

Manuscript Title

This manuscript ([permalink](#)) was automatically generated from [lubianat/fapesp_report_1@b0a999f](#) on June 13, 2021.

Authors

- **John Doe**

 [XXXX-XXXX-XXXX-XXXX](#) ·  [johndoe](#) ·  [johndoe](#)

Department of Something, University of Whatever · Funded by Grant XXXXXXXX

- **Jane Roe**

 [XXXX-XXXX-XXXX-XXXX](#) ·  [janeroe](#)

Department of Something, University of Whatever; Department of Whatever, University of Something

UNIVERSIDADE DE SÃO PAULO FACULDADE DE CIÊNCIAS FARMACÊUTICAS Departamento de Análises Clínicas e Toxicológicas

Study of transmission of diseases using cell phones data from infected people

Direct Doctorate Project - Report 1 FAPESP process number: 2019/16419-7 Project term: 01/09/2019 to 31/08/2023 Period of Scientific report: 01/09/2019 to 30/11/2020

Scholarship Holder: Jeevan Giddaluru Responsible Researcher: Prof. Dr. Helder Takashi Imoto Nakaya

São Paulo - SP November 2020

Project summary

The advent of single-cell technologies has ignited the desire of a deep knowledge on cells, the building blocks of life [1]. The Human Cell Atlas (HCA) project, has been a major player in the cell knowledge ecosystem, running since 2017 towards the task to characterize every cell type in the human body [2]. The HCA consortium recruited people from all over the world to tackle different parts of the project. In Brazil, Prof. Helder Nakaya (supervisor of this PhD project) is leading the national effort to contribute to HCA, with a focus on the roles of different cell types in the pathological processes of infectious and inflammatory diseases.

HCA is set to revolutionize the biomedical sciences, by creating tools and standards for basic research, as well as allowing better characterization of disease, and thus, ultimately, improving diagnostics and therapy. Its products (data, information, knowledge and wisdom) need to be FAIR: findable, accessible, interoperable and reusable. Data stewardship and data management are growing as core demands of the scientific community, ranging from data management plans [3] to specialized personnel [3].

The Human Cell Atlas has a dedicated team for organizing data: the Data Coordination Platform (DCP) [4] [5]. The DCP is responsible for tracing the plan for computational interoperability, from the data generators to the consumers.[5]. The Human Cell Atlas has its portal for data (<https://data.humancellatlas.org/>) which composes the data repository landscape with other resources, like the Broad Institute Single Cell Portal (https://singlecell.broadinstitute.org/single_cell) and the Chan-Zuckerberg Biohub Tabula Sapiens (<https://tabula-sapiens-portal.ds.czbiohub.org/>). In addition to its core team, the HCA is poised to grow by community interaction, and states in its opening paper that “As with the Human Genome Project, a robust plan will best emerge from wide-

ranging scientific discussions and careful planning.”[2]

Thus, this project inserts itself among the wide-ranging scientific discussions to improve data - and knowledge - interoperability.

The large amount of information generated by HCA creates the need for innovative knowledge management approaches. For the Human Cell Atlas Project (as presented in the chapter) to maximize its benefit for society, its knowledge products will need to be inserted into the main route of automated knowledge discovery .

The field denominated Literature Based Discovery [6] dedicates itself to this challenge: making actual discoveries (or at least very strong hypothesis) using as material plainly the existing literature. [7] The textbook example of Literature Based Discovery is described by Don Swanson’s so-called ABC model: If A is related to B, and B is related to C, then A and C are indirectly related [8] In a seminal paper, Swanson showed an hypothesis about using fish oil (A) to treat Raynaud’s disease (C), demonstrating that even though the specialized fish-oil (A) literature had shown its association (AB) with a set of blood parameters (B), and the specialized Raynaud’s disease literature had show its association (BC) with the same set of parameters (B), the AC link was never made in the literature, despite its seeming obviousness [8].

Modern advancements of literature-based discovery rely on Natural Language Processing, Machine Learning and Knowledge graphs to make inferences on literature knowledge. Word embeddings, for example, are leading inference of properties of compounds based on their shared neighbourhood of words (the words before and after their mentions) with known compounds, thus making use of latent knowledge in the body of knowledge. [9] Other, more explicit approaches, rely on extracted relations embedded in knowledge graphs, for example, the discovery of new RNA-binding proteins related to Amyotrophic Lateral Sclerosis by analysis of the Watson Drug Discovery gene-disease network. [10] Knowledge graphs have a set of characteristics that make them useful for Literature Based Discovery: the power of representing multiple relations, the power of making inferences on top of those relations, and provide human understandability at every step, allowing for a dialog between expert humans and computing systems. The field of biomedical ontologies explores that direction in depth, and the community is building many solutions, widely applicable for the biomedical sciences.

An ontology, as used here, is a formal computational representation of reality, which tries to represent each concept (and their relations) as precisely as possible. [11] Constructing an ontology is a process of selecting and defining terms of interest, selecting and defining relationships of interest and making statements about reality using terms and relationships. The Gene Ontology is probably the most well known biomedical ontology; it describes (among other things) different classes of biological process, related to each other by “is_a” and “part_of” relations. [12] [13].

The Gene Ontology is part of a much larger effort to formalize concepts across biology: the Open Biomedical and Biological Ontologies (OBO) Foundry. [14] Created in 2007, the OBO Foundry is a hub of biomedical ontologies that sets guidelines for the design and construction of high-quality ontologies. The initial OBO Foundry united several independent ontologies, like the Cell Ontology (CL), the Disease Ontology (DO) and the Protein Ontology (PRO) under a common framework, a great progress towards interoperability. At the same time, the creation of the Relation Ontology (RO) provided a go-to point for relations in biology that could then be reused by different ontologies.

Ontologies are powerful, but require a high degree of technical expertise to get started. Recently, a new approach for formal knowledge representation arose with the dawn of collaborative knowledge graphs. Wikidata, the collaborative knowledge graph of the Wikimedia foundation, allows users to contribute with classes and statements, in the same spirit of Wikipedia and share its “epistemic virtues, like power, speed and availability. [15] It is powerful because of its large community of contributors. With a community of more than 25,000 active editors (<https://www.wikidata.org/wiki/Wikidata:Statistics>) and growing, it is able to cover a much wider

number of concepts than any user individually. It is fast, because one does not need to install any software or ask for permissions to update it: any user can simply do it via a web interface. That speed makes it easier for newcomers to join and contribute, in contrast to OBO Foundry ontologies, which require extensive training on semantics and knowledge of Git/GitHub for contributions. Finally, the information on Wikidata is available via an user interface, via a SPARQL query service and as large, full-size database dumps, providing full extent reusability. The Wikidata model has been so successful that Google decided to migrate its own knowledge base, Freebase, fully into Wikidata.[\[16\]](#)

Several advances towards biological data integration and biological data analysis in Wikidata have been made before, yielding positive results [\[17\]](#) [\[18\]](#) and showcasing its potential for bioinformatics-related analyses, such as drug repurposing and ID conversion [\[19\]](#). Wikidata has been proposed as a unified base to gather and distribute biomedical knowledge, with more than 50 000 human gene items indexed and hundreds of biomedical-related properties [\[20\]](#).

The aim of this project is to study current understanding of cell types for development a comprehensive ontological model in Wikidata for cell types. We are reviewing the single-cell literature, refining and formalizing concepts for cell type delimitation and exploring their application in the Wikidata database. At the same time, we are exploring the use of Natural Language Processing tools, in combination with expert annotations tools to automate knowledge extraction from scientific articles in the scope of the Human Cell Atlas. The specific goals outlined in the approved project were:

- Build a data model to capture the main properties to describe a human cell. Provide working definitions of cell types and states and their characteristics.
- Extract and add to Wikidata pieces of information regarding Human Cell Atlas publications to build gold standards. Use this information to develop machine learning tools to extract knowledge from publications.
- Create tools to make data from the underlying knowledge graph accessible employing tools from network theory.

Achievements

In the first year of the project, we addressed 3 different facets of the project. The first facet dealt with the concept of cell type itself, in the goal of building theoretical models for the formalization of knowledge. The second facet was aimed at the practical application of the concepts. We adapted a public database of cell type markers, Panglao DB, to the Wikidata format, made the information available through Wikidata and performed SPARQL (SPARQL Protocol and RDF Query Language) queries to glance at the information on the database. Finally, we manually annotated articles in the Human Cell Atlas Publications collection to gain insights on the types of knowledge available about cell types and paving the way for automated and crowdsourced approaches.

Concept of cell types

The definition of the concept of “cell type” is currently a topic of debate by the biomedical community. [\[1,21,22,23,24,25,26,27,28,29,30,31\]](#). In an opinion article published in Cell Systems in 2017, a series of researchers presented their views on the conceptual definition of ‘cell type’ in the context of a mature organism [\[21\]](#). Many scientists believed that cell functions have a core role in defining cell types, which is a slippery road, as the very meaning of “function” in biology is elusive [\[32\]](#). The opinions were varied, and no consensus was achieved.

As much as different concepts of species coexist [\[33\]](#), our quest to define cell types may take various forms. The challenge of representing cell types in the context of evolution is conceptually different

from representing cell types in biomedical experimentation. In that second direction, the groundwork of the Cell Ontology [34,35,36] and CELDA [37] and the contributions of the International Workshop on Cells in Experimental Life Sciences series [38,39] are notable.

In this work, we targeted the question: which cell type definition allows crafting coherent biological statements? The goal is to not say what cell types *are*, but what they can be for a consistent representation on a knowledge graph, like Wikidata. We also assume that biological statements that derive from experimentation should be logically valid, and thus aim to distill ways to resolve ambiguities in natural language statements makes knowledge unresolvable (i.e. neither true nor false). Furthermore, for the practical purpose adopted here, we avoid the dissection of the differences between persistent classes of cells (often called “cell types”) or the transient, fugacious classes of cells (often called “cell states”) (see “Definition of cell identity” section in [40] for an example). Even though such a distinction is an essential topic for theoretical research, it is not required to represent formally biomedical experiments.

To facilitate communication among life scientists, in a preprint derived from this PHD project[41], we proposed naming conventions for different cell types classes. Much of the literature mixes cell types in one species (e.g., when dealing with a cell type as an evolutionary unit) or multiple species (e.g., in the Cell Ontology). It is useful to distill these different concepts into names. Given the importance of the species’ concept in biological classification [42], we derive a species-centric view on the naming of classes of cell types. The four classes (Figure 1) we propose are as follows:

- archetypes, for when the taxonomic scope of the type is beyond the level of species; for example, “mammal neutrophils.”
- *sensu stricto* cell types, for when the taxonomic scope of the type corresponds to a single species; for example, *Mus musculus* neutrophils.”
- infratypes, for when the taxonomic scope is below the level of species; for example, considering the mouse strain “C57BL/6J,” “neutrophils from C57BL/6J mice.”
- technotypes, for specific, experimentally defined cell types that harbor in their definition the precise conditions of the cells sampled; “2-month-old male C57BL/6J, Ly-6G⁺ CD11b⁺ M-CSF R⁻ CD244⁻ neutrophils.”

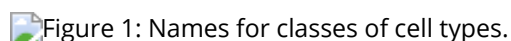
Figure 1: Names for classes of cell types.

Figure 1: Names for classes of cell types.

The 4 different categories of cell types help us to better organize the knowledge about cell types. Even though individual articles and databases often have species-neutral names, the information often comes from experiments with a single strain of a single species. Two articles might call by the same name cells that come from different animals, or were selected by different protocols. Large scale knowledge management requires an organized way of representing those details.

The division between archetypes and *sensu stricto* cell types is of special importance in the context of biocuration and annotation of data. Associations like the HUGO Gene Nomenclature Committee and UniProt organize names and identifiers for genes and proteins in single species. Thus, if we want to annotate marker genes, for example, we need to associate them to a species-specific cell type (a *sensu stricto* cell type) instead of the more vague association to a species-neutral type. That might seem obvious, but current standards still use identifiers that are species-neutral (e.g. in the reference HuBMAP app; <https://azimuth.hubmapconsortium.org/references/>)

The ontological discussion on the classes of cell types, thus, extends the current state-of-the-art and introduce new ways to organize our knowledge about cells. Notably, the technotype and the infratype are, currently, mostly theoretical constructs and almost no resources deal with cell types at the level of strains or below. The division of archetypes and *sensu stricto* cell types, on the other hand, was

already instrumental for the integration of the Panglao database of cell markers to Wikidata, described in the next session.

PanglaoDB integration to Wikidata

Annotation of the HCA corpus

References

1. **An era of single-cell genomics consortia**
Yoshinari Ando, Andrew T Kwon, Jay W Shin
Experimental and Molecular Medicine (2020-09-15) <https://www.wikidata.org/wiki/Q99418649>
DOI: [10.1038/s12276-020-0409-x](https://doi.org/10.1038/s12276-020-0409-x)
2. **The Human Cell Atlas.**
Aviv Regev, Sarah Teichmann, Eric Lander, Amir Giladi, Christophe Benoist, Ewan Birney, Bernd Bodenmiller, Peter Campbell, Piero Carninci, Menna R Clatworthy, ... Human Cell Atlas Meeting Participants
eLife (2017-12-05) <https://www.wikidata.org/wiki/Q46368626>
DOI: [10.7554/elife.27041](https://doi.org/10.7554/elife.27041)
3. **Everyone needs a data-management plan**
Nature
(2018-03-15) <https://www.wikidata.org/wiki/Q56524391>
DOI: [10.1038/d41586-018-03065-z](https://doi.org/10.1038/d41586-018-03065-z)
4. **About the Data Coordination Platform**
HCA Data Portal
<https://data.humancellatlas.org/about/>
5. **The Human Cell Atlas White Paper**
Aviv Regev, Sarah Teichmann, Orit Rozenblatt-Rosen, Michael JT Stubbington, Kristin Ardlie, Amir Giladi, Paola Arlotta, Gary D Bader, Christophe Benoist, Moshe Biton, ... Human Cell Atlas Organizing Committee
(2018-10-11) <https://www.wikidata.org/wiki/Q104450645>
6. **Literature-based discovery**
Wikipedia
(2021-01-24) https://en.wikipedia.org/w/index.php?title=Literature-based_discovery&oldid=1002467308
7. **Literature Based Discovery: models, methods, and trends.**
MSSam Henry, Bridget T McInnes
Journal of Biomedical Informatics (2017-08-21) <https://www.wikidata.org/wiki/Q38371706>
DOI: [10.1016/j.jbi.2017.08.011](https://doi.org/10.1016/j.jbi.2017.08.011)
8. **Online tools to support literature-based discovery in the life sciences.**
Marc Weeber, Marc Weeber, Jan A Kors, Jan A Kors, Barend Mons
Briefings in Bioinformatics (2005-09-01) <https://www.wikidata.org/wiki/Q36280460>
DOI: [10.1093/bib/6.3.277](https://doi.org/10.1093/bib/6.3.277)
9. **Unsupervised word embeddings capture latent knowledge from materials science literature**
Vahe Tshitoyan, John Dagdelen, Leigh Weston, Alexander Dunn, Ziqin Rong, Olga Kononova, Kristin A Persson, Gerbrand Ceder, Anubhav Jain
Nature (2019-07-03) <https://www.wikidata.org/wiki/Q91595456>
DOI: [10.1038/s41586-019-1335-8](https://doi.org/10.1038/s41586-019-1335-8)
10. **Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis.**

Nadine Bakkar, Tina Kovalik, Ileana Lorenzini, Scott Spangler, Alix Lacoste, Kyle Sponaugle, Philip Ferrante, Elenee Argentinis, Rita Sattler, Robert Bowser
Acta Neuropathologica (2017-11-13) <https://www.wikidata.org/wiki/Q47406275>
DOI: [10.1007/s00401-017-1785-8](https://doi.org/10.1007/s00401-017-1785-8)

11. **Ontologies for the life sciences**

Steffen Schulze-Kremer, Barry Smith
(2005-11-15) <https://www.wikidata.org/wiki/Q105870680>
DOI: [10.1002/047001153x.g408213](https://doi.org/10.1002/047001153x.g408213)

12. **The Gene Ontology resource: enriching a GOLD mine**

Gene Ontology Consortium
Nucleic Acids Research (2020-12-08) <https://www.wikidata.org/wiki/Q104130127>
DOI: [10.1093/nar/gkaa1113](https://doi.org/10.1093/nar/gkaa1113)

13. **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium**

M Ashburner, CA Ball, Judith A Blake, David Botstein, H Butler, JMichael Cherry, AP Davis, K Dolinski, Selina S Dwight, JT Eppig, ... Gavin Sherlock
Nature Genetics (2000-05-01) <https://www.wikidata.org/wiki/Q23781406>
DOI: [10.1038/75556](https://doi.org/10.1038/75556)

14. **The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration**

Barry Smith, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, Karen Eilbeck, Amelia Ireland, Christopher J Mungall, ... Suzanna Lewis
Nature Biotechnology (2007-11-01) <https://www.wikidata.org/wiki/Q19671692>
DOI: [10.1038/nbt1346](https://doi.org/10.1038/nbt1346)

15. **Toward an epistemology of Wikipedia**

Don Fallis
Journal of the Association for Information Science and Technology (2008-08-01)
<https://www.wikidata.org/wiki/Q101955295>
DOI: [10.1002/asi.20870](https://doi.org/10.1002/asi.20870)

16. **From Freebase to Wikidata: The Great Migration**

Thomas Pellissier Tanon, Denny Vrandečić, Sebastian Schaffert, Thomas Steiner, Lydia Pintscher
Proceedings of the 25th International Conference on World Wide Web (2016-01-01)
<https://www.wikidata.org/wiki/Q24074986>
DOI: [10.1145/2872427.2874809](https://doi.org/10.1145/2872427.2874809)

17. **Wikidata: A platform for data integration and dissemination for the life sciences and beyond**

Elvira Mitraka, Andra Waagmeester, Sebastian Burgstaller-Muehlbacher, Lynn M Schriml, Andrew I Su, Benjamin M Good
Cold Spring Harbor Laboratory (2015-11-16) <https://doi.org/gg9dk4>
DOI: [10.1101/031971](https://doi.org/10.1101/031971)

18. **Wikidata as a knowledge graph for the life sciences**

Andra Waagmeester, Gregory Stupp, Sebastian Burgstaller-Muehlbacher, Benjamin M Good, Malachi Griffith, Obi Griffith, Kristina Hanspers, Henning Hermjakob, Toby Hudson, Kevin Hybiske, ... Andrew I Su
eLife (2020-03-17) <https://www.wikidata.org/wiki/Q87830400>
DOI: [10.7554/elife.52614](https://doi.org/10.7554/elife.52614)

19. **Wikidata as a knowledge graph for the life sciences**

Andra Waagmeester, Gregory Stupp, Sebastian Burgstaller-Muehlbacher, Benjamin M Good, Malachi Griffith, Obi L Griffith, Kristina Hanspers, Henning Hermjakob, Toby S Hudson, Kevin Hybiske, ... Andrew I Su
eLife (2020-03-17) <https://doi.org/gggqgc6>
DOI: [10.7554/elife.52614](https://doi.org/10.7554/elife.52614) · PMID: [32180547](https://pubmed.ncbi.nlm.nih.gov/32180547/) · PMCID: [PMC7077981](https://pubmed.ncbi.nlm.nih.gov/PMC7077981/)

20. **Wikidata: A large-scale collaborative ontological medical database**
Houcemeddine Turki, Thomas Shafee, Mohamed Ali Hadj Taieb, Mohamed Ben Aouicha, Denny Vrandečić, Diptanshu Das, Helmi Hamdi
Journal of Biomedical Informatics (2019-11) <https://doi.org/gg9dnt>
DOI: [10.1016/j.jbi.2019.103292](https://doi.org/10.1016/j.jbi.2019.103292) · PMID: [31557529](https://pubmed.ncbi.nlm.nih.gov/31557529/)
21. **What Is Your Conceptual Definition of "Cell Type" in the Context of a Mature Organism?**
Paul Blainey, Hans Clevers, Cole Trapnell, Ed Lein, Emma Lundberg, Alfonso Martinez Arias, Joshua R Sanes, Jay Shendure, James Eberwine, Junhyong Kim, ... Mathias Uhlén
Cell systems (2017-03-01) <https://www.wikidata.org/wiki/Q87649649>
DOI: [10.1016/j.cels.2017.03.006](https://doi.org/10.1016/j.cels.2017.03.006)
22. **A periodic table of cell types**
Bo Xia, Itai Yanai
Development (2019-06-15) <https://doi.org/ggctwf>
DOI: [10.1242/dev.169854](https://doi.org/10.1242/dev.169854) · PMID: [31249003](https://pubmed.ncbi.nlm.nih.gov/31249003/) · PMCID: [PMC6602355](https://pubmed.ncbi.nlm.nih.gov/PMC6602355/)
23. **Exciting times to study the identity and evolution of cell types**
Maria Sachkova, Pawel Burkhardt
Development (2019-09-15) <https://doi.org/ghdb9v>
DOI: [10.1242/dev.178996](https://doi.org/10.1242/dev.178996) · PMID: [31537583](https://pubmed.ncbi.nlm.nih.gov/31537583/)
24. **The Human Cell Atlas: from vision to reality.**
Orit Rozenblatt-Rosen, Michael JT Stubbington, Aviv Regev, Sarah Teichmann
Nature (2017-10-01) <https://www.wikidata.org/wiki/Q47565008>
DOI: [10.1038/550451a](https://doi.org/10.1038/550451a)
25. **Human Cell Atlas and cell-type authentication for regenerative medicine**
Yulia Panina, Peter Karagiannis, Andreas Kurtz, Glyn N Stacey, Wataru Fujibuchi
Experimental and Molecular Medicine (2020-09-15) <https://www.wikidata.org/wiki/Q99418657>
DOI: [10.1038/s12276-020-0421-1](https://doi.org/10.1038/s12276-020-0421-1)
26. **A community-based transcriptomics classification and nomenclature of neocortical cell types**
Rafael Yuste, Michael J Hawrylycz, Nadia Aalling, Argel Aguilar-Valles, Detlev Arendt, Rubén Armañanzas, Giorgio A Ascoli, Concha Bielza, Vahid Bokharaie, Tobias B Bergmann, ... Ed S Lein
Nature Neuroscience (2020-08-24) <https://www.wikidata.org/wiki/Q98665291>
DOI: [10.1038/s41593-020-0685-8](https://doi.org/10.1038/s41593-020-0685-8)
27. **The evolving concept of cell identity in the single cell era**
Samantha A Morris
Development (2019-06-27) <https://www.wikidata.org/wiki/Q93086971>
DOI: [10.1242/dev.169748](https://doi.org/10.1242/dev.169748)
28. **Implications of Epigenetic Variability within a Cell Population for "Cell Type" Classification**
Inna Tabansky, Joel Stern, Donald W Pfaff
Frontiers in Behavioral Neuroscience (2015-12-16) <https://www.wikidata.org/wiki/Q26770736>
DOI: [10.3389/fnbeh.2015.00342](https://doi.org/10.3389/fnbeh.2015.00342)

29. **Geometry of the Gene Expression Space of Individual Cells**
Yael Korem, Pablo Szekely, Yuval Hart, Hila Sheftel, Jean Hausser, Avi Mayo, Michael E Rothenberg, Tomer Kalisky, Uri Alon
PLOS Computational Biology (2015-07-10) <https://www.wikidata.org/wiki/Q35688096>
DOI: [10.1371/journal.pcbi.1004224](https://doi.org/10.1371/journal.pcbi.1004224)
30. **Evolution of Cellular Differentiation: From Hypotheses to Models**
Pedro Márquez-Zacarías, Rozenn M Pineau, Marcella Gomez, Alan Veliz-Cuba, David Murrugarra, William C Ratcliff, Karl J Niklas
Trends in Ecology & Evolution (2020-08-20) <https://www.wikidata.org/wiki/Q98633613>
DOI: [10.1016/j.tree.2020.07.013](https://doi.org/10.1016/j.tree.2020.07.013)
31. **Inferring cell type innovations by phylogenetic methods-concepts, methods, and limitations**
Koryu Kin, Koryu Kin
Journal of Experimental Zoology. Part B: Molecular and Developmental Evolution (2015-10-14)
<https://www.wikidata.org/wiki/Q40436539>
DOI: [10.1002/jez.b.22657](https://doi.org/10.1002/jez.b.22657)
32. **The meanings of 'function' in biology and the problematic case of de novo gene emergence**
Diane Marie Keeling, Patricia Garza, Charisse Michelle Nartey, Anne-Ruxandra Carvunis
eLife (2019-11-01) <https://doi.org/ggjnmy>
DOI: [10.7554/elife.47014](https://doi.org/10.7554/elife.47014) · PMID: [31674305](https://pubmed.ncbi.nlm.nih.gov/31674305/) · PMCID: [PMC6824840](https://pubmed.ncbi.nlm.nih.gov/PMC6824840/)
33. **Species Concepts and Species Delimitation**
Kevin De Queiroz
Systematic Biology (2007-12) <https://doi.org/c34kzf>
DOI: [10.1080/10635150701701083](https://doi.org/10.1080/10635150701701083) · PMID: [18027281](https://pubmed.ncbi.nlm.nih.gov/18027281/)
34. **An ontology for cell types**
Jonathan Bard, Sue Rhee, Michael Ashburner
Genome Biology (2005-01-01) <https://www.wikidata.org/wiki/Q21184168>
DOI: [10.1186/gb-2005-6-2-r21](https://doi.org/10.1186/gb-2005-6-2-r21)
35. **Logical Development of the Cell Ontology**
Terrence F Meehan, Anna Maria Masci, Amina Abdulla, Lindsay G Cowell, Judith A Blake, Christopher J Mungall, Alexander D Diehl
BMC Bioinformatics (2011-01-05) <https://doi.org/c7kw6x>
DOI: [10.1186/1471-2105-12-6](https://doi.org/10.1186/1471-2105-12-6) · PMID: [21208450](https://pubmed.ncbi.nlm.nih.gov/21208450/) · PMCID: [PMC3024222](https://pubmed.ncbi.nlm.nih.gov/PMC3024222/)
36. **The Cell Ontology 2016: enhanced content, modularization, and ontology interoperability**
Alexander D Diehl, Terrence F Meehan, Yvonne M Bradford, Matthew H Brush, Wasila M Dahdul, David S Dougall, Yongqun He, David Osumi-Sutherland, Alan Ruttenberg, Sirarat Sarntivijai, ... Christopher J Mungall
Journal of Biomedical Semantics (2016-07-04) <https://doi.org/gg99b9>
DOI: [10.1186/s13326-016-0088-7](https://doi.org/10.1186/s13326-016-0088-7) · PMID: [27377652](https://pubmed.ncbi.nlm.nih.gov/27377652/) · PMCID: [PMC4932724](https://pubmed.ncbi.nlm.nih.gov/PMC4932724/)
37. **CELDA -- an ontology for the comprehensive representation of cells in complex systems**
Stefanie Seltsmann, Harald Stachelscheid, Alexander Damaschun, Ludger Jansen, Fritz Lekschas, Jean-Fred Fontaine, Throng Nghia Nguyen-Dobinsky, Ulf Leser, Andreas Kurtz
BMC Bioinformatics (2013-07-17) <https://www.wikidata.org/wiki/Q21284308>
DOI: [10.1186/1471-2105-14-228](https://doi.org/10.1186/1471-2105-14-228)

38. **Cells in experimental life sciences - challenges and solution to the rapid evolution of knowledge**
Sirarat Sarntivijai, Alexander D Diehl, Yongqun He
BMC Bioinformatics (2017-12-21) <https://doi.org/gg99b7>
DOI: [10.1186/s12859-017-1976-2](https://doi.org/10.1186/s12859-017-1976-2) · PMID: [29322916](https://pubmed.ncbi.nlm.nih.gov/29322916/) · PMCID: [PMC5763506](https://pubmed.ncbi.nlm.nih.gov/PMC5763506/)
39. **Cells in Experimental Life Sciences (CELLS-2018): capturing the knowledge of normal and diseased cells with ontologies**
Sirarat Sarntivijai, Yongqun He, Alexander D Diehl
BMC Bioinformatics (2019-04-25) <https://doi.org/gg99b8>
DOI: [10.1186/s12859-019-2721-9](https://doi.org/10.1186/s12859-019-2721-9) · PMID: [31272374](https://pubmed.ncbi.nlm.nih.gov/31272374/) · PMCID: [PMC6509796](https://pubmed.ncbi.nlm.nih.gov/PMC6509796/)
40. **The Human Cell Atlas: Technical approaches and challenges.**
Chung Chau Hon, Jay W Shin, Piero Carninci, Michael JT Stubbington
Briefings in functional genomics (2017-10-28) <https://www.wikidata.org/wiki/Q48563763>
DOI: [10.1093/bfpg/elx029](https://doi.org/10.1093/bfpg/elx029)
41. **Towards a pragmatic definition of cell type**
Tiago Lubiana, Helder I Nakaya
Authorea, Inc. (2021-01-04) <https://doi.org/ghrxwf>
DOI: [10.22541/au.160979530.02627436/v1](https://doi.org/10.22541/au.160979530.02627436/v1)
42. **PhyloCode** <https://www.wikidata.org/wiki/Q1189395>