QLattes: An Open-Source Tool for Qualis Annotation and Visualization in the Lattes Platform

Nabor C. Mendonça1, Maria Andréia F. Rodrigues1, Lucas R. Mendonça2

1Programa de Pós-Graduação em Informática Aplicada (PPGIA)  
Universidade de Fortaleza (UNIFOR)  
Av. Washington Soares, 1321, Edson Queiroz  
60811-905 Fortaleza – CE

2Instituto Eldorado  
Av. Alan Turing, 275, Cidade Universitária  
13083-898 Campinas – SP

{nabor,mafr}@unifor.br, lucas.rodrigues@eldorado.org.br

Abstract. QLattes is an open-source browser extension for Chrome and Firefox that automatically classifies and annotates journal publications in CNPq’s Lattes Platform according to their Qualis categories. It also provides filtering, analysis, and visualization of Qualis-classified publication data extracted from Lattes CVs based on multiple metrics and perspectives. This paper presents a comprehensive description of QLattes, covering its user interface, architecture, implementation decisions, Qualis data sources, and performance. Additionally, the paper provides some preliminary data on the usage and impact of QLattes within the Brazilian academic community. The source code of QLattes is freely available at https://github.com/nabormendonca/qlattes.

1. Introduction

The Lattes Platform is a database of academic CVs created by the Brazilian National Council for Scientific and Technological Development (CNPq) in 1999 [CNPq 2022b]. The platform was developed to provide a uniform and comprehensive way for Brazilian researchers and students to showcase and share their academic and professional achievements, including publications, research projects, and awards. Over the years, the Lattes Platform has become a central component of the Brazilian research ecosystem, with nearly seven million registered CVs as of March 2023 [CNPq 2022a].

Although the Lattes Platform has been instrumental in facilitating the evaluation of Brazilian researchers and their institutions, e.g., [Dias et al. 2013, Perlin et al. 2017, Dias and Moita 2018, Cota et al. 2021], providing a unified repository of academic information was an important yet insufficient step towards a standardized way of assessing academic output. To that end, in 1998 the Coordination for the Improvement of Higher Education Personnel (CAPES), an agency of the Brazilian Ministry of Education, created Qualis, Brazil’s current official system for ranking and evaluating scientific publications [CAPES 2020].

Qualis’ main goal is to offer a more precise and transparent process for evaluating the quality of scientific publications in Brazil, as well as the scientific output of Brazilian graduate programs [Barradas Barata 2016]. The system classifies scientific publications
into several categories according to their academic impact and relevance. In its most recent update, introduced in 2020, the Qualis system classifies scientific publication vehicles according to four A-level (A1–A4) categories and four B-level (B1–B4) categories, in addition to a single C category [CAPES 2020]. The category of a specific publication vehicle is defined based on multiple objective metrics, such as the citation percentile computed by the Scopus database, the impact factor computed by Clarivate Analytics, and the h5 index computed by Google, among other metrics [CAPES 2023].

It is worth noting that Qualis has faced some criticism from Brazilian researchers since its inception, who argue that it may incentivize the publication of research in vehicles that prioritize quantity over quality and that are not necessarily aligned with the specific needs and interests of the Brazilian scientific community, e.g., [Rocha-e-Silva 2009, Fernandes and Manchinin 2019, Jaffé 2020]. Another criticism is that Qualis is often used to evaluate researchers individually, while CAPES’s official recommendation is that it should be used only for evaluating the scientific output of research programs [Barradas Barata 2016]. Notwithstanding its potential limitations and misuse, Qualis is widely recognized as another critical component of the Brazilian research ecosystem.

Given their widespread adoption by the Brazilian academic community, the Lattes Platform and Qualis are frequently used together by researchers and institutions, e.g., to identify the Qualis classification of journal and conference publications in Lattes CV pages. However, despite being commissioned by Brazil’s federal government, these systems were developed independently and are maintained by different government agencies, resulting in a cumbersome and predominantly manual cross-referencing process. This lack of integration between the two systems is particularly challenging for research program coordinators and administrative staff who typically have to analyze the Lattes CV of many researchers within a short time frame. This situation presents a compelling opportunity for independent developers and researchers to create automated tools that integrate the Lattes Platform with Qualis, directly benefiting their combined user bases.

This paper introduces QLattes, an open-source browser extension for Chrome and Firefox that automatically classifies and annotates the Qualis category of journal publications in the Lattes Platform.1 Moreover, QLattes facilitates the filtering, analysis, and visualization of Qualis-classified publication data extracted from Lattes CVs using multiple metrics (e.g., average, median, trend) and perspectives (e.g., tabular and graphical views). To our knowledge, QLattes is the first open-source tool to seamlessly integrate, filter, analyze, and graphically display publication data from both the Lattes Platform and the Qualis system. By doing so, QLattes can greatly accelerate the task of analyzing and collecting publication data from many Lattes CVs, thus providing a welcome productivity boost to research program coordinators and administrative staff. By releasing the source code for QLattes, we also aim to contribute to the research community with an open extension infrastructure, which could facilitate the exploration and development of novel Lattes-based publication analytics and visualization tools.

QLattes was launched in the Chrome Web Store on January 3, 2023, the same day it was first announced to the general public through the Brazilian Computer Society (SBC)

1QLattes’ source code is freely available at https://github.com/nabormendonca/qlattes.
mailing list. As of March 31, 2023, QLattes has approximately 10,500 weekly active users, according to usage data provided by Google, and this number continues to grow steadily. These statistics, along with other performance and user experience indicators presented later in this paper, suggest that QLattes is an effective tool that is already being utilized by a significant segment of the Brazilian academic community.

The rest of the paper is organized as follows. Section 2 gives an overview of the Chrome extension development platform. Section 3 describes QLattes. Section 4 offers some preliminary usage statistics and community feedback about the tool. Section 5 compares and contrasts QLattes with other related tools. Finally, Section 6 concludes the paper and outlines future developments.

2. Chrome Extension Platform

Chrome extensions are small software programs that are designed to add or modify the functionality of Chromium-based web browsers (e.g., Chrome, Edge, Opera, and Brave) [Google 2023]. They are written using standard web technologies like HTML, CSS, and JavaScript and can interact with web pages, browser tabs, and other Chrome APIs.

The development of Chrome extensions has evolved over the years. The most recent version of the Chrome extension platform, introduced in 2021, is manifest version 3 (v3). This version further tightened security measures, simplified the development process compared to previous versions, and is also supported by Mozilla Firefox browsers, making v3 the first cross-browser extension platform.2

Below we briefly describe the main components and resources required to develop Chrome extensions with manifest v3. The next section explains how we have built upon some of those components and resources to design and implement QLattes.

Manifest file This is the main configuration file for a Chrome extension. It defines the name, version, description, and permissions of the extension, as well as its content scripts, service workers, and web resources.

Content scripts Content scripts are JavaScript files that can be injected into web pages. They can modify a web page’s HTML structure and hence change its appearance and behavior. Content scripts are loaded when a user visits a web page the extension has permission to run on.

Service workers Service workers are JavaScript files that run in the background of the browser. They can be used to listen to browser events (e.g., navigating to a new page, removing a bookmark, and closing a tab), manage the extension state, and communicate with the other components via message-passing.

Chrome APIs Chrome provides a range of APIs that allow extensions to interact with the browser, web pages, and other resources. Some of the most commonly used APIs include the storage API and the message-passing API. The former allows extensions to persistently store data within the browser while the latter allows extension components to exchange messages asynchronously at run time.

HTML pages An extension can have multiple HTML pages. The two most common HTML pages are the popup page and the options page. The popup page is a small

2https://blog.mozilla.org/addons/2022/05/18/manifest-v3-in-firefox-recap-next-steps/
window that appears when the user clicks on the extension icon. The options page provides a way for users to customize an extension, such as choosing which sites the extension will run on. HTML pages can interact with service workers and the other pages via the storage and message-passing APIs.

**Web resources** An extension can include many types of files as in-browser web resources, such as images (e.g., icons and banners), fonts, data files, and third-party JavaScript libraries.

3. QLattes

In this section, we describe QLattes in terms of its user interface (UI), architecture, implementation decisions, Qualis data sources, and performance.

3.1. UI

QLattes’ UI comprises two web pages, as shown in Figure 1: an annotated Lattes CV page (Figure 1a), and a visualization page (Figure 1b). Both pages support multiple windows, enabling users to work with multiple CVs and visualizations simultaneously.
To enhance the original Lattes CV page, QLattes adds a new header with two links (items 1 and 2): one to the section containing the annotated journal publications, and another to open the tool’s visualization page. Additionally, each journal publication is annotated with the journal’s ISSN, Qualis classification, and a link to the Scopus portal (when the journal is indexed by Scopus), among other information (item 3).

The visualization page consists of four main sections (items 5–8): a navigation bar, an author information display and removal section, a control panel, and a visualization canvas. The navigation bar allows users to switch between the visualization page’s multiple tabs, including tabs for data visualization (the default tab), usage instructions, frequently asked questions, privacy policy, credits, and an inline user feedback form. The visualization tab contains UI elements to control and display the selected author’s publication data. The control panel includes options to select the view type (tabular, graphical, or top papers lists), the publication period, and whether to display publication statistics such as the average, mean, and trend of the author’s number of publications per year. Both the tabular and graphical views are updated automatically in response to changes in the control panel options.

3.2. Architecture

In addition to the two HTML pages that are part of its UI, QLattes’ architecture includes a popup page, a content script, and multiple web resources, as shown in Figure 2.

The popup page serves no useful content to the user and only functions as a means of opening the visualization page once the user clicks the extension’s icon. The content script is responsible for extracting, classifying, and annotating the Qualis categories of journal publications that are registered on the targeted Lattes CV page. To achieve this, the content script initially loads three data files into memory, which contain the three Qualis data sources distributed with QLattes (see Section 3.4). The script also saves the classified publication data that it extracts from the Lattes CV page in an in-browser repository via Chrome’s storage API.

Whenever there is a change in the data stored in the repository, the storage API

---

3 The user can also open the visualization page by clicking on the tool’s icon (item 4) in the browser’s toolbar.
notifies the visualization page. There are two occasions when this can happen: first, when the content script saves the publication data that it extracted from a new Lattes CV page into the repository, and, second, when the user removes an author’s publication data from the repository using a different visualization page window. In both instances, the visualization page updates the list of available authors for selection. Once the user chooses a particular author, the visualization page retrieves that author’s publication data from the repository and displays them in the visualization canvas based on the user’s chosen view type and period.

QLattes also contains other types of web resources that are not shown in Figure 2, such as image icons of various sizes and a third-party chart library (Chart.js) that is utilized to generate the graphical view of the visualization page.

3.3. Implementation Decisions

During the development of QLattes, we made some important implementation decisions that directly affected the tool’s scope and usage. In the following, we describe those decisions, explain the rationale behind them, and discuss their trade-offs.

The first decision was that QLattes should not call any external service or API during its execution. The rationale behind this decision was to guarantee the tool’s responsiveness and resilience to possible communication delays. To enable that decision, we preemptively downloaded and included as local web resources all the data QLattes needs to classify the Qualis categories of journal publications (see the next subsection). One consequence of that decision is that we will have to manually update QLattes to include new classification data every time a new version of Qualis becomes available. Although updating QLattes may require a non-trivial effort on the part of the tool’s developers, on the user’s side no further action is needed for an update to take effect. This happens because both Chromium and Firefox browsers keep track of the version history of all installed extensions and automatically update them when a new version becomes available in their respective extension stores.

The second decision was that QLattes should only allow filtering and visualization of publication data of a single Lattes author at a time. This decision was meant to simplify QLattes’ user interface and thus expedite the development and launch of its first (beta) version to the public. As QLattes’ user base grows, the need to aggregate and visualize publication data extracted from groups of authors may lead us to revisit this decision in future versions of the tool.

The third decision was that QLattes’ user interface should be implemented in plain HTML5 and should not rely on any external web development framework, such as JQuery or Reactive. With this decision, we avoided having to distribute QLattes with non-essential third-party libraries, which could further increase the tool’s size and memory footprint. As QLattes’s user interface features grow in size and complexity, we may revisit this decision if the higher-level user interface abstractions provided by existing web development frameworks turn out to compensate for their extra run-time cost.

The final decision was that QLattes should not (attempt to) classify the Qualis categories of conference and workshop publications. The rationale behind this decision is two-folded: (i) conference publications are not considered for the evaluation of research programs in many knowledge areas, and (ii) accurately identifying conferences
and workshops solely based on the information available on Lattes CV pages can be challenging. In contrast to scientific journals, which can be unambiguously identified based on their ISSN, conferences and workshops have no generally accepted unique IDs. This means that any naive attempt to uniquely identify conferences and workshops on Lattes CV pages is deemed to generate many false positives and false negatives. We believe having QLattes misclassifying Lattes CV publications would hurt the tool’s credibility, especially in knowledge areas where conference publications are highly regarded, such as Computer Science. Again, we may revisit this decision in the future as more accurate approaches for conference identification are developed.

3.4. Qualis Data Sources

QLattes includes three Qualis data sources to classify journal articles:

**CAPES data source** This data source was extracted from CAPES’ Sucupira Platform[^4] at the end of 2022 and includes the Qualis classification of 31,337 journals from all knowledge areas, as classified by CAPES in the 2017-2020 quadrennium.

**PUC-RS data source** This data source was extracted from the PUC-RS’s Computer Science Qualis web page[^5] in December 2022 and includes the Qualis classification of 2,190 journals (predominantly in Computer Science), as classified by PUC-RS researchers in November 2021.

**Scopus data source** This data source was extracted from the Scopus database[^6] in May 2021 and includes the citation percentiles of 25,990 journals from all knowledge areas, as computed by Scopus.

When QLattes annotates a Lattes CV page, it searches for a journal’s Qualis classification in each of the three data sources, one at a time, in the order listed above, based on the journal’s ISSN. If the journal’s ISSN is not found in the first data source, QLattes repeats the search in the second data source, and so on. If a journal’s ISSN is only found in the Scopus data source, QLattes applies the citation percentile distribution rules established by CAPES to identify the journal’s Qualis classification, e.g., A1 if the citation percentile is $\geq 87.5$, A2 if the citation percentile is $< 87.5$ and $\geq 75$, and so on [CAPES 2020]. If a journal’s ISSN is not found in any of the three data sources, QLattes annotates it as *Non-Classified*.

To ensure the accuracy and efficiency of the tool, we filtered out all redundant entries from each of the three data sources, keeping only the entries with the highest Qualis classification (in the case of the CAPES and PUC-RS data sources) or the highest citation percentile (in the case of the Scopus data source). This approach, consistent with the current journal classification policy adopted by CAPES, has significantly reduced the size of each data source compared to the original dataset extracted from the respective sites. However, since the three data sources were developed independently, there are still many common journal entries among them. In future versions of QLattes, we plan to eliminate all redundant journal entries across the three data sources, further reducing the tool’s storage and memory requirements.

[^6]: https://www.scopus.com/sources.uri
3.5. Performance

We conducted a preliminary evaluation of QLatte’s performance by annotating and visualizing the Qualis categories of all journal publications in the Lattes CV of a senior Brazilian computer science researcher. We selected this CV for its relatively high number of journal publications (over 150) spanning over a long (47-year) period. This extended publication period is particularly relevant to evaluate QLatte’s visualization performance since the tool displays publication data aggregated by year. Since most current Brazilian researchers are expected to have a shorter publication history, and many certainly are not as prolific, we believe this CV is likely to be at the high end of QLatte’s common use cases in terms of the number of classified journal publications. We ran the performance evaluation experiment ten times on a MacBook Air with macOS Big Sur 11.7, featuring a 1.6 GHz Intel Core i5 processor and 8 GB LPDDR3 memory. We selected this relatively low-end hardware configuration because we believe it is also representative of the devices QLatte’s users typically use. Finally, to measure QLatte’s performance, we computed the average execution time and standard deviation of QLatte’s content and visualization scripts across all runs.

On average, the content script took 1,894ms to annotate the target CV with a standard deviation of 367ms. As for the visualization script, it took an average of 110ms to render the CV’s graphical view and 6ms to render the tabular view, with standard deviations of 32ms and 2ms, respectively.\(^7\) These results demonstrate that QLatte can reliably classify and annotate over one hundred journal publications in just a few seconds. Interestingly, the Lattes CV page takes longer to load than QLatte takes to annotate its journal publications. Moreover, QLatte’s visualization script has negligible execution times, allowing it to render and update the CV’s graphical and tabular views in near real-time. These findings suggest that QLatte is an effective tool for annotating and visualizing the Qualis categories of reasonably sized Lattes CVs containing over one hundred journal publications. Conducting a more systematic evaluation of QLatte’s performance and scalability is outside the scope of this paper and will be addressed in our future work.

4. Usage Statistics and Community Feedback

Usage statistics for QLatte are displayed in Figure 3, which includes information on the tool’s daily installations (Figure 3a) and weekly active users (Figure 3b), based on usage data provided by Google’s Developer Dashboard.\(^8\) As of March 31, 2023, QLatte has been installed over 9,500 times. As expected, the number of daily installations peaked during the initial few weeks after the tool’s public announcement and has since stabilized at roughly 50–90 new installations per business day. However, the number of weekly active users has continued to grow steadily, with an average of 6,260 users per week since its launch and a recent peak of about 10,500 weekly active users during the most recent week of data collection (March 25–31, 2023).

QLatte’s feature of registering more weekly active users than total installations is likely due to the tool’s installation on devices shared by multiple users, a common occurrence in administrative work environments. Moreover, the tool’s current number of daily active users (roughly 7,500 per business day) remaining relatively close to its total

\(^7\)The full experimental results are available at https://tinyurl.com/yvytsf87
\(^8\)https://chrome.google.com/webstore/devconsole/
number of installations suggests that most users are utilizing the tool regularly. These figures bolster our confidence that QLattes is already providing valuable benefits to a significant and expanding subset of the Brazilian academic community.

To preserve user privacy, we have disabled all options for tracking user location upon deploying QLattes in the Chrome Web Store. As a result, we lack fine-grained data regarding the geographic location of QLattes’ users within Brazil. However, users may provide their institutional affiliation and geographic location, as well as other personal data, voluntarily by using the feedback form embedded in the tool or by contacting QLattes’ developers directly via email.

As of this writing, we have received over 70 feedback messages from QLattes users across all five geographic regions of Brazil, representing a diverse range of knowledge areas, from Computer Science to Law Studies, Archeology, and Medicine. Figure 4 displays a map of Brazil highlighting the locations of these users, as well as a selection of their most common quotes and impressions about the tool. As one can see, the feedback we have received on QLattes from those users has been overwhelmingly positive, with several users praising the tool’s usefulness and effectiveness, including some self-declared research program coordinators. The only “negative” feedback has been from users who initially encountered issues installing or running the tool, which were later solved and found to be due to misunderstandings or misconfigurations on the users’ part.

Despite receiving encouraging feedback from tens of QLattes’ users, we acknowledge that such a small sample (less than 1% of the tool’s weekly active users) only provides anecdotal evidence. Nevertheless, the voluntary nature of all the user feedback we have received implies that our sample is likely randomized, and thus, we believe it is suffi-
What a fantastic tool! People have been loving the tool. It's a very relevant tool that will help researchers and graduate program coordinators. Great initiative and very useful! It does in seconds what used to take weeks. The tool works really well! It greatly facilitated my work as a coordinator of a graduate program. This tool will make my job so much easier. Thank you very much for sharing it for free!!!

Figure 4. City locations of a sample of QLattes’ users in Brazil and a selection of their shared impressions about the tool

icient to provide an early but reliable indication that QLattes’ target audience is embracing it as a valuable and effective companion to the Lattes Platform.

5. Related Work

Various independent researchers have developed analytics tools to extract, analyze, and visualize academic data from the Lattes Platform over the years. These tools include scriptLattes [Mena-Chalco and Cesar Junior 2009], an open-source system that generates academic reports of research groups; LattesMiner [Alves et al. 2011a], a multilingual domain-specific language that automatically extracts individual and group-based academic information; Sucupira [Alves et al. 2011b], a system that identifies and visualizes academic social networks, which utilizes LattesMiner for data extraction; and Science Tree [Cota et al. 2021], a user-oriented platform that generates and explores the academic genealogy trees of Brazilian researchers. Despite the diversity of analytics and visualization features in these independent tools, none were designed to automatically classify Lattes CV publications based on Qualis.

Three existing browser extensions [Batista 2023, Enders Neto 2022, Marcel 2023] offer somewhat related functionalities to those of QLattes. ChromeQualis [Batista 2023] enables users to search the Qualis classification of journals and conferences directly from the extension popup window, without the need to open CAPES’s Qualis search page. Qualis Capes [Enders Neto 2022], in turn, automatically annotates the Qualis classification of journals on their Scopus database page. Finally, Qualis no Lattes [Marcel 2023], launched shortly after QLattes, offers a similar set of functionalities, including the ability to annotate the Qualis classification of journal publications directly on an author’s Lattes page. In contrast to those extensions, QLattes provides several data analytics and visualization features that seamlessly integrate with the Lattes Platform. Moreover, because its HTML pages are loosely coupled via an asynchronous notification mechanism, the tool can offer more advanced interaction modalities, providing a much richer user experience.
6. Conclusion and Future Work

This paper introduced QLattes, an open-source browser extension that automatically annotates and displays the Qualis classification of journal papers on the Lattes Platform. QLattes’ preliminary performance results, official usage statistics from Google, and voluntary user feedback indicate that an increasing number of members of the Brazilian academic community are adopting the tool as an effective companion to the Lattes Platform. We encourage readers not yet familiar with QLattes to try it out and share their experiences via the feedback form embedded in the tool.

Some of our current and future work on QLattes include extending it to extract and visualize other publication metrics, such as JCR and h5; supporting the aggregation and visualization of publication data from groups of Lattes authors; adding the capability to export/import publication data to/from external files; and investigating more accurate approaches to identifying and classifying conference publications, for example, by using transformer-based language models [Vaswani et al. 2017]. We believe that with the support of the Brazilian academic community, QLattes has the potential to evolve into a full-fledged publication management and analytics tool based on the Lattes Platform.

Acknowledgements

The human authors thank ChatGPT for assisting in writing and revising several sections of the paper, as well as generating the initial version of the Python program that produced the map plot featured in Figure 4. The authors also are grateful to the tens of QLattes users who generously shared their early experiences with the tool.

References


Dias, T. M. R. et al. (2013). Modelagem e caracterização de redes científicas: um estudo sobre a Plataforma Lattes. In Anais do II Brazilian Workshop on Social Network Analysis and Mining, pages 116–121. SBC.


