Inverting {{Citation needed}}: Critical Design Reflection of a Citation Learning Game

Simon Knight[a,c*], Antonette Shibani[a], Heather Ford[b,c], Simon Chambers[a]

[a]TD School, University of Technology Sydney, Sydney, Australia; [b]School of Communications, University of Technology Sydney, Sydney, Australia; [c]Centre for Research on Education in a Digital Society (CREDS), University of Technology Sydney, Sydney, Australia

* simon.knight@uts.edu.au, TD School, PO BOX 123, UTS, Broadway, NSW 2007, Australia.

Inverting {{Citation needed}}: Critical Design Reflection of a Citation Learning Game

Abstract

Discussion of the potential of open data for impact - including in education - has grown. Wikipedia, the world's largest open educational resource, is a prime source of open data. This data is used in a variety of contexts to understand learning resources, inform collaborative learning contexts, and to evaluate the quality of individual encyclopaedic content. A key aspect of the sociotechnical character of this Wikipedia data is that it arises in a platform-based social dynamic, governed by norms and conventions, and negotiated by sets of volunteers for the particular purpose of writing an encyclopedia.

Verifiability through external sources is one central tenet of Wikipedia. Similarly, scholarly writing shares focus on citation as a crucial signifier of warrants for claims. This paper uses citation as a *boundary object*, through the exegesis of a reflective design case regarding a citation learning game (Citelearn). This case description highlights the continuities and discontinuities in citation across sites (Wikipedia and scholarly writing), and the significance of translation of such data across sites. The paper aims to exemplify a method for close case analyses of learning data, its representation, and re-representation, drawing attention to the need for such analysis in adopting data-infrastructure across contexts.

Keywords: open data; boundary objects; infrastructure; socio-material; case study

Introduction

The promise of ‘big data’ and ‘open data’ relies on assumptions regarding the ways in which data should be translated to move from one context to another (Koesten et al., 2020). While data is increasingly available, a relatively small number of datasets form the basis of significant work in machine learning (Koch et al., 2021). In this paper, we (1) introduce the role of open data and Wikipedia as learning infrastructure and resource, (2) outline a method – the reflective design case – that supports probing of design artefacts, (3) apply this method

to the boundary object of citation, a socio-technical feature of scholarly and Wikipedia writing, using the design of the novel CiteLearn tool to do so. By framing citation as a boundary object and the game as a learning tool, the focus is on how we understand the ways that data, practices such as citation, and infrastructure, are entangled in ways that may be hidden by processes of moving (data and practices) from one context to another. As the following sections will thus unpack through discussion of data infrastructures and the details of our case, the work addresses two key objectives:

1. To use the boundary object of citation – a feature of writing that is readily observable across contexts – to probe assumptions regarding the socio-technical character of citation data from two knowledge communities (Wikipedia and academic writing in our case);
2. In doing this, to demonstrate a critical design reflection method using an algorithm game focused on citation, as an approach to consider the ways that data is taken up, adopted, and applied in new contexts that produce data, and the implications of this for learning and technologies.

Understanding when and how citation data can move across contexts is critical for understanding and developing learning technologies and research that is built on open data – or indeed, datasets that are merged from multiple contexts. Merging datasets offers enormous potential for developing new insights and for advancing learning technologies, but data transfer often surfaces important context information that makes transfer to new contexts difficult, sometimes impossible. It is our goal in this paper to articulate these considerations in the context of citation data in order to inform future efforts for data reuse.

Data Infrastructures for Learning

Open data – data that is freely available for re-use and re-distribution – underpins much of our information infrastructure. It is used in cases from mapping traffic and public transport

journey times via your choice of app, to the sharing of data on COVID-19 cases, adverse events, and hospitalisations, and to provide insights to both policy makers and the public. Education has also seen a growth in open data, as part of a broader open education movement (Edwards, 2015; Winn, 2015). Here, cases range across for example, data and representations regarding (contested) national and local accountability structures, and 'linked data' structures that map out open educational resources and their connections to features such as curricula content (Atenas & Havemann, 2019; d'Aquin, 2016). While the original promise of open data rested on public availability and permissive licenses, and there have been releases of such data for educational purposes (Kuzilek et al., 2017), recent shifts in data infrastructure have been less open, and more commercially oriented (Williamson, 2019).

Closely tied to this growing use of data, an area of particular recent interest has been 'learning analytics' and associated fields such as Artificial Intelligence in Education (AIEd), educational data mining (EDM), and Learning at Scale (L@S). These fields, in various ways, attempt to design or observe learning environments for the collection of data about learning, applying data science approaches to analyse the data for the purpose of understanding and supporting that learning. Open data is useful for learning analytics projects because it provides efficient access to datasets that are often well described by the communities that produce them.

Wikipedia and education

A particular context of interest in recent open data and educational technology discussion has been the use of wikis, platforms that support collaborative authoring of rich hypermedia content (Zheng et al., 2015). This interest has included focus on the world's largest open educational resource (by visitors, editors, or content) in the form of Wikipedia (Hylén, 2006; Voss, 2005). Wikis in general have been of interest for the potential to offer models for collaborative learning (Chen et al., 2021). There has been further interest from education

because of potential lessons from the crowdsourcing model used by Wikipedia – in which 'anyone' may contribute knowledge. Wikipedia provides a model for developing an updatable knowledge base for a particular project – and demonstrates the potential for creating other forms of resources such as problem sets (Heffernan et al., 2016; Prpic et al., 2015).

Moreover, Wikipedia specifically has been of interest to learning researchers and others, with a particular focus on three interrelated concerns. Firstly, the coverage and accuracy of the content has been a significant focus, as discussed in a systematic review of Wikipedia research (Mesgari et al., 2015), with generally favourable outcomes compared to other encyclopaedic sources. However, this research thread has also noted the significant gaps in coverage, particularly around language (English is by far the largest Wikipedia), culture and geography (Graham et al., 2015), and gender (Redi et al., 2021). These gaps intersect with the second theme of interest in readership and community participation (Redi et al., 2021), with significantly more men (and people from the Global North) editing the encyclopaedia. These features also interact with the third area of focus, namely, the community norms and structures through which editors build and quality control the encyclopaedia, and into which new editors are inducted. These rules intersect with Wikipedia's gaps insofar as, for example, they may create barriers for new editors (who do not know the core rules), or/and they restrict the kinds of sources that can be used to support a claim's addition, in such a way as to perpetuate other biases including linguistic and gender gaps (Redi et al., 2021).

Despite these concerns, Wikipedia has succeeded in building the largest encyclopaedia, grounded in a largely self-governing volunteer community, who work together, and create the rules that underpin it. For this reason, learning researchers are interested in the practice of editing *Wikipedia* as a specific form of collaboration platform,

both to understand how people learn to become 'Wikipedians' – a part of the community of editors – and what skills and knowledge this might enable for broader purposes (abilities to verify claims, knowledge of content, community discussion, etc.), and the potential of editing Wikipedia as an authentic output to support engagement in students (Brailas et al., 2015; Carver et al., 2012; Haren et al., 2020; Y. Li et al., 2016; Oliver, 2015). These issues are, of course, also of central interest to Wikimedians (volunteers involved in all Wikimedia projects, including Wikipedia), and the charitable body, the Wikimedia Foundation and its global chapters and user groups. In particular, issues of skills and knowledge are important in relation to their goals of addressing the gaps described above, and in supporting new editors to become part of the community of editors (e.g. A. Li et al., 2020; Vetter et al., 2020).

Wikipedia and Data Infrastructures

These intersections of critical foci on Wikipedia – content, community, and norms – are also of wide significance on the web. While Wikipedia is an encyclopaedia, with its sister projects (particularly WikiCommons, which hosts media files, and WikiData, which hosts structured data), this open knowledge acts as underpinning infrastructure for a significant portion of the web. That is in part because the resources and their history may be obtained and mined freely, and have thus provided a rich dataset for computer scientists, including in information retrieval, natural language processing, and ontology building (Mehdi et al., 2017). These advances have, in particular, provided infrastructure for automated Q&A and knowledge graph systems, including those encountered by users in the form of 'info boxes' containing key facts about entities provided in response to many queries (Ford, 2015; Tiwari et al., 2021). These applications are possible due to the large and structured nature of the data, and its open license that allows commercial use and the building of computational models.

This ability and willingness to build on Wikipedia's open knowledge resources is bound up with its content, community, and norms. Wikipedia's three core content policies

comprise of: neutral point of view, verifiability, and no original research ("Wikipedia," 2021c). These intersect in ways we will unpack later in this piece, but here we particularly flag the significance of the verifiability principle which requires that statements in Wikipedia be supported by external sources ("Wikipedia," 2021e), and where no such sources are provided, a {{Citation needed}} tag should be applied ("Wikipedia," 2021b). These tags are applied manually by the community of editors, following a set of norms around acceptable claims and sources.

However, there has been some interest in developing automated approaches to identifying statements in need of a citation, building on the open dataset of sentences labelled {{Citation needed}} (Chou et al., 2020; Redi et al., 2019). Such analysis would serve multiple purposes, including automated fact-checking, prioritisation for editor contributions, support for editing poorly expressed claims, and support for onboarding new editors into the norms and culture of the editor community. Citation data, seen through a socio-material lens of practices and artefacts, is also of significant interest for education, given the role of citation in scholarly writing, and the importance of learning to cite for students, which we further unpack through the critical design reflection discussed next.

This Paper: A Critical Design Reflection of Citations

Citations are an important socio-technical feature of the Wikipedia system, and an important practice for new editors to learn as they become a part of the community. In parallel, citation is of course also crucial to most higher education contexts, with some lecturers engaging students in the practice of editing wikis or indeed Wikipedia for its range of benefits for learning. This engagement is bound up with the wider socio-technical infrastructure of Wikipedia, and the variety of synergies and tensions at play in the aims and means of university education and Wikipedia with its wider relationship as web infrastructure.

This paper unpacks this relationship, using the shared feature of citation as a boundary object (Akkerman & Bakker, 2011; Bowker & Star, 1999; Star & Griesemer, 1989), providing a particular kind of artefact in writing that implies commitments to knowledge systems (Van der Veer Martens, 2001). Citation is a boundary object in the sense that there is a common recognisable feature – the citation and references – across sites, while holding characteristics that vary in sometimes hidden, yet important, ways across these sites. We explore the socio-technical features of citation in the context of a citation learning game (CiteLearn), built on the Wikipedia {{Citation needed}} dataset. In taking this approach we seek to foreground the ways in which data – citation data – and its classification, application, and use in one context have implications for classificatory processes elsewhere. We highlight the entanglement of learning and wider infrastructures with these classificatory processes.

We surface these characteristics using a 'game' in the dual sense of a learning technology targeting feedback on a particular set of skill or knowledge (Plass et al., 2015), and in the sense of an 'algorithm game' or 'toy algorithm' (Gulson, Benn, et al., 2021; Gulson, Swist, et al., 2021; Hao & Stray, 2019; Stray & Hao, n.d.) – an algorithmic artefact designed to help understand the underlying structure of the tool and its application. In taking this approach, we draw on the concept of a 'data journey' (Bates et al., 2016; Howard et al., 2021; Leonelli, 2020) to highlight the socio-material construction, use, and importing into contexts, of data from systems such as Wikipedia, to draw specific attention to citation, through the Citelearn tool. We use the Citelearn game as a case study for the critical design reflection method (Knight et al., 2023), highlighting the complexity of the 'taking from' Wikipedia to 'apply to' other educational contexts. In doing so, a mutual lens is provided, with the learning context providing a lens onto the original data, its context within Wikipedia, and potential for learning; and that original Wikipedia context providing a lens onto our learning spaces. As such, we highlight how, through a close analysis of our feature – citation

– brought about in moving data from one context to another, we can draw attention to key aspects of our two contexts and the role of citation in both of them.

This study has important implications for learning and education using open data and for educational initiatives involving Wikipedia and other platforms (ranging from social media tools, community software platforms such as Github, online fora, and so on). This is particularly significant given that Wikipedia acts as both an open *data* source (for modelling data), and *learning* source (through its content, community structures, and norms), with instructors adopting the platform for assignments. Our aim is to highlight the ways in which datasets can be adopted and adapted, and to foreground the ways in which such systems trained on 'big data' rest both on detailed technical features (about how to process and output data) and the socio-technical characteristics of their ongoing development (including policy, norms and practices). The study enables us to think through the considerations for designing algorithmic learning tools using open data and to understand the local peculiarities of communities like Wikipedia when engaging students in learning activities on the platform itself.

In the remainder of this paper we present a design case, using a game built on Wikipedia data as the unit of analysis for our case study. Following the steps for critical design reflection (Knight et al., 2023) we first discuss the norms and conventions in both Wikipedia and scholarly citation practices. We then provide focal points in our discussion of citation as a boundary object. The paper does not focus on evaluation of the tool itself, or its use with students; the significance of the work is not grounded in its 'success' or otherwise in this regard. Rather, the contribution of this paper is to highlight how open data transfer can occur through seemingly paralleled features (citation in this case) and how the challenges of this transfer inform both our navigation of data itself, and of the practices that the data encodes (in Wikipedia) and is intended to represent (in learning contexts).

Methodology: Inverting {{Citation needed}} Through a Reflective Design Case

Designs encode judgements about a range of features of the world and user needs, and thus deliberate analysis of design processes and artefacts can provide insights into both purpose and delivery for different groups (Bardzell et al., 2014). Our reflective design case (Knight et al., 2023) draws out *socio-material* aspects of design and the roles of technology designs as – in a sense – making 'claims' about the world (Carroll & Rosson, 1992). Our approach provides a re-representation of a design as a case, through the lens of reflection-on-action, inviting us to explore thick descriptions, and imagine alternative designs.

This approach follows a tradition in infrastructure and platform studies that have emerged as useful ways to understand our relation to media and data technologies, and their mediational properties. These approaches recognise that software, media technologies, and tools such as predictive algorithms are *designed* objects that emerge from, and act in, contexts of use (Williamson, 2015). The central claim here is not one of technological-determinism, but rather a sociotechnical perspective that *designs* involve *people and communities* and their norms, and therefore that technologies can be studied as objects that instantiate or materialize views of the world, just as they "in turn can influence and shape how people act in and on that world." (Williamson, 2015, p. 84) Wikipedia has been one such object of study, through the lens of knowledge infrastructures, recognising the digital mediation and production of knowledge and the role of human actors, their tools, and community norms (or 'managerial dynamics') in this (Ford & Wajcman, 2017). Wikipedia can thus be read both as a representation of knowledge, but also an active site for knowledge production in which decisions are made about what constitutes credible knowledge and who are considered authorities (Ford & Wajcman, 2017).

We use a reflective design case as a structured design-oriented case study, following steps outlined in (Knight et al., 2023) to:

1. Identify a unit of analysis: For learning technologies, this should comprise an inter-connected design proposition or propositions connecting the technology experience and the learning. For example, a specific feature (e.g., plagiarism checking), and learning design (e.g., pre- or post-submission access to the plagiarism reports).
2. Situate the unit of analysis in relation to extant conventions and norms: In learning contexts, this will comprise familiar sets of features around learning outcomes, social and material mediation, roles of teachers and students, and so on.
3. Isolate the focal aspects of the design: In learning contexts these may be features that are specific to the intended learning outcome, construct, or social relation, or – as in our case – related to the intent, implied 'claim', or continuities and discontinuities, in relation to the design features.

This reflection describes our own navigation of citation as a boundary object, using the CiteLearn algorithm game which we designed with the intent of testing with learners. As such, we are not neutral observers of the tool but rather provide reflection-on-action drawing on our experience in the design process. Nevertheless, it is worth highlighting that even if we had deliberately sought to create a game that represented particular features of citation or its use in one of our contexts, or/and to obscure any discontinuities, the aim of the reflection here is to consider how this instantiation of continuities and discontinuities into the material artefact of the CiteLearn game helps us to make visible these sociotechnical features. Our case approach helps make concrete and highlight tensions, in order to understand and navigate these.

Reflective Design Case

Unit of Analysis: The CiteLearn Tool

The CiteLearn tool was designed and built as an online interface to support learning how to cite sources. The tool was specifically designed to address the principle of 'verifiability' in Wikipedia, and scholarly citation conventions in university level coursework assignments. The tool was instigated by a grant call from WikiCred, a North America based organisation of "Wikimedians Strengthening Knowledge and News Credibility on the Internet", whose primary activity is the delivery of seed grants to support their mission.

The aim of the CiteLearn project was to build learning tasks for the improvement of citation practices that were scalable across teaching contexts. To do so, the tool draws on the open data of Wikipedia's citation practices – specifically the Citation Need dataset of 'Featured Articles' which was compiled to train the development of machine learning models (Chou et al., 2020; Redi et al., 2019). Featured Articles represent those Wikipedia articles which are regarded by editors as best exemplifying the norms of excellence which govern content published on Wikipedia. These criteria are explicitly articulated ("Wikipedia," 2021a) and include not only aspects relating to the standard of writing (including the appropriate citation of high-quality sources to support claims) but also the subject of the article – which should be "stable" and "not subject to ongoing 'edit wars'".

In addition to incorporating the Citation Need models developed by Redi et al. to predict which sentences required a citation, CiteLearn also used this same Featured Articles data to train alternative models to further investigate the applicability of the data to teaching about citation practices more generally. There are a range of other datasets that have been developed from this data for use in tools such as this, including Cite Unseen (which detects bias in citations) ("User," 2021), and Citation Reason (which classifies reasons for a citation being given) (Chou et al., 2020). We anticipated this game being used to onboard people into

Wikipedia citation practices, and more broadly to develop the skills associated with citation (credibility checking, verifiability, etc.).

We aimed to build a tool that provided students with authentic learning experiences for understanding the purpose and usefulness of citation practice, targeting both novice Wikipedians being onboarded into the community and its norms, and university students learning about citation. The project aimed to support verifiability processes – and the WikiCred mission – by providing learning support for the addition of citations to Wikipedia. The CiteLearn game situated students in the context of Wikipedia editing and gets them to think about where to place citations, for which purposes and through which conventions, which contributes to their development of information verifiability and fact-checking skills more broadly. The tool was designed with two core tasks, which reflected a model that was already in use in the teaching practice of the authors. In both tasks, users work with text via a simple text editor per Figure 1.

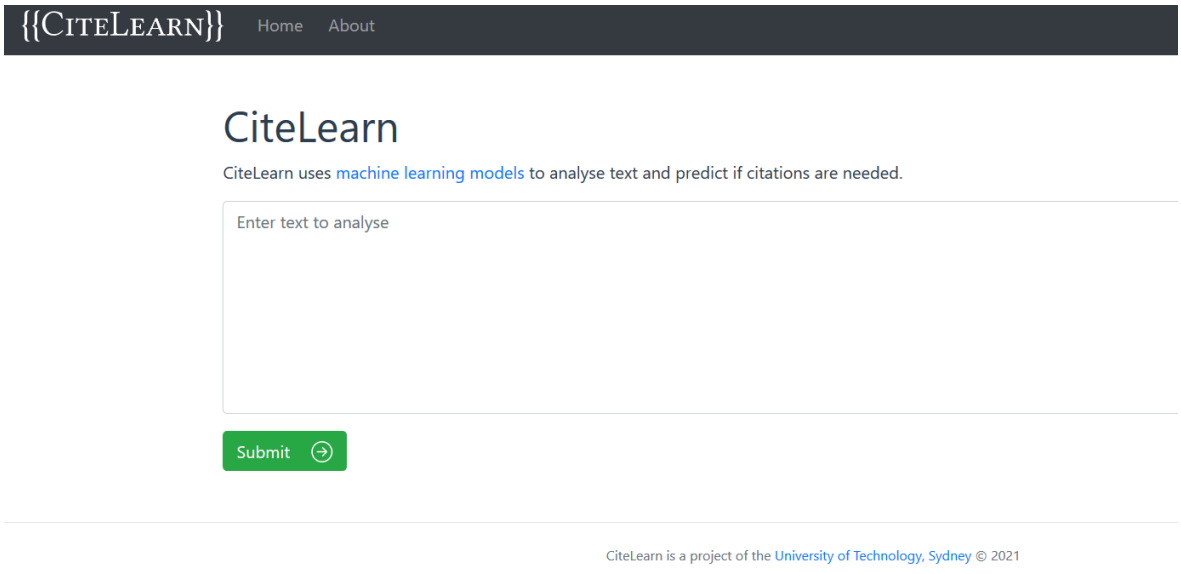
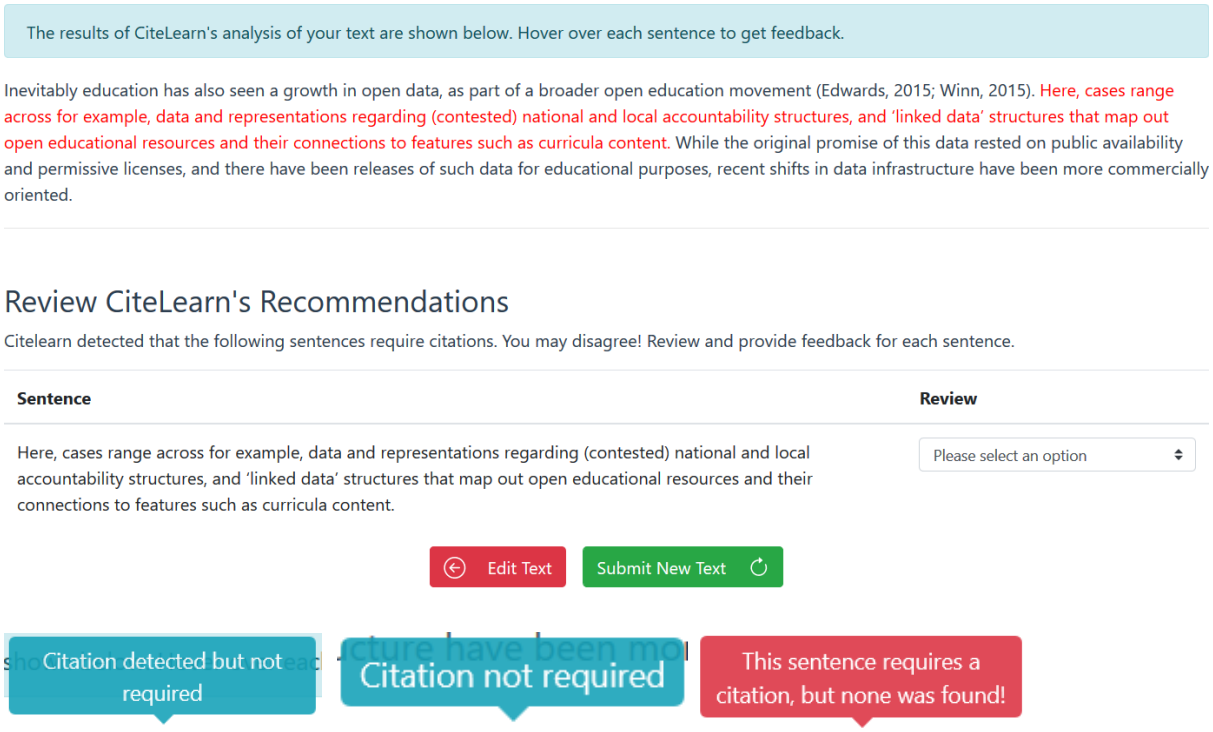


Figure 1 CiteLearn text editor

Figure 1 Alt-text: A screenshot of the CiteLearn tool with key features including a user text-entry box, and a submit button.

In round 1, users locate where citations should be placed in articles that are provided to them. This allows discussion of appropriate citation based on prior knowledge about the texts and where citations are required in it. In the main task (round 2), users must write their own articles, including citations. In this latter round, predictive models trained on the Citation Need dataset were used to identify sentences that require a citation and provide automated feedback to students, with the overall aim of developing their citation practices. As indicated in Figure 2, the tool provided multi-level feedback on the sentences, labelling sentences to indicate where (1) a citation is required but not given, (2) a citation is not required but detected, and (3) where a citation is not required and not given. Users review this feedback, labelling each sentence with their decision or implied action for the sentence. Users may also re-edit the text to receive new feedback.



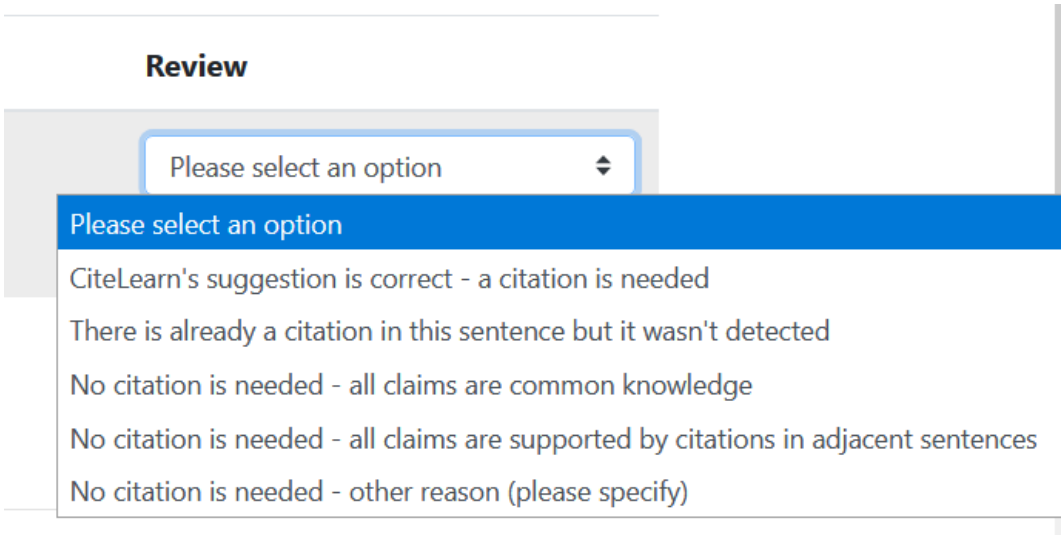


Figure 2 CiteLearn feedback model, top pane shows the feedback provided under the text editor; second row shows the 'hover over' text provided when users highlight an individual sentence; bottom row shows the "review" options that indicate the users review of the citation feedback.

Figure 2 Alt Text: A sample of CiteLearn analysis, with a sample paragraph text including one sentence highlighted in red to indicate a citation is missing. Users can hover over the sentences to see detail: "citation detected but not required", "citation not required", "this sentence requires a citation, but none was found!" The Recommendations panel displays the sentence, with a 'review' dropdown list with options: "CiteLearn's suggestion is correct - a citation is needed; There is already a citation in this sentence but it wasn't detected; No citation is needed - all the claims are common knowledge; No citation is needed - all claims are supported by citations in adjacent sentences; No citation is needed - other reason (please specify). Underneath this users can select to "edit text" or "submit new text".

The tool is thus intended to use existing data infrastructure (in the form of the citation needed dataset) to provide feedback on citation use, with the aim to support conventions regarding verifiability in both Wikipedia and scholarly academic writing genres.

Norms and Conventions: Citation in Wikipedia and Scholarly Writing

Wikipedia as verifiability infrastructure

Wikipedia has existed since 2001, currently hosting ~6.5 million English articles, along with other language versions, and other media on its various sister projects. Wikipedia's content is written and curated by volunteer editors, who have created a range of technical infrastructure including bots and structured templates, alongside norms and conventions, that govern practices on the platform (Ford, 2015, 2022). As this work has made clear, Wikipedia should thus be thought of as combining a technological infrastructure in mutual interaction with human activity, both shaping and shaped by the norms and conventions of the community. At the core of the Wikipedia community's content polices are three central principles ("Wikipedia," 2021c): neutral point of view (NPOV), no original research (NOR), and verifiability. These three all motivate inclusion of, and centre the importance of, external citations to display independently held views on issues in a fair proportionate way, to ensure this is grounded in independent sources rather than editorial opinion, and that this can be verified through attribution to these reliable sources. In this way, verifiability acts both as a central policy, and a statement of values that "effectively defines the communicative acts that are possible within Wikipedia's sociotechnical system." (Ford & Wajcman, 2017, p. 518).

A key aspect of citation practice is the tagging of *claims* within articles using the flag "citation needed" (often stylised on and off Wikipedia as {{Citation needed}}). At the time of writing, 444,518 articles have at least one such flag, and this is likely to understate both the number of articles in need of citations (because some articles have not been reviewed, and others are flagged with parallel templates), and the number of such flags within articles.

Learning to Cite at University

Citation practices are also a central feature of the genre of scholarly academic writing. When university students learn about citation practices, they are effectively being socialised into a particular professional genre with its own institutional contexts and norms (Lea & Street, 2006). A key component of these social norms in academic writing involves a broad set of academic integrity practices of which citation is a key part. In this context, both appropriate paraphrasing, and high quality synthesis writing, are challenging skills to learn but are important for accurate citations and referencing (Keck, 2006). However, we often focus on citation only from a plagiarism perspective, rather than their driving purposes for supporting and contextualising the claims being made, grounding claims in warrants (McGowan, 2005). Thus, understanding the need for citations and learning to attribute to appropriate sources can lead to improved writing across academic and external settings.

Despite the importance of citation in both Wikipedia and scholarly contexts, there are few learning tools to develop citation competencies. Previous game-based citation tools rely on multiple-choice-quiz (MCQ) assessments largely focusing on issues of plagiarism and academic integrity, and do not directly assess the authentic practices of writing and citing (Bradley, 2015). Nevertheless, evaluation instruments do exist to explore students' understanding of source use, competence, and citation motivation, to support such learning (Childers & Bruton, 2015; Ma & Qin, 2017).

Focal aspects: Citations and Learning Tools as Objects at the Boundary

We use the development of the CiteLearn tool – particularly the different models and algorithms employed to analyse when citations are required – to illuminate the nature of citations in both scholarly and Wikipedia writing. While we decided that the predictions generated by CiteLearn with respect to when citations were needed were not yet passing a face validity test (i.e., it was not clear it would provide useful feedback to students on where

citations were needed), that is not our focus here. In this discussion of the focal aspects of the case, we use "citation" as a boundary object, informing our understanding of citation in Wikipedia and scholarly learning environments. This view of "citation" as a boundary object in scholarly and encyclopaedic writing, makes visible the ways citation is grounded in norms and conventions of the two sites, through the Citelearn tool. The Citelearn tool thus acts as an algorithm game, a lens onto our feature of interest (citation). These considerations are not intended as a criticism either of work on citation in Wikipedia, or of the use of data from such environments to inform others (i.e., a misalignment of pedagogy or technical scope). Rather, we point out how an analysis of citation in the CiteLearn tool illuminates citation practices, and the potential for data infrastructures to obfuscate these practices in adapting and adopting data from one place and using it in another.

{{Citation needed}} as infrastructure and as practice

Across both Wikipedia and scholarly writing, 'citation' can be thought of as an artefact in the text that fulfils a particular social purpose for the genre, a data trace that signifies a particular feature to be used in manual and automated processes, and as the practices underpinning this. Just as Wikipedians tag articles {{citation needed}}, academics in peer review and assignment marking ask for references to be provided, indeed, sometimes adopting the "citation needed" tag as an annotation. It is from this, and the need for students to learn these practices, that the tool emerges.

Wikipedia practices: The Wikipedia context imbues {{citation needed}} with a particular status, clearly flagging the lack of citation on a factive claim to users. While {{citation needed}} is perhaps the most commonly known tag, and is the most common maintenance tag, there are others. When a user tags something they can add a reason, and the tag automatically links to a page that provides information on the meaning of the tag and how to resolve it. Other tags include: "clarify" (for sentences that may be confusing, used in ~46k

articles); “examples needed” (for sentences that need examples, used in ~700 articles); “more citations needed” (for whole articles in which more citations are needed, used in ~361k articles); and “unreferenced section” (for sections within articles, which contain no references, used in ~33k pages).

Alongside these, a set of discussions around the claims and their warrants (or verifiability) exist, with artefacts emerging from this practice including a list of commonly used sources and their quality – for Wikipedia purposes – in relation to particular topics (Lewoniewski et al., 2020; “Wikipedia,” 2021d). Such lists reflect (1) the range of actors engaging in editing, and drawing on sources – reliable and unreliable – to add citations; and (2) the range of appropriate sources drawn on in Wikipedia, including but not limited to scholarly academic literature. That is, Wikipedia has a rich set of practices and infrastructure to classify claims made in articles, from poor sourcing, to lack of clarity in the writing, ambiguity regarding which claim the source is a warrant for, and so on.

Scholarly practices: Writing is central to the work of universities (Hyland, 2013), with citation playing a central role in the provision of warrants for claims. However, there are a set of discontinuities around the particular character or purpose of citation in scholarly writing, and the kinds of sources drawn on for these purposes. Citation in Wikipedia is about verification that a view (external to Wikipedia) is held. Citation in scholarly writing is about provision of warrants for claims, often drawing together concepts and issues in ways that are novel, to make the argument (Hyland, 2013). These scholarly citations are, thus, typically external, scholarly, peer reviewed sources, or original research data (broadly construed), to forward a particular argument. Citations, then, fulfil a role in providing credit for ideas, and in tallying the frequency of individual papers (citation counts) and authors being cited by others (e.g. the H-index) (Hyland, 2013; West et al., 2017; Zhu et al., 2015).

Citations are also a data infrastructure, particularly in contexts where citation styles that apply 'digital object identifiers' (DOIs) are used. This represents "a technical and social infrastructure for the registration and use of persistent interoperable identifiers" (doi, 2021). They also act as infrastructure for tools such as Turnitin, providing 'feedback reports' to students and academics regarding then inclusion of uncited material, and features such as quotation marks acting to exclude text from these algorithms (Turnitin, n.d.-b, n.d.-a). The technologies of citation in scholarly contexts overlap with those in Wikipedia, with citation standards (APA, Harvard, etc.) and tools (Zotero, Mendeley, etc.) used in both, alongside standards for source reliability (as noted above). Ironically, although students may not believe Wikipedia is a reliable source, they may have weaker definitions of reliability than those used on Wikipedia (MacIntyre, 2020).

Citations situated in genres

The purported practices of citation in both Wikipedia and scholarly writing are situated in wider genres of text. In both contexts, model practices of citation are encoded in documents and resources such as teaching and learning resources, guidelines, and policies. However, these tend to be generalised, and typically do not reflect disciplinary differences in citation practices.

Wikipedia genres: Wikipedia articles themselves belong to the encyclopaedic genre, drawing on reliable sources for verifiability, presenting no original research and a neutral point of view. These features alone are distinctive insofar as much writing outside Wikipedia makes a particular argument, and draws out implications beyond those directly referenced in the source material. This results in a genre of text that is often densely populated with citations, because each claim should have verification and the content should be driven by these secondary sources limiting any additional text. This has implications for citation data in

that Wikipedia articles, especially those that meet quality criteria, are *unbalanced*; i.e., it is more likely that a sentence needs a citation (and either has one, or is labelled {{citation needed}}) than not, and this means training models on this data is presented with a challenge, because there are simply fewer sentences to learn from that correctly do not have citations. This is compounded by the fact that – as in scholarly writing – citations for a claim are often found in the surrounding linked sentences, and indeed conversely, that often sentences might have multiple claims in them.

Within this, too, Wikipedia has sub-genres both within articles – for example, lead sections do not contain citations, and are a summary of the article – and across them. Across articles differences include articles that are deemed to meet particular quality standards, disciplinary differences (e.g. historic events vs scientific articles), and topics including biographies and current event pieces In analysing the data for this work, we observed that many of the articles that were rated high quality (i.e., they have been reviewed, and thus we would expect that sentences not containing citations are appropriate) were also following particular conventions described above, such as using a single citation over multiple sentences, or a full paragraph.

Scholarly genres: Citations also fulfil a slightly different role in scholarly writing insofar as while in Wikipedia they are provided for the verification of claims, scholarly writing also includes authorial discussion, or – intuitively enough – what Wikipedians might refer to as 'original research'. While methodological critique and weighing up of alternatives may occur in Wikipedia, typically this is in the context of external sources having engaged in such critique, and that being reported. That is, original research and taking a point of view, grounded in warrants and support that contextualises claims, is core to scholarship, from faculty to students (Hyland, 2013). Citing is done across a range of types of scholarly writing, from assignments, magazine articles, peer review journals articles, page constrained grant

applications, and so on. Each of these has distinct norms and practices, that intersect with those of the discipline they are situated in.

Hidden technical infrastructure: From Wikipedia to Scholarly structures

Alongside issues of genre and practice, the boundary object of citations, can be analysed through our design case with respect to the technical infrastructure of both contexts.

Wikipedia structures: The genre and practices of Wikipedia are bound up with the technical infrastructure in a number of ways that ground the model development. The tool is grounded in a machine learning model to predict the likelihood a sentence in Wikipedia requires a citation. The model embeds aspects of Wikipedia's structures and norms both through its use of a corpus of exemplar 'Featured Articles' together with the particular approach used to translate these articles into features which train the model to predict when citations are required.

As noted above, two features in the training data demonstrate this point:

1. Claims in Wikipedia should be externally verifiable, leading to a far greater number of sentences with citations, than without. This is particularly true in the 'high quality' articles the model was trained on, and if not accounted for can result in 'unbalanced' data, i.e. data in which one class (sentences with citations) is more common. Unbalanced datasets are an issue for machine learning approaches as the algorithms do not have enough sample data to learn from, contributing to incorrect predictions on unseen data.
2. Genre may be related to structural features such as citations in unexpected ways. A surface analysis of sentences that lack citations suggests that many of these are from particular genres – often history topics – where a single source is used to ground a large portion of the text, and thus although there is an overarching citation for the claims made, it is not considered proper citation as per Wikipedia editing norms.

Thus, the annotated data on whether a sentence requires a citation derived from Wikipedia is likely to be applicable for specific purposes rather than being a universal definition for claims.

Models turn input data into 'features' on which they operate. For example, 'number of characters', 'the presence of a citation', and so on. In engaging in this feature engineering, we turn input data into an abstraction, that can change aspects of the representation. To give an example, a common problem across research is the appropriate unit of analysis or segmentation. For example, detecting sentence boundaries can be complex when non-standard sentence breaks are used, lists given, and long quotes (that encompass full stops) are provided. To derive features to be used as inputs to the model, we sampled an even number of sections that contained, and did not contain citations from high quality articles, with the assumption that any negative cases correctly omit citations. An even number of each were selected to counteract the imbalance in the data. For claims without a citation, the whole paragraph was programmatically checked for citations, and where none was provided, these paragraphs were used as negative cases. Here, we can see that a relatively minor decision – how to split sentences – may result in changes in a model, with the average text length of positive sentences at 40 words, and negative at 96. Given the genre issues noted above, choosing to use sentences, looking at paragraph breaks and citation density within paragraphs, or a range of other approaches are likely to impact modelling.

Model development also involves making decisions about how to develop a numerical representation of texts which can be used in machine learning. The original Citation Need model generated a word embedding using Word2Vec, in which each individual word maps to a single vector representation. More recently, BERT-based models have emerged which are able to both incorporate the contextual use of words while also having the capacity to go beyond the original vocabulary used to train the data. The BERT models are themselves pre-

trained on Wikipedia data (together with the BooksCorpus dataset), but can also incorporate models which have been fine-tuned to support specific tasks such as sentiment analysis, and question-answering. Rather than simply a decision which gives different model performance and 'accuracy', the choice of how a model generates word embeddings implicitly embeds particular structures – which impact both the generalisability of the model (e.g. to learn new words from new domains) and which use particular domains as the basis for generalising how text is understood (i.e. by pre-training models on a corpus of Wikipedia data).

Finally, model development also involves decisions regarding the outputs. Binary classifiers, including the one used here, assign a True or False, or "in the class: should have a citation" vs "*not* in the class: should have a citation", to input sentences using probabilistic models. However, how we interpret these outputs is ambiguous, including issues around the appropriate threshold at which to assign a True, if a lower score indicates greater certainty that a sentence should not have a citation, and whether the probabilities can be used to provide useful information in place of the binary class output (whether or not a citation is required). As noted above, these decisions have perfectly reasonable rationales, and so our aim here is not to critique the work, but rather to show how such decisions interact with the social context of the data, the structure of the data itself, and the translation of that data for other purposes in which it is represented and re-represented.

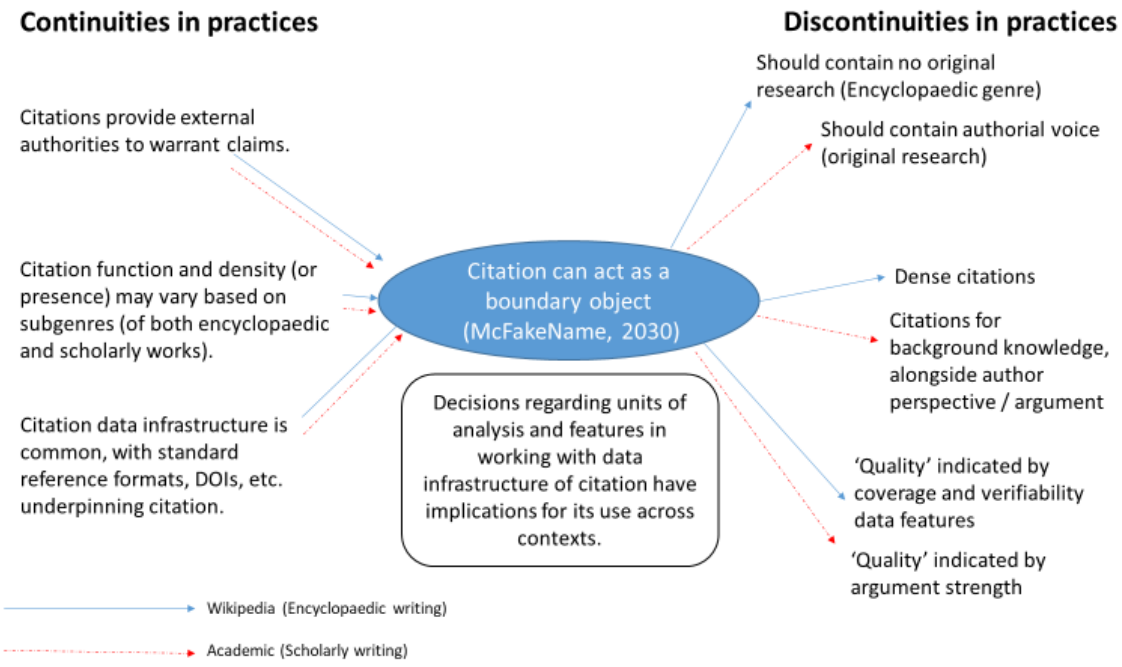
Scholarly structures: What, then, the investigation of citation as a boundary object indicates is nuanced differences in citation practices and norms, through the data of citation, across Wikipedia and learning contexts. Scholarly writing has a focus on warrants for claims, whether those warrants are from external sources or original data. These are often in the context of original argument, supported and contextualised by this evidence. That is, verifiability of individual claims by external citation is not key, with a focus instead on following the norms of argumentation for the discipline to put forward an original point of

view, with evaluation of this writing focusing on the overall structure rather than individual sentences.

Some language in this writing follows particularly scholarly patterns for rhetorical moves, such as those that may be familiar to academic readers from the Manchester Academic Phrasebank (Morley, n.d.). These provide one source for alternative model development in searching for phrases such as "As [authors] note", that indicate a citation is required[1]. However, notably the analysis of citations given these features of citation in scholarly writing is likely to be significantly more complex, presenting challenges for development of citation infrastructure from scholarly texts (as indeed anyone who has looked at plagiarism checking reports will know).

Key Implications

Across these focal aspects, some key implications can be highlighted as described below, and summarised in Figure 3.



[1] We note the original work did also test a phrase-based approach, but that these and use of other indicators such as quotation marks were not included in the final models because for the particular data they did not improve performance.

Figure 3 – Key Focal Points in Citation as Infrastructure: Material Representation; Genre Positioning Object; and Data Unit

Figure 3 Alt-Text. A diagram summarising the preceding section. On the left ‘continuities in practices’ are indicated; on the right ‘discontinuities in practices’, in the centre a sample citation – the boundary object – is displayed.

Implication of Citation as Infrastructure: Continuities and Discontinuities in Material Representation – citation as a written *artefact* or infrastructure has been used as a signal of continuities in citation as a *practice* across scholarly and Wikipedia writing. There are continuities in these infrastructures and practices across sites insofar as citation representations act as an indicator of authoritative external sourcing, and shared data infrastructure around common citation standards and DOIs (Carver et al., 2012). Through analysis of this data and its use we can identify key discontinuities. These include the nature of the claims being supported (verification of a non-original view in Wikipedia vs warrants in original research in scholarly practice), the sources themselves (secondary consensus pieces in Wikipedia vs peer review sources in scholarly practice), and their purpose (verifiability and article credibility in Wikipedia vs warrants alongside credit for claims to be used in metrics such as the h-index in scholarly practice). Although the material of a citation is similar (i.e., a text representation), across the two socio-technical systems it fulfils different purposes. This focal aspect illustrates described the analysis of data features as material artefacts with corresponding practices, such as citations. Such analysis holds promise for critical studies that seek to understand data journeys through and across sites (Bates et al., 2016; Howard et al., 2021; Leonelli, 2020),and the ways materials are imbued with and obscured from new and original meanings.

Implications of Citation as Genre Positioning Boundary: Material Objects in Social Data Structures – Common across both sites, citations may serve a range of purposes and these purposes may vary with the genre and disciplinary conventions of any given piece

of writing. Capturing these features that represent externally imbued meanings in the given feature – a citation – is challenging. One expression of this, is that continuity in the material expression of citation may not indicate continuities in the *absence* of such features and their meaning across contexts and genres.

Implications of Citation as Data Unit: Making Visible Technical Infrastructures – finally we highlighted technical infrastructure (1) underpinning data (such as the way Wikipedia articles are categorised and tagged, and scholarly practices around teaching particular forms of argumentative writing); (2) processing that data, including through decisions regarding features to focus on and units of analysis; and (3) re-present that data, including in decisions regarding what categories to assign to different types of citation, or how to present this information to users. Again, our aim here is to highlight how the use of our boundary object (citation) and analysis of technical infrastructure provide a lens onto the practices and materials of citation across contexts.

Conclusion

Open data sources, such as Wikipedia, match the largest datasets produced by commercial vendors, while making this data freely available for secondary uses (§Wikipedia and Data Infrastructures). This resource offers significant potential in learning contexts (§Wikipedia and education). The increasing datafication of education – and growing presence of collaborative learning tools such as wikis – lends itself to the adoption of this data for understanding and supporting learning. However, in abstracting models from their original contexts we must maintain a critical eye on how these models are used, and what differences there are between those original contexts and any given learning environment. Features that matter to educators, such as use of scholarly sources, hold similarities to features outside formal learning (§Norms and Conventions: Citation in Wikipedia and Scholarly Writing). However, these features – translated into data structures – cannot be transferred without

consideration of their local context. This very consideration provides a helpful lens onto open data practices.

This paper demonstrates this need, using a critical design reflection of a learning game, 'CiteLearn', and taking citation as a boundary object that helps us understand the continuities and discontinuities between Wikipedia and scholarly citation practices. We unpack this through our case context, the broad norms and conventions each site (Wikipedia and scholarly contexts) work under, and the ways these interact across our focal points. These focal points show how the 'data' of {{citation needed}} while representing a form of infrastructure, is bound up with practices and norms in both the encyclopedia and scholarly writing.

This characteristic of citation interacts with features of genre that both share similarities and distinctions across context. The explicit citation infrastructure – our first focal point – is also bound up with implicit concerns particularly in the translation of text into model, in which decisions around the units of analysis (such as sentence division) have significant implications, and where scholarly practices change the interpretation of features such as citation absences or use of key phrases. These points are nuanced, and complex, representing an interaction between data infrastructures, practices and norms, and genres across contexts.

While the paper provides a single case study, the intent is to provide an approach and illustration in order to foster further explication of such cases and to provide an important lens onto learning data, its representation and re-representation. Simply, design case analysis provides a method for analysis of the *materials* in a design-context (or citation game), the imbued meaning of those materials in practice (e.g., genre), and how data as a particular form of material represents meanings across contexts.

Acknowledgement

We would like to acknowledge funding from the WikiCred group, and the helpful discussions and feedback from both them and the researchers on the original Citation Need model: Miriam Redi (Wikimedia Foundation Senior Research Scientist) and Aiko Chou (Wikimedia Foundation Research intern).

We have also benefited from useful discussion with colleagues, including Kalervo Gulson and Teresa Swist on related projects. Our thanks to Kal particularly, who offered feedback on an earlier version of the manuscript.

The tool is implemented in a pilot form at https://citelearn.wmcloud.org/ documented and shared under open license at https://github.com/thelondonsimon/citelearn with a separate repository available for the models themselves https://github.com/thelondonsimon/citelearn-model. The pilot tool is hosted on Wikimedia Cloud VPS, documented at https://wikitech.wikimedia.org/wiki/Portal:Cloud_VPS

Two other significant points: The modelserver component of CiteLearn which gets predictions from pretrained models can be configured to use either the original CitationNeeded model (TensorFlow 1) or one of the BERT trained models tentatively developed using TensorFlow 2. Because of the different Python package dependencies, having the capacity to readily switch between models would require a more substantial refactoring. The model repository includes an alternative approach to developing a training set compared to the published data set for the Citation Needed Paper. This new data set includes contextual details of the sentence in terms of whether the preceding/subsequent sentences and containing paragraphs have citations. This extra feature is not included in the models at this stage.

References

Akkerman, S., & Bakker, A. (2011). Boundary Crossing and Boundary Objects. *Review of Educational Research*, *81*(2), 132–169. https://doi.org/10.3102/0034654311404435

Atenas, J., & Havemann, L. (2019). *Open Data & Education*.

Bardzell, J., Bardzell, S., & Stolterman, E. (2014). Reading critical designs: Supporting reasoned interpretations of critical design. In M. Jones, P. Palanque, A. Schmidt, & T. Grossman (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1951–1960). ACM. https://doi.org/10.1145/2556288.2557137

Bates, J., Lin, Y.-W., & Goodale, P. (2016). Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society*, *3*(2), 2053951716654502. https://doi.org/10.1177/2053951716654502

Bowker, G. C., & Star, L. S. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press.

Bradley, E. G. (2015). Using Computer Simulations and Games to Prevent Student Plagiarism. *Journal of Educational Technology Systems*, *44*(2), 240–252. https://doi.org/10.1177/0047239515617653

Brailas, A., Koskinas, K., Dafermos, M., & Alexias, G. (2015). Wikipedia in Education: Acculturation and learning in virtual communities. *Learning, Culture and Social Interaction*, *7*, 59–70. https://doi.org/10.1016/j.lcsi.2015.07.002

Carroll, J. M., & Rosson, M. B. (1992). Getting Around the Task-artifact Cycle: How to Make Claims and Design by Scenario. *ACM Transactions on Information Systems (TOIS)*, *10*(2), 181–212. https://doi.org/10.1145/146802.146834

Carver, B. W., Davis, R., Kelley, R. T., Obar, J. A., & Davis, L. L. (2012). Assigning Students to Edit Wikipedia: Four Case Studies. *E-Learning and Digital Media*, *9*(3), 273–283. https://doi.org/10.2304/elea.2012.9.3.273

Chen, B., Håklev, S., & Rosé, C. P. (2021). Collaborative Learning at Scale. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 163–181). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_9

Childers, D., & Bruton, S. (2015). "Should It Be Considered Plagiarism?" Student Perceptions of Complex Citation Issues. *Journal of Academic Ethics*, *14*(1), 1–17. https://doi.org/10.1007/s10805-015-9250-6

Chou, A.-J., Gonçalves, G., Walton, S., & Redi, M. (2020). *Citation Detective: A Public Dataset to Improve and Quantify Wikipedia Citation Quality at Scale*.

d'Aquin, M. (2016). On the use of Linked Open Data in education: Current and future practices. In *Open data for education* (pp. 3–15). Springer.

doi. (2021). *The DOI® System*. https://www.doi.org/

Edwards, R. (2015). Knowledge infrastructures and the inscrutability of openness in education. *Learning, Media and Technology*, *40*(3), 251–264. https://doi.org/10.1080/17439884.2015.1006131

Featured article criteria. (2021a). In *Wikipedia*. https://en.wikipedia.org/wiki/Wikipedia:Featured_article_criteria

Ford, H. (2015). Infoboxes and cleanup tags: Artifacts of Wikipedia newsmaking. *Journalism*, *16*(1), 79–98. https://doi.org/10.1177/1464884914545739

Ford, H. (2022). *Writing the Revolution: Wikipedia & the Survival of Facts in the Digital Age.* MIT Press.

Ford, H., & Wajcman, J. (2017). 'Anyone can edit', not everyone does: Wikipedia's infrastructure and the gender gap. *Social Studies of Science*, *47*(4), 511–527. https://doi.org/10.1177/0306312717692172

Graham, M., Straumann, R. K., & Hogan, B. (2015). Digital divisions of labor and informational magnetism: Mapping participation in Wikipedia. *Annals of the Association of American Geographers*, *105*(6), 1158–1178. https://doi.org/10.1080/00045608.2015.1072791

Gulson, K., Benn, C., Kitto, K., Knight, S., & Swist, T. (2021). Algorithms can decide your marks, your work prospects and your financial security. How do you know they're fair? *The Conversation*. http://theconversation.com/algorithms-can-decide-your-

marks-your-work-prospects-and-your-financial-security-how-do-you-know-theyre-fair-171590

Gulson, K., Swist, T., Knight, S., & Kitto, K. (2021). *Technical democracy, fairness and the UK exam algorithm: Making a 'design Thing' to explore bias in automated grading systems* [Conference presentation]. AARE. https://www.aare.edu.au/events/2021-conference/

Hao, K., & Stray, J. (2019). Can you make AI fairer than a judge? Play our courtroom algorithm game. *MIT Technology Review*.

Haren, S., Romero, S., Hottinger, P., & DeJonghe, E. (2020). Using Wikipedia for project-based learning in information literacy instruction. *Journal of New Librarianship*, *5*(1), Article 1. https://doi.org/10.33011/newlibs/9/13

Heffernan, N. T., Ostrow, K. S., Kelly, K., Selent, D., Van Inwegen, E. G., Xiong, X., & Williams, J. J. (2016). The Future of Adaptive Learning: Does the Crowd Hold the Key? *International Journal of Artificial Intelligence in Education*, *26*(2), 615–644. https://doi.org/10.1007/s40593-016-0094-z

Howard, S., Gašević, D., Knight, S., Swist, T., Bartimote, K., Gulson, K., Apps, T., Peloche, J., Hutchison, N., & Selwyn, N. (2021). *Data journeys and lives: Making visible the datafication of educational policy, teaching and learning* [Conference presentation]. AARE. https://www.aare.edu.au/events/2021-conference/

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching*, *46*(01), 53–70. https://doi.org/10.1017/S0261444811000036

Hylén, D. J. (2006). *Open Educational Resources: Opportunities and Challenges*. 10.

Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, *15*(4), 261–278. https://doi.org/10.1016/j.jslw.2006.09.006

Knight, S., Shibani, A., & Buckingham Shum, S. (2023). A Reflective Design Case of Practical Ethics in Learning Analytics. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13323

Koch, B., Denton, E., Hanna, A., & Foster, J. G. (2021). *Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research* (arXiv:2112.01716). arXiv. https://doi.org/10.48550/arXiv.2112.01716

Koesten, L., Vougiouklis, P., Simperl, E., & Groth, P. (2020). Dataset Reuse: Toward Translating Principles to Practice. *Patterns*, *1*(8), 100136. https://doi.org/10.1016/j.patter.2020.100136

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data*, *4*(1), 170171. https://doi.org/10.1038/sdata.2017.171

Lea, M. R., & Street, B. V. (2006). The" academic literacies" model: Theory and applications. *Theory into Practice*, *45*(4), 368–377. https://doi.org/10.1207/s15430421tip4504_11

Leonelli, S. (2020). Learning from data journeys. In *Data journeys in the sciences* (pp. 1–24). Springer, Cham.

Lewoniewski, W., Węcel, K., & Abramowicz, W. (2020). Modeling Popularity and Reliability of Sources in Multilingual Wikipedia. *Information*, *11*(5), Article 5. https://doi.org/10.3390/info11050263

Li, A., Yao, Z., Yang, D., Kulkarni, C., Farzan, R., & Kraut, R. E. (2020). Successful Online Socialization: Lessons from the Wikipedia Education Program. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 050:1-050:24. https://doi.org/10.1145/3392857

Li, Y., Lovitt, C. F., McNeil, A., & Shuyler, K. (2016). Improving Information Literacy through Wikipedia Editing in the Chemistry Classroom: Lessons Learned. In

Integrating Information Literacy into the Chemistry Curriculum (Vol. 1232, pp. 247–264). American Chemical Society. https://doi.org/10.1021/bk-2016-1232.ch013

Ma, R., & Qin, X. (2017). Individual factors influencing citation competence in L2 academic writing. *Journal of Quantitative Linguistics*, *24*(2–3), 213–240. https://doi.org/10.1080/09296174.2016.1265793

MacIntyre, R. (2020). Citation machines: The use of evidentials in the academic writing of Japanese university students. *Changing Perspectives in Academic Genres*, 59.

McGowan, U. (2005). *Plagiarism detection and prevention: Are we putting the cart before the horse*. https://digital.library.adelaide.edu.au/dspace/handle/2440/28598

Mehdi, M., Okoli, C., Mesgari, M., Nielsen, F. Å., & Lanamäki, A. (2017). Excavating the mother lode of human-generated text: A systematic review of research that uses the wikipedia corpus. *Information Processing & Management*, *53*(2), 505–529. https://doi.org/10.1016/j.ipm.2016.07.003

Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. Å., & Lanamäki, A. (2015). "The sum of all human knowledge": A systematic review of scholarly research on the content of Wikipedia. *Journal of the Association for Information Science and Technology*, *66*(2), 219–245. https://doi.org/10.1002/asi.23172

Morley, J. (n.d.). *Academic Phrasebank*. Academic Phrasebank. Retrieved January 11, 2022, from https://www.phrasebank.manchester.ac.uk/

Oliver, J. T. (2015). One-shot Wikipedia: An edit-sprint toward information literacy. *Reference Services Review*, *43*(1), 81–97. https://doi.org/10.1108/RSR-10-2014-0043

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Prpic, J., Melton, J., Taeihagh, A., & Anderson, T. (2015). MOOCs and Crowdsourcing: Massive Courses and Massive Resources. *First Monday*, *20*(12). https://doi.org/1584023090

Redi, M., Fetahu, B., Morgan, J., & Taraborelli, D. (2019). Citation Needed: A Taxonomy and Algorithmic Assessment of Wikipedia's Verifiability. *The World Wide Web Conference*, 1567–1578. https://doi.org/10.1145/3308558.3313618

Redi, M., Gerlach, M., Johnson, I., Morgan, J., & Zia, L. (2021). A Taxonomy of Knowledge Gaps for Wikimedia Projects (Second Draft). *ArXiv:2008.12314 [Cs]*. http://arxiv.org/abs/2008.12314

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology,translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387–420. https://doi.org/10.1177/030631289019003001

Stray, J., & Hao, K. (n.d.). *Interactive Visualization of Fairness Tradeoffs*.

Tiwari, S., Al-Aswadi, F. N., & Gaurav, D. (2021). Recent trends in knowledge graphs: Theory and practice. *Soft Computing*, *25*(13), 8337–8355. https://doi.org/10.1007/s00500-021-05756-8

Turnitin. (n.d.-a). *Excluding quotes and bibliography*. Retrieved January 17, 2023, from https://web.archive.org/web/20230116182548/https://help.turnitin.com/feedback-studio/turnitin-website/student/the-similarity-report/excluding-quotes-and-bibliography.htm

Turnitin. (n.d.-b). *Interpreting the Similarity Report*. Retrieved January 17, 2023, from https://web.archive.org/web/20230116182434/https://help.turnitin.com/originality-check/turnitin-website/instructor/the-similarity-report/interpreting-the-similarity-report.htm

User:SuperHamster/CiteUnseen. (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=User:SuperHamster/CiteUnseen&oldid=1020991889

Van der Veer Martens, B. (2001). Do citation systems represent theories of truth? *Information Research*, *6*(2), 6–2.

Vetter, M. A., Sarraf, K. S., & Woods, E. (2020). Assessing the Art + feminism Edit-a-thon for Wikipedia literacy, learning outcomes, and critical thinking. *Interactive Learning Environments*, *0*(0), 1–13. https://doi.org/10.1080/10494820.2020.1805772

Voss, J. (2005). *Measuring Wikipedia*. http://eprints.rclis.org/6207/

West, R., Stenius, K., & Kettunen, T. (2017). Use and abuse of citations. *Addiction Science: A Guide for the Perplexed*, *191*. https://doi.org/10.5334/bbd.j

Wikipedia:Citation needed. (2021b). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Citation_needed&oldid=1047094854

Wikipedia:Core content policies. (2021c). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Core_content_policies&oldid=1029195166

Wikipedia:Reliable sources/Perennial sources. (2021d). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Reliable_sources/Perennial_sources&oldid=1058737261

Wikipedia:Verifiability. (2021e). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&oldid=1049369660

Williamson, B. (2015). Governing software: Networks, databases and algorithmic power in the digital governance of public education. *Learning, Media and Technology*, *40*(1), 83–105. https://doi.org/10.1080/17439884.2014.924527

Williamson, B. (2019). Policy networks, performance metrics and platform markets: Charting the expanding data infrastructure of higher education. *British Journal of Educational Technology*, *50*(6), 2794–2809. https://doi.org/10.1111/bjet.12849

Winn, J. (2015). Open education and the emancipation of academic labour. *Learning, Media and Technology*, *40*(3), 385–404. https://doi.org/10.1080/17439884.2015.1015546

Zheng, B., Niiya, M., & Warschauer, M. (2015). Wikis and collaborative learning in higher education. *Technology, Pedagogy and Education*, *24*(3), 357–374. https://doi.org/10.1080/1475939X.2014.948041

Zhu, X., Turney, P., Lemire, D., & Vellino, A. (2015). Measuring academic influence: Not all citations are equal. *Journal of the Association for Information Science and Technology*, *66*(2), 408–427. https://doi.org/10.1002/asi.23179