных в виде пар ввода-вывода, для обучения субсимвольной модели. Тренировка с примерами логического вывода в первую очередь полезна для математических задач, и она на удивление эффективна, хотя иногда допускает ошибки [Karanipathi, 2021].

6. Другая архитектура использует субсимвольные методы, такие как обратное распространение, для обучения символьной системы. Обратное распространение вызывается всякий раз, когда система делает ошибку. Тренировка символьной системы может быть полезна для вопрос-ответных систем.

Ещё одна причина, по которой субсимвольные методы нуждаются в методах, основанных на знаниях, - это интуиция, основанная на человеческом поведении, что интеллект обязательно включает обучение, знания на основе опыта и рассуждения, которые можно выразить следующим уравнением:

Интеллект = обучение на основе данных + знания / опыт + рассуждения,

которое было использовано в качестве основы для обучения, основанного на знаниях (*Knowledge-Infused Learning, KIL*) [Sheth, Gaur, Kursuncu, and Wickramarachchi, 2019]. Понятие KIL - это набор архитектур, которые варьируются от поверхностных до глубоких, основанных на знаниях, позволяющих согласовать архитектуру с конкретным приложением. *KIL* может оказать влияние на робототехнику, когнитивную науку, автономные транспортные средства и личных помощников [Sheth, 2021].

Ещё одна архитектура должна начинаться с «начальной» онтологии и постепенно увеличивать её, используя методы обработки сигналов из графов знаний, извлечённых из различных источников, таких как документы.

Обработка сигналов - это очень высокоразвитая и активная область с большим сообществом, и эта связь между этим сообществом и сообществом онтологов может иметь много преимуществ [Majumdar, 2021; Sowa, 2021; Bacławski et al, 2020b].

5 Устойчивость онтологий

Многие организации, включая правительственные агентства, органы по стандартизации и коммерческие фирмы, используют онтологии и разработали инструменты для различных онтологических операций, таких как создание, развитие, отображение и другие формы гармонизации.

Устойчивость предполагает решение гораздо большего, чем просто обеспечение достаточного финансирования. Для достижения устойчивости необходимо строить прочный фундамент. Наиболее важными аспектами такого фундамента являются следующие три «столпа» [Dickerson, 2021].

1. *Экономическая жизнеспособность*. Необходимо обеспечить достаточное финансирование для поддержки онтологии до тех пор, пока её цель остаётся актуальной. Способ распределения и мониторинга ресурсов определяет, будет ли достигнута экономическая жизнеспособность. Без надлежащего надзора невозможно сохранить экономическую жизнеспособность.

2. *Социальная справедливость*. Онтологии могут иметь предубеждения, в том числе онтологии, созданные с помощью МО. Данные неизменно имеют в большей или меньшей степени предвзятость, и методы МО не могут найти или исправить их самостоятельно [Suresh and Guttag, 2019]. Хотя сочетание символьных знаний с МО может помочь обнаружить и

смягчить её, важно признать, что устранение предвзятости остаётся жизненно важной частью разработки и сопровождения онтологий.

Использование стандартов и строгой методологии также может помочь в обеспечении надлежащего решения вопросов социальной справедливости. Стандарты обеспечивают основу для вовлечённых сторон, поэтому более широкая группа людей может в полной мере участвовать в разработке. Тщательный подход помогает установить качество, что также способствует беспристрастному подходу [Dickerson, 2021].

3. *Охрана окружающей среды.* Термин «среда» в этом столпе относится к среде человека, которая окружает конкретное сообщество, разработавшее онтологию. Важно понимать, что сообщества и их онтологии не существуют изолированно. Необходимо поддерживать возможности общения и сотрудничества с соседними и другими родственными сообществами.

Хорошо продуманные определения, документация и согласование, обсуждаемые в разделах 3 и 4, могут способствовать как социальному равенству, так и защите окружающей среды.

Столпы устойчивости показывают, что разработка онтологии - это гораздо больше, чем просто создание онтологии. Первый этап разработки онтологии - это только начало; нужно запустить её в производство.

Во-первых, для выполнения релизации необходимо планирование. Необходимо планировать инфраструктурные проблемы, участие заинтересованных сторон и планирование связанных задач [Franch and Ruhe, 2016]. После того, как онтология выпущена и активно используется теми, кто не входит в группу разработки, возникает необходимость в пересмотрах. Команда разработчиков может предвидеть не так много потенциальных проблем; реализация также содержит базовый уровень содержания [Kotis, Vouros, and Spiliotopoulos, 2020]. Этими проблемами необходимо управлять надлежащим образом, иначе жизнеспособность онтологии после выпуска окажется под угрозой.

Должны быть реализованы механизмы, облегчающие пересмотр и, при необходимости, расширение исходной модели. Хорошо спроектированные каналы обратной связи и редактирования, включая шаблоны, способствуют развитию активной среды для совершенствования онтологии [Blasko, Kremen, and Kouba, 2015]. Заинтересованные стороны могут быть ресурсами для редактирования содержания онтологии, а также для обеспечения справедливости путём продвижения стандартов.

Поддержка технической инфраструктуры способствует развитию интеллектуальной и совместной инфраструктуры, необходимой для поддержки онтологий в долгосрочной перспективе. Постоянно меняющиеся форматы, языки, платформы и инструменты также затрудняют поддержку репозитория онтологий, что было темой Саммита Онтологий 2008 [Obrst et al., 2008].

Пример *EnvO* в разделах 3.1 и 3.2 показывает, как сообщество решает проблему устойчивости своей онтологии. Результаты анализа глоссария *GCW* были согласованы с *EnvO* и с соответствующими терминами в онтологии *SWEET*. *SWEET* - это лёгкая онтология с широким охватом, но с отдельными определениями, которые исторически служили отправной точкой для концепций в науках о Земле. Для определённых *PrO* часто добавлялась более богатая семантика и создавались побочные продукты от *SWEET*. По сравнению с концепциями *EnvO*, концепции *SWEET* менее аксиоматизированы и меньше терминов имеют определения. Многие устаревшие термины, используемые в *SWEET*, были взяты из онлайн-источников, таких как Википедия, и не подвергались анализу экспертами в *PrO*. Недавно части *SWEET* были обновлены новым выпуском в 2021 году. Уроки, извлечённые сообществом *EnvO*, ценны для других сообществ.

По мере того, как методология и область исследования расширялись с развитием технологий, управление онтологиями и другими связанными семантическими ресурсами стало критически важным компонентом жизненного цикла (ЖЦ) онтологии. Масштаб и разнообразие новых семантических ресурсов, таких как графы знаний, нейросимвольные генерируемые продукты и словари ПрО, требуют пересмотра практик онтологической инженерии и различных ролей онтологий в общем семантическом исследовательском предприятии.

6 Выводы и заключение

Распространение онтологий самых разных типов, целей и ролей создало острую потребность в гармонизации онтологий для улучшения коммуникации между людьми, между людьми и машинами, а также между машинами. В этом коммюнике рассмотрены вопросы генерации и гармонизации онтологий. Поскольку существует множество заинтересованных сторон, которые оказывают влияние на эти проблемы, ниже следующие выводы сформулированы для различных заинтересованных сторон.

Заинтересованные стороны на самом высоком уровне - это сообщества и организации, которые спонсируют проекты разработки онтологий либо самостоятельно, либо как часть других проектов. На этом уровне важно убедиться, что все три столпа устойчивости хорошо обоснованы. Для расширений и исправлений онтологий необходимо согласие сообщества. Важны хорошие механизмы для обсуждения в сообществе, а также партнерские соглашения с группами, владеющими словарями ПрО. Для координации и согласования используются различные инструменты, такие как *Slack* и *Github*, но члены сообщества не обязательно обладают навыками использования этих инструментов.

Руководители проекта, включающего разработку онтологий, являются важными заинтересованными сторонами для создания и гармонизации онтологий. Усилия по контролю значения терминов в словарях требуют собственного ЖЦ, которым нужно управлять, как и другими ЖЦ цифровых данных. Руководители проектов несут ответственность за выбор и соблюдение соответствующих руководящих принципов стиля в целом и руководящих принципов стиля для определений в частности, таких как Руководство по написанию определений в онтологиях и рекомендации *MIRO*. Другая задача управления - поддержание доступа к словарю для повторного использования или согласования с онтологиями.

Руководители проектов и разработчики должны сотрудничать, чтобы принимать важные высокоуровневые решения в процессе разработки онтологии. Хорошо построенная базовая онтология может помочь в онтологическом анализе, но также может повлиять на другие решения и процесс разработки. Итак, выбор базовой онтологии должен выполняться очень тщательно. Ещё одно важное решение - нейросимвольная архитектура, которая будет использоваться для включения символьных рассуждений в субсимвольную систему. Хорошая интуиция относительно решения нейросимвольной архитектуры - это уравнение:

Интеллект = Обучение на основе данных + Знание / Опыт + Рассуждение.

Термины в словаре ПрО должны включать не только существительные и глаголы, но также лексические модификаторы, такие как прилагательные и наречия. Все термины словаря должны быть организованы в систему понятий, которая не обязательно должна быть иерархией.

Конечные пользователи онтологий и систем на основе онтологий являются заинтересованными сторонами, и разработчики онтологий должны сотрудничать с конечными пользователями, чтобы гарантировать соответствие технического языка разработчика естественному языку конечного пользователя. Использование естественного языка для символов онтологии может привести к путанице из-за существующих значений, которые люди имеют для

символов. Соответственно, онтологии должны распознавать различия между потребностями людей и машин. Одна из функций определения заключается в корректировке выводимых компетенций читателей или систем, то есть того, что они делают, когда сталкиваются с определённым термином. Одним из эффективных средств коммуникации является использование теории нарратива и событийной лингвистики. Логика, лежащая в основе онтологии, должна выбираться в соответствии с требованиями пользователей онтологии. Особенно важно включить соответствующее понятие неопределённости.

Есть заинтересованные стороны другого типа, которые лишь косвенно участвуют в разработке онтологии, но оказывают значительное влияние; а именно - исследователь онтологий. Гармонизация онтологий может занять очень много времени, поэтому инструменты важны для упрощения требуемых усилий, а также для управления усилиями с течением времени. Задача состоит в том, чтобы разработать более совершенные инструменты для гармонизации. Масштаб и разнообразие новых семантических ресурсов, таких как графы знаний, нейросимвольные генерируемые продукты и словари ПрО, требуют пересмотра практик онтологической инженерии. Изучение и обучение этим новым семантическим ресурсам, а также их использование в практике оперативной онтологии, очень важно. Были созданы некоторые предварительные механизмы для поддержания доступа к словарям для повторного использования или для согласования с онтологиями, но необходимы более совершенные механизмы. Ещё одна задача - разработать более совершенные инструменты и методы для эффективного достижения соглашения с разнообразным сообществом.

Разработчики онтологий несут важные обязанности по обеспечению правильного понимания и документирования ПрО онтологии. Одним из важных предварительных условий для понимания ПрО является изучение всех существующих релевантных ссылок. Собрав соответствующую терминологию, разработчик должен избегать соблазна таксономии, т.е. преждевременного помещения сущностей в таксономию. Важно помнить, что термины могут быть стандартизированы, но значение должно быть то же. Написание хороших определений необходимо для стандартизации смысла, но первый шаг в написании определений - это признать, что человек не знает, как это делать. Следующий шаг - научиться тому, как это делать. Теперь есть отличные рекомендации по написанию определений. Некоторые другие проблемы при разработке определений включают избежание нечётких сравнений и чувствительность к уровням детализации.

7 Благодарности

В документе описаны некоторые коммерческие программные системы. Такая идентификация не подразумевает рекомендации или одобрения Национальным институтом стандартов и технологий (*NIST*) или организациями авторов или сторонников этого коммюнике; это также не означает, что идентифицированные продукты обязательно являются лучшими из имеющихся для этой цели. Любые мнения, выводы или рекомендации, выраженные в этом материале, принадлежат авторам и не обязательно отражают точку зрения *NIST* или любых других поддерживающих правительства США или Европы или корпоративных организаций.

Доктор Селья Сеппала является научным сотрудником программы Марии Склодовской-Кюри *Career-FIT* под номером MF20180003. Она благодарна за финансирование, полученное в рамках грантового соглашения Европейского Союза № 713654 имени Марии Склодовской-Кюри.

Мы хотим поблагодарить сообщество онтологов за поддержку, особенно приглашённых докладчиков и участников, которые внесли свой вклад в саммит онтологий. Полный список сессий, докладчиков и ссылки на слайды презентаций и видеозаписи доступны по адресу <https://bit.ly/3gKukvC>.

References

- Baclawski, K. (2007a). Ontology Summit 2007 Survey Response Analysis. (Retrieved 9 June 2021 from <http://bit.ly/2DmZn0i>)
- Baclawski, K. (2007b). Ontology Summit 2007 Survey Response - Issues. (Retrieved 9 June 2021 from <http://bit.ly/2DoEa6i>)

- Baclawski, K. (2021). Shimmering Semantics: Some Insights for Ontologies from Lexicography. (Retrieved 1 June 2021 from <https://bit.ly/32WdO51>)
- Baclawski, K., Bennett, M., Berg-Cross, G., Casanave, C., Fritzsche, D., Ring, J., Schneider, T., Sharma, R., Singer, J., Sowa, J., Sriram, R.D., Westerinen, A., Whitten (2018a). Ontology Summit 2018 Communique: Contexts in Context. Applied Ontology, 13 (3): IOS Press, 181-200. (DOI: 10.3233/AO-180200 Retrieved 25 May 2020 <https://bit.ly/2LYPsz1>)
- Baclawski, K., Bennett, M., Berg-Cross, G., Fritzsche, D., Schneider, T., Sharma, R., Sriram, R.D., Westerinen, A. (2018b). Ontology Summit 2017 Communique: AI, Learning, Reasoning and Ontologies. Applied Ontology, 13 (1), 3–18. (DOI: 10.3233/AO-170191 Retrieved 9 September 2021 from <https://bit.ly/2PWRTrO>)
- Baclawski, K., Bennett, M., Berg-Cross, G., Fritzsche, D., Sharma, R., Singer, J., Singer, J., Sowa, J., Sriram, R.D., Underwood, M., Whitten, D. (2020a). Ontology Summit 2019 Communique: Explanation. Applied Ontology. (DOI: 10.3233/AO-200226 Retrieved 9 September 2021 from <http://bit.ly/2TDDyTG>)
- Baclawski, K., Bennett, M., Berg-Cross, G., Fritzsche, D., Sharma, R., Singer, J., Sowa, J., Sriram, R.D., Whitten, D. (2020b, April). Ontology Summit 2020 Communique: Knowledge Graphs. Applied Ontology, 16 (2), 229–247.
- Baclawski, K., & Duggar, V. (2007). Ontology Summit 2007 Assessment Report. (Retrieved 9 June 2021 from <http://bit.ly/2KTTD2J>)
- Bennett, M. (2021). The Landscape of Ontology Purpose. (Retrieved 1 June 2021 from <https://bit.ly/3r1s4K4>)
- Berg-Cross, G. (2021a). Introduction to Harmonizing Definitions and the EnvO Ontology. (Retrieved 1 June 2021 from <https://bit.ly/3pnkwdG>)
- Berg-Cross, G. (2021b). Semantic harmonizations of concepts revisited: work in the context of the ontology development landscape. (Retrieved 1 June 2021 from <https://bit.ly/3vsfJuY>)
- Blasko, M., Kremen, P., & Kouba, Z. (2015). Ontology evolution using ontology templates. Open Journal of Semantic Web (OJSW), 2 (1), 16–29.
- Breiner, S. (2021). Ontology and the Bayesian Brain. (Retrieved 1 June 2021 from <https://bit.ly/3o0UOfV>)
- Carvalho, R., Laskey, K., & Costa, P. (2017). PR-OWL – a language for defining probabilistic ontologies. International Journal of Approximate Reasoning, 91, 56–79.
- Cassidy, P., Obrst, L., Ray, S., Soergel, D., & Yim, P. (2006). Ontology Summit 2006: Upper Ontology Summit. (Retrieved 9 September 2021 from <http://bit.ly/2fQ6XWh>)
- Ceylan, I., & Penaloza, R. (2017). The Bayesian Ontology Language BEL. Journal of Automated Reasoning, 58, 67–95.
- Chen, J., Ludwig, M., Ma, Y., & Walther, D. (2019). Computing minimal projection modules of description logic terminologies. In Proceedings of Logics in Artificial Intelligence - 16th European Conference, JELIA (Vol. 11468, pp. 355–370). Rende, Italy: Springer. (https://doi.org/10.1007/978-3-030-19570-0_23)
- Dickerson, L. (2021). Integrating Sustainability into Ontology Development: The Case of GAO’s Fraud Ontology. (Retrieved 1 June 2021 from <https://bit.ly/3x3XrBX>)
- EnvO GitHub Site. (n.d.). (Retrieved 1 June 2021 from <https://bit.ly/2TqP6sq>)
- ESIP 2021 Advances in Semantic Harmonization: from the Cryosphere, to the Earth System. (2021). (Retrieved 1 June 2021 from <https://bit.ly/3wQbHgx>)
- Franch, X., & Ruhe, G. (2016). Software release planning. In 2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C) (pp. 894–895).
- Gaur, M., Faldu, K., & Sheth, A. (2021). Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable? IEEE Internet Computing, 25 (1).
- Hanks, P., & Jezek, E. (2008). Shimmering lexical sets. In Euralex XIII Proceedings.
- Hitzler, P. (2021a). Neural-Symbolic Integration and Ontologies. (Retrieved 1 June 2021 from <https://bit.ly/2ZPKRk>)
- Hitzler, P. (2021b). A review of the semantic web field. Communications of the ACM, 64 (2), 76–83.
- Kapanipathi, P. (2021). Getting AI to Reason: Using NeuroSymbolic AI for Knowledge-based Question Answering. (Retrieved 1 June 2021 from <https://bit.ly/3f2PDba>)
- Kautz, H. (2021). Toward a Taxonomy of Neuro-Symbolic Systems.
- Kotis, K., Vouros, G., & Spiliotopoulos, D. (2020). Ontology engineering methodologies for the evolution of living and reused ontologies: status, trends, findings and recommendations. The Knowledge Engineering Review, 35.
- Lamb, L. (2021). Neural-symbolic AI: From Turing to Deep Learning. (Retrieved 1 June 2021 from <https://bit.ly/2OGesXT>)
- Majumdar, A. (2021). Ontology from Graph Signal Processing. (Retrieved 1 June 2021 from <https://bit.ly/3tS4T0w>)
- Masolo, C. (2021). Experiences in developing with DOLCE. (Retrieved 1 June 2021 from <https://bit.ly/2OGusCt>)

- Matentzoglou, N., Malone, J., Mungall, C., et al. (2018). MIRO: guidelines for minimum information for the reporting of an ontology. *J Biomed Semant*, 9 (6).
- Neuhaus, F., Smith, B., Florescu, E., Galton, A., Gruninger, M., Guarino, N., Obrst, L., Sanchez, A., Vizedom, A., Yim, P. (2010). *Ontology Summit 2010 Communiqué: Creating the Ontologists of the Future*. (Retrieved 9 June 2021 from <http://bit.ly/2fOGuIO>)
- NIEM website. (2017). National Information Exchange Model of the Office of the Director of National Intelligence. (Retrieved 2 June 2018 from <http://bit.ly/2Lcm1re>)
- Obrst, L., Cassidy, P., Ray, S., Smith, B., Soergel, D., West, M., & Yim, P. (2006). The 2006 Upper Ontology Summit Joint Communiqué. *J. Appl. Formal Ontology*, 1 (2), 203–211.
- Obrst, L., Musen, M., Smith, B., Neuhaus, F., Olken, F., Gruninger, M., Raymond, M., Hayes, P., Sharma, R. (2008). *Ontology Summit 2008 Communiqué: Towards an Open Ontology Repository*. *Ontolog Forum*. (Retrieved 9 June 2021 from <http://bit.ly/2fOGs3E>)
- Ontological commitment. (2014). *Stanford Encyclopedia of Philosophy*. (Retrieved 15 June 2021 from <https://stanford.io/2SssY0V>)
- Ontology Term List. (2020). (Retrieved 30 June 2021 from <https://bit.ly/3AoAPxE>)
- Ontology Tools and Resources. (2021). (Retrieved 1 June 2021 from <https://bit.ly/2UdoSKC>)
- Schneider, T. (2021). *Ontology Landscape: One View*. (Retrieved 1 June 2021 from <https://bit.ly/3cYRKxw>)
- Searle, J. (2010). *Making the social world: The structure of human civilization*. Oxford: Oxford University Press.
- Seppala, S. (2021). *Systematizing definitions in ontologies*. (Retrieved 1 June 2021 from <https://bit.ly/3eUnHrQ>)
- Seppala, S., Ruttenberg, A., Schreiber, Y., & Smith, B. (2016). *Definitions in Ontologies*. *Cahiers de lexicologie*, 2 (109), 173–206. (Retrieved 1 June 2021 from <https://bit.ly/3g4QWXN>)
- Seppala, S., Ruttenberg, A., & Smith, B. (2016, July). *The Functions of Definitions in Ontologies*. In R. Ferrario & W. Kuhn (Eds.), *Formal Ontology in Information Systems: Proceedings of the 9th International Conference (FOIS 2016)* (Vol. 283, pp. 37–50). Annecy, France: IOS Press (Frontiers in Artificial Intelligence and Applications). (Retrieved 1 June 2021 from <https://bit.ly/3yTPsYZ>)
- Seppala, S., Ruttenberg, A., & Smith, B. (2017). *Guidelines for Writing Definitions in Ontologies*. *Ciencia da Informacao (Information Science Journal)*, 46 (1), 73–88. (Retrieved 1 June 2021 from <https://bit.ly/3c9G5dL>)
- Sharma, R. (2021). *Ontology Generation and Harmonization*. (Retrieved 15 June 2021 from <https://bit.ly/2YX3qsh>)
- Sheth, A. (2021). *Semantics of the Black-Box: Can knowledge graphs help make deep learning systems more interpretable and explainable?* (Retrieved 1 June 2021 from <https://bit.ly/3wG36Oc>)
- Sheth, A., Gaur, M., Kursuncu, U., & Wickramarachchi, R. (2019). *Shades of knowledge-infused learning for enhancing deep learning*. *IEEE Internet Computing*, 23 (6), 54–63. (doi: 10.1109/MIC.2019.2960071)
- Silva, A., Perez-Alcazar, J., & Kofuji, S. (2019). *Interoperability in semantic web of things: Design issues and solutions*. *International Journal of Communication Systems*, 32 (6).
- SKOS (2009). *SKOS Simple Knowledge Organization System Reference*. World Wide Web Consortium. (Retrieved 8 September 2021 from <http://www.w3.org/TR/skos-reference>)
- Sowa, J. (2021). *Automated and semi-automated methods for developing and using ontologies*. (Retrieved 1 June 2021 from <https://bit.ly/2P3ciLB>)
- Sriram, R. (2021). *Neuro-Symbolic Learning Technologies*. (Retrieved 1 June 2021 from <https://bit.ly/3kh3Aog>)
- Suresh, H., & Guttag, J. (2019). *A framework for understanding unintended consequences of machine learning*. (arXiv preprint arXiv:1901.10002)
- Uschold, M. (2021). *Commercial Ontology Development*. (Retrieved 1 June 2021 from <https://bit.ly/3dHXfAJ>)
- Westerinen, A. (2021). *Experiences with Foundational, Domain and Application Ontologies in OWL2*. (Retrieved 1 June 2021 from <https://bit.ly/3bSRbEu>)
- Woods, C., & Low, E. (2021). *Representing ontologies in Natural Language and First-Order Logic: Why do we need both?* (Retrieved 1 June 2021 from <https://bit.ly/3vKPfG2>)