The visibility of open access publications

Results of sub-project 5 by Austrian Transition to Open Access 2 (AT2OA2)

Imprint

Cite as: Can, E., Ferus, A., Gumpenberger, C., Hölbling, L., Holzner, B., Pacher, A., Reitbrecht, C., Sams, B., & Schilhan, L. (2024). The visibility of open access publications – Results of sub-project 5 of Austrian Transition to Open Access 2 (AT2OA2). Zenodo. https://doi.org/10.5281/zenodo.14413232. Online (including various supplement materials) also available at: https://phaidra.univie.ac.at/o:2095102.

Abstract: Subproject 5 – Visibility of Open Access Publications of Austrian Transition to Open Access 2 (AT2OA2) was dedicated, among other things, to whether Open Access, in addition to improving the availability of scholarly publications, also contributes to increasing visibility. Alternative metrics (altmetrics) served as the primary hook and vehicle for investigating this and other questions, as they are no longer mainly citation-based metrics that provide information about the reception of scientific achievements. A particular focus of the activities was realizing a study in cooperation with Altmetric (Digital Science). For this purpose, a data set generated as part of the predecessor project AT2OA and covering the publication years 2015-2018 was expanded in several steps with data on publications from 2019-2021, enriched with the help of Altmetric, and then analyzed based on specific research questions. It should be noted in particular that a unique feature of this study is the use of normalized altmetrics data (Relative Index Metric). In addition to the study and its results, the final recommendations for increasing the visibility of scholarly achievements are probably the most important outcome.

Keywords: scholarly publications, Open Access, scholarly outputs, visibility, altmetrics

Contact:
c/o University of Vienna, Library and Archives
Austrian Transition to Open Access 2 (AT2OA2)
E-Mail: office-at2oa@univie.ac.at
Website: https://at2oa.at/

Contributors

Main authors	Type of contribution
Eylem Can, University of Vienna	SP5 overview of research questions, methodology, results, annex
Andreas Ferus, Academy of Fine Arts Vienna	Executive summary, methodology, results, summary of results, recommendations, references
Christian Gumpenberger, University of Vienna	Contributors to the report, SP5 overview and organisation, introduction, SP5 overview of research questions, methodology, results, summary of results, recommendations, outlook, acknowledgements
Lothar Hölbling, University of Vienna	SP5 overview of research questions, methodology, recommendations
Birgit Holzner, University of Innsbruck	Recommendations, imprint, table of contents
Andreas Pacher, TU Wien	Methodology, results
Caroline Reitbrecht, Medical University of Vienna	Recommendations
Birgit Sams, Johannes Kepler University Linz	Literature overview, references
Lisa Schilhan, University of Graz	Recommendations

Supporting authors

Andreas Brandtner, University of Vienna	Foreword
Brigitte Kromp, University of Vienna	Foreword
Horst Mayr, BOKU Vienna	Annex (use case)

Other contributors

Eylem Can, University of Vienna	SQL queries, statistical analyses, graphics and tables
Carmen Drolshagen, University of Innsbruck	Editorial work, typesetting, layout
Andreas Ferus, Academy of Fine Arts Vienna	Graphics and tables, editorial work, typesetting, layout, proofreading
Christian Gumpenberger, University of Vienna	Overall coordination, proofreading
Lothar Hölbling, University of Vienna	Data management plan, proofreading
Birgit Holzner, University of Innsbruck	Editorial work, layout, proofreading
Martin Rösel	Database programming
Lisa Schilhan, University of Graz	Proofreading
Jennifer Wright, University of Graz	English translation

Table of contents

Executive Summary

Background

Like its predecessor *AT2OA*, the *Austrian Transition to Open Access 2* (AT2OA^2) project aims primarily to drive forward the transformation of scientific publications from closed to Open Access throughout Austria. The intention is to contribute significantly to the global transformation of the scientific publishing system, facilitate access to scientific knowledge and enable social participation.

AT2OA^2's *sub-project 5 (TP5)*, *Visibility of Open Access Publications*, is dedicated, among other things, to the question of whether Open Access, in addition to improving the availability of scientific literature, also promotes an increase in visibility.

Methodology

Alternative metrics (altmetrics) served as the primary angle and the vehicle for investigating this and other questions, as information about the reception of scientific achievements is no longer provided solely by citation-based metrics.

A particular focus of the activities of TP5 was a study carried out in cooperation with Altmetric (Digital Science). For this purpose, a dataset for the years 2015–2018, generated as part of AT2OA sub-project 1, was expanded in several steps, with data on publications from 2019–2021, enriched with the help of Altmetric and then analysed based on specific research questions. This study's particularly noteworthy unique feature is the use of normalised altmetrics data (Relative Index Metric). This normalisation function is currently under development and is not yet a feature of the commercial Altmetric application (Altmetric Explorer).

Results

With a growth rate of almost 70 % within just seven years, the number of publications published increased significantly between 2015 and 2021. The proportion of publications recognised by Altmetric remained relatively constant from 2015 to 2018, increased substantially in 2019 and declined only slightly in 2020 and 2021. A similar trend can be seen in the successfully enriched records, whereby the disciplines "Medical & Health Sciences" and "Physical & Mathematical Sciences" have the highest proportion, while the "Social Sciences" and the "Humanities" are underrepresented. Here, too, is a recognisable trend over time. The proportion of successfully enriched publications is growing continuously. The number of enriched publications at least doubled between 2015 and 2021 in all disciplines.

In terms of verifiable online attention via altmetrics, no significant difference between Open Access and closed-access publications can be found. Likewise, there is no correlation between the different Open Access statuses and the different forms of online attention. However, differences can be seen in other measurements, such as the median and the density distribution.

Taking an overall view and considering individual parameters (such as the strongly diverging number of publications per discipline, etc.), it is clear that the "Humanities" have the highest normalised altmetrics values in terms of citations and with regard to most online media. The "Social Sciences" achieved the second-best values for citations, blogs and policy, and the third-best for Facebook, Twitter, news and Wikipedia. "Engineering & Technology" was the frontrunner for patents and policy documents. This shows that the disciplines least represented in Web of Science and Scopus have the best normalised altmetrics values.

In most cases, it was not possible to demonstrate any correlation between citations and online attention. Only a few cases showed a very slight correlation. The highest values were found in the "Medical & Health Sciences" cluster between citations and news indices and citations and policy indices.

Generally, considerable differences are identifiable between absolute and normalised altmetrics data. Using normalised altmetrics values significantly increased the informative value of the results in all cases.

Regarding alternative tools, it was found that there are some differences between Crossref Event Data and Altmetric. As a result, Crossref Event Data cannot be considered a replacement. It can, however, be seen as a free complementary product for carrying out altmetrics analyses, at least to a limited extent.

Initial experiments with data, kindly made available to us by DOI Service Austria, were rather disappointing and did not deliver any "positive" results at first glance, but did raise numerous questions that warrant further investigation.

Deliverables

No doubt, the most important outcome of TP5, in addition to the study conducted and its results, are the final recommendations for increasing the visibility of scientific achievements, which were in part derived from them.

These were formulated primarily from the point of view of ease of implementation and relate to the following topics: altmetrics, publications, Open Access, repositories and FIS/CRIS, affiliations, persistent identifiers, academic search engine optimisation, EU projects and academic social networks.

Foreword

As project manager of the two Austrian Transition to Open Access (AT2OA) projects, I am delighted that, after eight years of intensive collaboration, we can present an extremely successful balance sheet to which the sub-project "Visibility of Open Access publications" has made a significant contribution. Over the course of seven years (2015–2021), the project collected, processed, enriched and analysed data on scientific output in Austria. The result is a study of remarkable scope that breaks new ground in terms of the questions it asks.

In recent years, in addition to their core business as literature providers, major international publishers have developed services and tools to position themselves as data providers in the global science market. At the same time, new specialised companies are increasingly entering the market with innovative solutions. The most recent development is the increasing availability of extensive publication-related databases under open licences.

These new services and tools also include alternative metrics (altmetrics). As a comparatively young discipline in the context of research evaluation, they record the attention to and discussion of scientific work in social media, news articles, blogs and other digital platforms and provide a real-time picture of the reception and dissemination of research results.

This study exclusively used normalised altmetrics as part of a project partnership with the company Altmetric, one of the new data providers. This made it possible for the first time to use altmetrics to analyse aspects of the visibility of publications in relation to Open Access.

I would like to thank Christian Gumpenberger and Andreas Ferus, who have fulfilled their roles as sub-project leaders with exceptional competence and commitment, and their team of colleagues from nine Austrian universities for the great work they have done over the past four years in addition to their daily workload. Thanks are also due to the scientific advisor Mike Taylor and his colleague Carlos Areia, both from Altmetric, who were responsible for the data enrichment and gave critical support during the editing of the final report.

The completion of AT2OA2 marks the end of a formative and ground-breaking chapter in the study and promotion of Open Access in Austria. This is a good moment for the project to end because it leaves Austria as a centre for science and research with a tight-knit national network of knowledge and experts.

Brigitte Kromp
Project Manager AT2OA and AT2OA2
University of Vienna

Foreword

As the director of Vienna University Library, I am particularly pleased that the AT2OA² project, the successor to AT2OA, has come to a successful and fruitful conclusion. This success would not have been possible without the invaluable contributions of a national network of more than a hundred colleagues involved in these two initiatives, who addressed and worked intensively on key issues in the field of Open Access over a period of eight years.

The projects have developed recommendations and practical guidelines based on extensive surveys and data-based studies. These are important for researchers, research support, and, ultimately, university administrations. The number of contracts with academic publishers with Open Access components (transformative agreements) has increased, and a national data hub has been established for nationwide Open Access monitoring.

From my perspective as a library manager, it is all the more pleasing that with sub-project 5, an Open Access "side issue" has also been considered, as altmetrics have increasingly gained momentum in recent years. They have become as important in the context of Responsible Research Assessment as they are as a central component of Narrative Bibliometrics. In a rapidly changing scientific landscape, it is valuable and necessary to shine a light on the question of whether Open Access also increases the visibility of publications in altmetrics. A better understanding of how we can use publication-related metrics responsibly and beneficially is a key to success – for individual researchers and universities as a whole.

I want to take this opportunity to thank the two sub-project leaders, Andreas Ferus and Christian Gumpenberger and their team, as well as my colleagues at the data provider Altmetric, for their commitment and willingness to cooperate. With this project report, we have reached an end point in answering the questions for the present, although I am confident that the curiosity is not at an end. I am sure, therefore, that the network will continue to exchange information about current developments and support library and university management in fulfilling their roles as well as possible in research and developing new internal services, not least with regard to an interested public.

Andreas Brandtner
Director of the Vienna University Library

Austrian Transition to Open Access 2 (AT2OA2)

Like its predecessor AT2OA, the project Austrian Transition to Open Access 2 (AT2OA2) aims to drive the transformation from closed to Open Access in scientific publication throughout Austria. This should, among other things, contribute to the global transformation of the scientific publication system, facilitate access to scientific knowledge, and enable social participation.

The following topics will be addressed in five subprojects (TPs):

1. Transformative contracts (TP1): increasing the number of contracts with scientific publishers, which include Open Access components.
2. Austrian datahub (TP2): establishing a national datahub to monitor Open Access nationwide and to support negotiations with scientific publishers.
3. Publication costs (TP3): developing sample workflows and recommendations for recording publication costs at Austrian universities.
4. Predatory publishing (TP4): considering the phenomenon of predatory publishing in the context of changing scientific communication.
5. Altmetrics (TP5): examining the visibility of Open Access publications using alternative metrics.

Overview

Key Information:
Start: January 1st, 2021
End: December 31st, 2024
Project leader: University of Vienna
Project management: University of Vienna
Project partners: 24 (all public universities in Austria + IMP & ISTA)
Client: Federal Ministry of Education, Science and Research (Bundesministerium für Bildung Wissenschaft und Forschung, BMBWF)
Funding track: "Vorhaben zur digitalen und sozialen Transformation in der Hochschulbildung" ("Digital and social transformation projects in higher education", German only)

Project supervision:
HR Mag. Brigitte Kromp
Head of the Austrian Central Library for Physics and Chemistry Library
e-mail: brigitte.kromp@univie.ac.at

Mag. Dr. Andreas Brandtner, MAS MSc MBA (from April 2024)
e-mail: andreas.brandtner@univie.ac.at
Head of Libraries and Archives at the University of Vienna

HR Mag. Maria Seissl (until March 2024)
Former Head of Libraries and Archives at the University of Vienna

Project management:
Laura Still, BA BA
e-mail: laura.still@univie.ac.at
Melanie Stummvoll, BA
e-mail: melanie.stummvoll@univie.ac.at

Mag. Ursula Ulrych
e-mail: ursula.ulrych@univie.ac.at

Mag. Lothar Hölbling,
e-mail: lothar.hoelbling@univie.ac.at
(until September 2022)

Technical support:
Tobias Zarka
e-mail: tobias.zarka@univie.ac.at

Contact:
e-mail: office-at2oa@univie.ac.at
Website: https://at2oa.at/en/

Sub-project 5 – TP5 – Visibility of Open Access publications

TP5 addresses the question of whether Open Access, in addition to improving the availability of scientific literature, also promotes an increase in visibility.

Time frame:
Start: 12th November 2021
End: 31st December 2024

Project leader:
Dr. Christian Gumpenberger, MAS MSc
Head of the Department for Bibliometrics and Publication Strategies, University of Vienna
e-mail: christian.gumpenberger@univie.ac.at

Deputy project leader:
Andreas Ferus, MSc
Head of University Library, Academy of Fine Arts Vienna
e-mail: a.ferus@akbild.ac.at

TP5 working groups:

Working group "Altmetric analyses – WoS & Scopus"
Coordination: Christian Gumpenberger

Working group "Alternatives and workflows"
Coordination: Andreas Ferus

Working group "Visibility recommendations"
Coordination: Dr. Lisa Schilhan, University of Graz
e-mail: lisa.schilhan@uni-graz.at

Working group "Report creation"
Coordination: Dr. Birgit Holzner, University of Innsbruck
e-mail: birgit.holzner@uibk.ac.at

Scientific advisor:
Mike Taylor
Head of Data Insights, Digital Science
e-mail: m.taylor@digital-science.com

TP5 is supported by Altmetric as part of a partnership.

Contact:
e-mail: tp5-at2oa2.ub@univie.ac.at
Website: https://www.at2oa.at/teilprojekte/#tp5

I. Introduction

The why

The issue of visibility is unavoidable for all scientists aiming to have successful academic careers in the wake of the exponential growth in research output. Increased visibility raises the probability of being indexed, found, read and ultimately – in whatever form – received as a scientist. Digitalisation, progressing at breathtaking speed, offers challenges and opportunities.

It is no longer just citation-based metrics that provide information about the reception of academic publications and are limited to the "publish or perish" community. Meanwhile, the spectrums of what can be quantified and of interested stakeholders have become much broader.

The how

Within the scope of AT2OA[2], TP5 (sub-project 5) aimed to highlight the topic of visibility and how to increase it. Alternative metrics (altmetrics) served as both a peg on which to hang the analysis and a vehicle. Chapter III, "TP5 research questions at a glance" details the questions we addressed. The framework for our analyses can be found in Chapter IV, "Methodology".

The what for

This report is a revealing snapshot in many ways. On the one hand, it provides information on the overall Austrian publication output indexed in WoS and Scopus for the years 2015–2021 regarding the availability of altmetrics data. Based on this, it is possible to follow up on the topic of altmetrics in Austria at both the disciplinary and institutional levels.

Furthermore, the report also provides information on randomised supplementary studies using alternative data sources.

The most important deliverables, however, are the final recommendations for increasing visibility, which were primarily formulated with a view to ease of implementation.

With this report, we would like to contribute an instrument for raising awareness and, ultimately, also for change towards increased visibility for scientific publications.

II. Literature overview TP5

The Open Access movement arose from the call to make research results freely accessible worldwide on the internet so that full texts of research articles could be read, downloaded, copied, shared, printed, searched and linked for lawful purposes at no cost (Open Society Foundations, 2002).

Important milestones in this movement are the "Budapest Open Access Initiative" (Open Society Foundations, 2002), the "Bethesda Statement on Open Access Publishing" (Bethesda Statement on Open Access Publishing, 2003) and the "Berlin Declaration on Open Access" (Max Planck Society, 2003) (Suber, 2012).

Other initiatives and funding programmes relevant to Austria include the cOAlition S and Plan S initiatives (cOAlition S, 2018), the Open Access policies of Austrian funding bodies (FWF, n.d.), the Open Science Policy Austria (Open Science Policy Austria, 2022), the Austrian Academic Library Consortium (Kooperation E-Medien Österreich, n. d.) Open Access transformative agreements, the Open Science Austria (formerly Open Access Network Austria, Open Science Austria, 2022), and the previous project Austrian Transition to Open Access (AT2OA) (Austrian Transition to Open Access, 2017).

Although the spread of Open Access has great momentum, considerable differences exist between the various disciplines (Laakso & Björk, 2022) and countries (Simard et al., 2022; Basson et al., 2022).

Due to significant differences in the methodology, sample size and period under analysis of the existing studies, it is scarcely possible to make a universal statement about the proportion of OA publications.

However, if one follows the development of this proportion over the years, it seems reasonable to assume that almost half of all research articles are now available free of charge (Piwowar et al., 2018; Basson et al., 2022; Taylor, 2024; European Commission, 2019; Simard et al., 2022).

In comparison, the OA share of Austrian publications in the SciVal tool, which uses data from the Scopus database, is 55.5 % (Elsevier, 2024).

Numerous studies have used citations as an indicator to show the influence and quality of research in OA publications compared to fee-based publications.

These studies have shown that OA publications tend to have higher citation counts than subscription-based publications (Wang et al., 2015; Piwowar et al., 2018; Holmberg et al., 2020).

With the spread of social media and the digitalisation of science, alternative metrics or altmetrics have become increasingly important as a new data source for scientific communication (Sugimoto et al., 2017).

The term altmetrics was introduced by Priem et al. (2011) in "altmetrics: a manifesto": "... the growth of new, online scholarly tools allows us to make new filters; these altmetrics reflect the broad, rapid impact of scholarship ... ".

Erdt et al. (2016, p. 1118) summarise the definition of altmetrics: "... the common understanding across all definitions is that altmetrics are new or alternative metrics to the established metrics for measuring scholarly impact. The main difference in the definitions however is in how and where altmetrics can be found – activities on Social Media, based on the Social Web, observing activity in online tools and systems, engagement with research output, based on social media platforms and tools, scholarly activities or various user activities in social media environments."

For Sugimoto et al. (2017, p. 2051), the use of social media in an academic context includes: "... social networking, social bookmarking, social data sharing, video, blogging, microblogging, as well as social recommending, rating and reviewing."

Several studies have investigated the relationship between alternative and traditional metrics to determine whether and how altmetrics differ from citations. Most studies confirm a positive, albeit weak, correlation, suggesting that while there is a relationship, different types of impact are measured (Costas et al., 2015; Stephen & Stahlschmidt, 2024; Thelwall et al., 2013; Haustein et al., 2014; Zahedi et al., 2014; Erdt et al., 2016).

Most authors, therefore, do not regard altmetrics as a substitute but rather as a supplement to traditional metrics due to their different impact and greater reach (Thelwall, 2020; Bornmann, 2014; Costas et al., 2015; Haustein et al., 2014; Zahedi et al., 2014; Erdt et al., 2016; Fang et al., 2020; Wouters & Costas, 2012).

Stephen and Stahlschmidt (2024) subdivide the impact depending on whether the social media and platforms are used more by the academic community or by a broader audience. Social media and platforms used by an academic audience, such as Mendeley and ResearchGate, tend to have a stronger connection to the traditional metrics than those used by a non-academic audience and are, therefore, also characterised by different communication patterns.

The advantages of the alternative metrics compared to the classic metrics are their immediate effectiveness, the diversity of the research output and the greater scope, which applies not only to research (scientific impact) but also to other areas of society (societal impact) (Holmberg et al., 2019; Thelwall, 2020; Bornmann, 2014; Stephen & Stahlschmidt, 2024; Fang et al., 2020; Piwowar, 2013; Zahedi et al., 2014; Thelwall et al., 2013; Wouters & Costas, 2012; Erdt et al., 2016; Holmberg et al., 2020).

However, these advantages are offset by some challenges related to the small amount of data from most altmetric sources (except Mendeley and X (Twitter)), data quality and stability, data manipulability, limitations in data collection (API, DOI) and the lack of quality standards and controls (Thelwall et al., 2013; Zahedi et al., 2014; Fang et al., 2020; Mohammadi et al., 2015; Mohammadi et al., 2018; Thelwall, 2020; Costas et al., 2015; Haustein et al., 2014; Erdt et al., 2016; Haustein, 2016; Bornmann, 2014; Fang et al., 2020).

Research shows that even with the alternative metrics, Open Access articles tend to receive more online activity than subscription-based articles (Holmberg et al., 2020; Wang et al., 2015).

In their study, Holmberg et al. (2020) point out that discipline- and platform-specific differences can lead not only to an "Open Access advantage" but also to an "Open Access disadvantage".

Research areas with the highest proportion of alternative metrics are medicine, the natural sciences and the social sciences and humanities (Zahedi et al., 2014; Costas et al., 2015; Thelwall et al., 2013; Stephen & Stahlschmidt, 2024; Fang et al., 2020).

Alternative metrics, therefore, have a potential added benefit for the social sciences and humanities, which could previously only be analysed to a limited extent using traditional citation analyses (Costas et al., 2015; Wouters & Costas, 2012; 2014; Fang et al., 2020).

Haustein et al. (2015) find that disciplines that are more embedded in society, such as the social sciences and humanities, or disciplines that deal with specific problems in people's daily lives, such as the biomedical and health sciences, as well as natural and earth sciences, are more likely to be present on social media platforms than technical and applied disciplines such as the natural and engineering sciences or disciplines that deal with more complex and technical topics such as mathematics and computer science.

A comprehensive study by Taylor (2024) shows that some disciplines, such as medicine and health sciences, life sciences and the humanities, are strengthened by the introduction of OA, while other disciplines, such as the social sciences, may be weakened. It is also noteworthy that some disciplines gain visibility and socio-economic influence through their OA status, while other disciplines, such as the social sciences and humanities, may not benefit in the same way.

The alternative metrics are very heterogeneous in terms of distribution in the various disciplines, target groups, areas of application, methodology and availability (Haustein, 2016; Taylor, 2023; Zahedi et al., 2014; Fang et al., 2020; Ortega, 2018; Holmberg et al., 2020).

This heterogeneity makes it difficult to create a common conceptual framework (Haustein, 2016; Sugimoto et al., 2017).

Only by normalising the altmetrics, taking into account the year of publication and the discipline, is it possible to compare the alternative research output on different topics from different periods (Thelwall, 2017; Taylor, 2023; Bornmann, 2014).

III. TP5 research questions at a glance

1. **Ascertaining the absolute and percentage shares of records with altmetrics in comparison to the total data volume (incl. breakdowns by publication year and discipline).**

   Background to the question
   On the one hand, the aim is to determine what proportion of data has been successfully enriched by altmetrics in relation to the total volume of Web of Science (WoS) and Scopus data analysed. On the other hand, we are also interested in the development over time, which is presented by discipline based on the "Fields of Research".

2. **Does Open Access correlate with increased online attention in altmetrics (and thus increased visibility)? Can differences be identified between the various Open Access statuses in relation to altmetrics?**

   Background to the question
   The fundamental aim of TP5 is to establish whether Open Access publications gain more visibility in terms of altmetrics (overall and differentiated by the level of online attention) than closed access publications. Potential differences in relation to the various Open Access statuses are also of interest.

3. **Are there any differences between disciplines with regard to altmetrics, especially in the various forms of online attention?**

   Background to the question
   The insights gained are not only of academic interest but could also potentially be integrated into practical publication strategy consultations. This would allow research output to be disseminated more directly to different disciplines and target groups.

4. **Can correlations be identified between citations and online attention?**

   Background to the question
   Correlations between citations and altmetrics have been analysed several times in the past. The results of these studies varied depending on both the discipline and the type of online attention. Of particular interest here is the specifically Austrian situation in comparison to international results.

5. **Do the results become more meaningful through the use of normalised altmetrics values?**

Background to the question
Normalising data on publication years, subject areas, and publication types is already common practice in citation analyses and is often available as a separate function in conventional analysis tools. TP5 will explore the potential of applying this concept to altmetrics to obtain more relevant results.

6. **How useful is Crossref Event Data as a free altmetrics provider? How expedient are altmetrics analyses based on alternative data sources?**

Background to the question
It is known that only a section of the actual publication output is captured when restricting the publication data used to that indexed in WoS and Scopus, and that this leads to certain (sub)disciplines being comparatively underrepresented. Using a fee-based tool to carry out the altmetrics analysis based on this data leads to a further limitation. For this reason, TP5 at least uses alternative data sources and a free altmetrics analysis tool on a random basis.

7. **How can visibility be increased outside of altmetrics?**

Background to the question
The aim of answering all the previous questions is to develop suitable recommendations for increasing the visibility of scientific publications beyond altmetrics.

IV. Methodology (research questions 1–5)

Data preparation and enrichment

AT2OA2[1] sub-project 5 (TP5) builds in part on data from sub-project 1 (TP1) from the previous project AT2OA (duration 2017–2020). It expands it and supplements the basic publication data collected with data from Altmetric. Here, we will briefly explain the methodology for data collection and enrichment.

Data on scientific publication output in Austria was collected and processed for 2015–2018 as part of a publication output analysis (TP1 AT2OA). The data was taken from the two citation databases, Web of Science Core Collection (WoS CC) and Scopus. It was the most detailed and comprehensive survey of scientific publications in Austria to date, and it enabled analysis down to the level of individual publications.

The publication information was downloaded in several stages from WoS CC and Scopus. Incorrect entries were corrected, duplicates removed, inconsistencies corrected, and finally, the data was prepared for the project-specific requirements of TP1. Further information on this can be found in a 2019 publication (Hölbling, 2019).[2]

The data collection on scientific publication output in Austria was carried out in TP5 in essentially the same way for 2019 to 2021. As part of a collaboration between TP5 and the data provider Altmetric[3] (Digital Science[4]) Mike Taylor, Head of Data Insights at Digital Science, took on the role of scientific advisor. Together with his colleague Carlos Areia, they were responsible for enriching data with altmetrics and Open Access information. DOIs and PubMed IDs were used as unique identifiers for data synchronisation between the AT2OA2 datasets and the data from Altmetric. In 2022 and 2023, the individual AT2OA2 datasets were periodically enriched with information from Altmetric seven times.

The main unique feature of this study is the use of normalised metrics for Altmetric data (relative index metric). Normalised metrics are still under development and are not yet part of the commercial Altmetric application. They were provided exclusively to TP5 as part of the project collaboration. Taylor and Areia calculated the relative index metric based on analyses using Google BigQuery.

The relative index metric uses an algorithm similar to the FWCI (Elsevier's Field-weighted Citation Impact[5]).[6] It is calculated by dividing the actual number of mentions in so-called attention sources by the expected number for publications of the same document type, publication year, and subject area. A value of 1.00 means that an article performs "as well as expected"; >1 is better; <1 is worse.

1 https://www.at2oa.at/en/
2 Hölbling, Lothar: Datenerhebung und Analyse des Publikationsoutputs von Forschenden an österreichischen Universitäten und außeruniversitären Forschungseinrichtungen 2015 bis 2017 im Rahmen von AT2OA – Werkstattbericht zu einer bibliometrischen Studie. In: Mitteilungen der VÖB 72 / 2019, Nr. 1: AT2OA, S. 50–58. https://doi.org/10.31263/voebm.v72i1.2290
3 https://www.altmetric.com/
4 https://www.digital-science.com/
5 https://helpcenter.pure.elsevier.com/en_US/data-sources-and-integrations/field-weighted-citation-impact-fwci-metrics
6 Relative index metric differs from FWCI in three aspects:
- The calculation is not static, but is repeated weekly using live data.
- The factor "subject area" is applied at publication level and not at journal level.
- It calculates the arithmetic mean and not the harmonic mean.

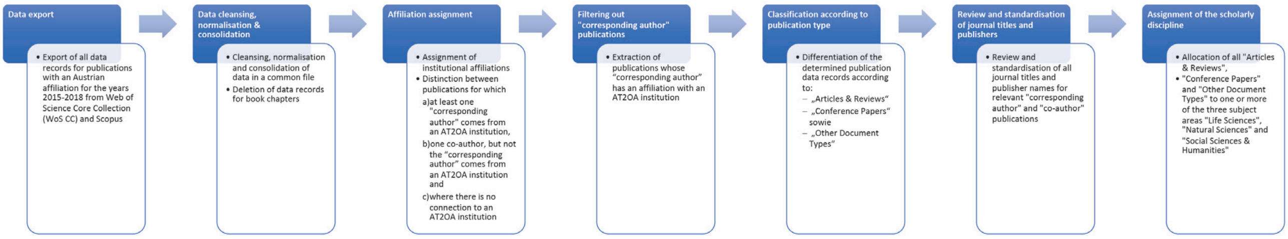

Fig. 1: Data processing in AT2OA TP1 (Hölbling, 2019)

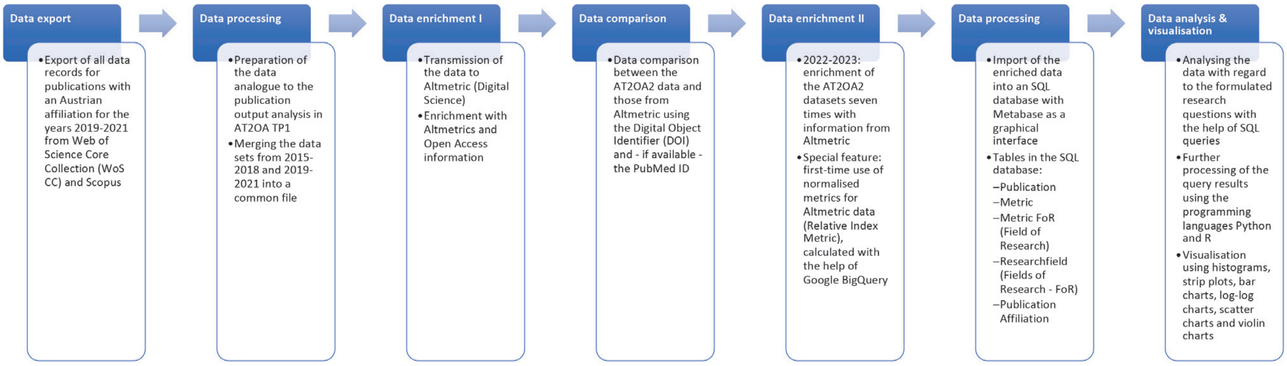

Fig. 2: Data processing in AT2OA² TP5 (Can & Hölbling, 2024)

Examples:

Altmetric score[7]

- The absolute Altmetric score for an article is 245.
- The score index (relative index metric) is 54.68.
- This means that the score for this article is 55 times better than for an average article (same document type, publication year and subject area).

Score for individual mentions

- An article has 24 mentions in news sources
- Its news index (relative index metric) is 85.38.
- This means that this article is 85 times better than an average article in terms of news mentions (same document type, year of publication and subject area.

Excel was used to collect and enrich data. Metabase[8] was used for the subsequent data analysis and creating visualisations.

SQL data queries

The AT2OA2 project programmer Martin Rösel imported the collected and enriched data on publication output into an SQL database with the graphical interface "Metabase" to be evaluated and visualised. It was summarised in various tables.

Database "Metabase": when users log in to Metabase, they see the dashboard (Figure 3. The database tp5 and the corresponding tables are found via the menu on the left under "Daten durchsuchen" (search data).

7 https://www.altmetric.com/about-us/our-data/donut-and-altmetric-attention-score/ and https://help.altmetric.com/support/solutions/articles/6000233311-how-is-the-altmetric-attention-score-calculated-

8 https://www.metabase.com/

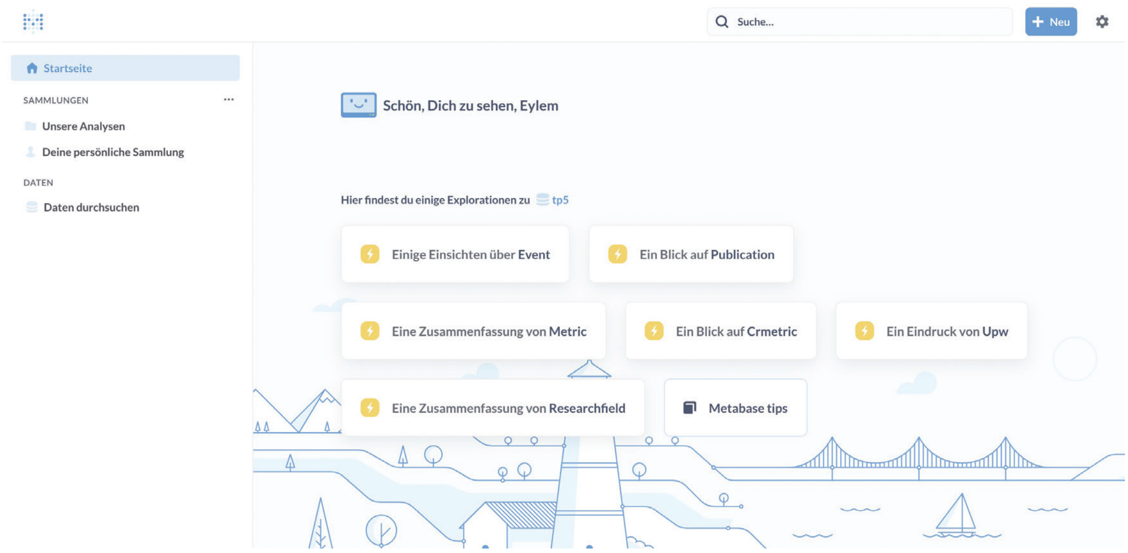

Fig. 3: Metabase website start page

Tables in the database:

- Publication: all publication data from Web of Science and Scopus. Each entry has a unique ID: Atid + year. Atid stands for AT2OA2 ID and corresponds to a consecutive identification code assigned while creating the publication output datasets.
- Metric: the data enriched by Altmetric. Unique ID: Atid + year + quarter
- Metric FoR: FoRids (Field of Research IDs) are assigned to the Metric IDs.
- Researchfield: a breakdown of all of the Fields of Research
- Publication Affiliation: institutional affiliations are assigned to the IDs from the Publication table. The type of authorship (corresponding author or co-author) is also listed.

SQL query: for database queries, click on "+Neu" at the top right of the start page and then on "SQL-Abfrage" and select the tp5 database. Now, you can create a query using SQL syntax.

A detailed introduction to the database and the procedure for creating a query can be found in the TP5 project wiki.

Each TP5 research question was translated into one or more SQL queries to analyse the data accordingly. All the tables were linked with data relevant to the research question using SQL syntax.

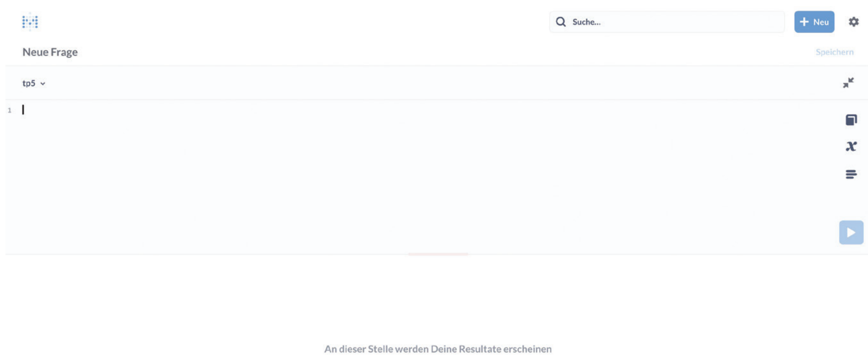

Fig. 4: SQL query

Research question 1:

This query only used the data from the last three altmetrics enrichments from 2023.

Anything with an online attention score of 0 was excluded in order to capture only the successfully enriched data. The sum of each form of online attention (Twitter, Facebook, etc.) and each enrichment was then calculated and summarised in a corresponding table. The number of publications in the table for the "Fields of Research" is higher than the total number of publications recorded, as some were assigned more than one "Field of Research ID".

A table was chosen to visualise the results rather than a graph due to the large size of the dataset.

Research question 2:

This query only took the data from the last enrichment into account. All records with an index of zero were excluded, as were any that were not Open Access.

Strip plots and violin diagrams were chosen for the visualisation.

Strip plots

In this form of visualisation, the filtered data points are plotted in a strip for each category, making it also possible to assess the distribution and density of the data. It is similar to a scatter plot, but the points are slightly offset so they do not overlap. The categories are plotted on the x-axis.

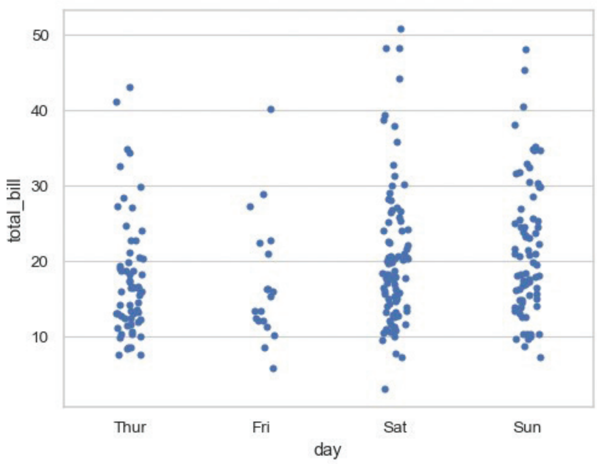

Fig. 5: Example of a strip plot (Waskom, 2024)

Strip plots make it easy to identify outliers and patterns (Michael Messer, 2019). The median, the value that separates the top 50 % of the data from the bottom 50 %, was also added to the graphs. Figure 5 shows an example of a strip plot.

Violin diagram

A violin diagram is a form of visualisation that embeds a box plot in the density distribution to combine the two.

Structure

Density distribution: The violin's shape provides information about the data distribution, as a wide point indicates a high density and a narrow point a low density. The symmetry or asymmetry indicates an even or uneven data distribution (Healy, 2018).

Comparability: The violin diagram offers a good way to make comparisons, particularly regarding differences in distribution and the median.

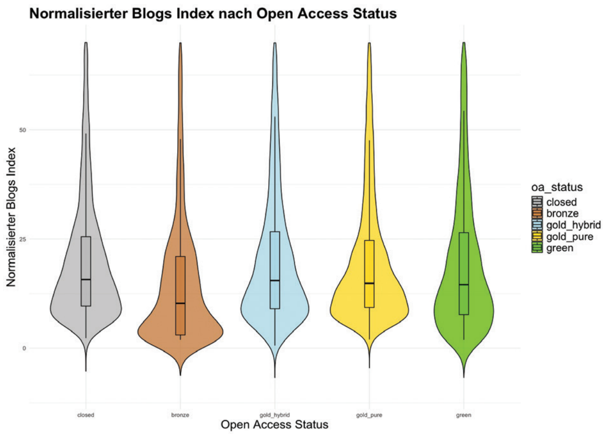

Fig. 6: Example of a violin diagram

Research question 3 and 5:

The data from the last enrichment was used for research questions 3 and 5. Additional filters were set here, according to which neither the absolute nor the normalised attention score could equal zero. Each publication entry is assigned a so-called FoRid (Field of Research ID) corresponding to a discipline.

The FoRids were divided into six groups (clusters) to allow a clearer presentation of the results. This grouping is illustrated in the following table.

Two forms of graph were chosen for the visualisation. Firstly, histograms with equidistant intervals in relation to the normalised and absolute values.

There is a histogram for each of the six discipline clusters so that each interval range is subdivided into six diagrams.

The number of publications is displayed along the x-axis with a bar representing the value for that range. The histogram example in Figure 7 was taken from (Nuzzo, 2019).

Table 1: Composition of the „Fields of Research“

FoRid	NAME	Discipline
30	AGRICULTURAL, VETERINARY AND FOOD SCIENCES	Life Sciences
31	BIOLOGICAL SCIENCES	Life Sciences
41	ENVIRONMENTAL SCIENCES	Life Sciences
40	ENGINEERING	Engineering & Technology
46	INFORMATION AND COMPUTING SCIENCES	Engineering & Technology
32	BIOMEDICAL AND CLINICAL SCIENCES	Medical & Health Sciences
42	HEALTH SCIENCES	Medical & Health Sciences
37	EARTH SCIENCES	Physical & Mathematical Sciences
34	CHEMICAL SCIENCES	Physical & Mathematical Sciences
49	MATHEMATICAL SCIENCES	Physical & Mathematical Sciences
51	PHYSICAL SCIENCES	Physical & Mathematical Sciences
52	PSYCHOLOGY	Physical & Mathematical Sciences
33	BUILT ENVIRONMENT AND DESIGN	Social Sciences
35	COMMERCE, MANAGEMENT, TOURISM AND SERVICES	Social Sciences
38	ECONOMICS	Social Sciences
39	EDUCATION	Social Sciences
44	HUMAN SOCIETY	Social Sciences
48	LAW AND LEGAL STUDIES	Social Sciences
50	PHILOSOPHY AND RELIGIOUS STUDIES	Social Sciences
36	CREATIVE ARTS AND WRITING	Humanities
43	HISTORY, HERITAGE AND ARCHAEOLOGY	Humanities
45	INDIGENOUS STUDIES	Humanities
47	LANGUAGE, COMMUNICATION AND CULTURE	Humanities

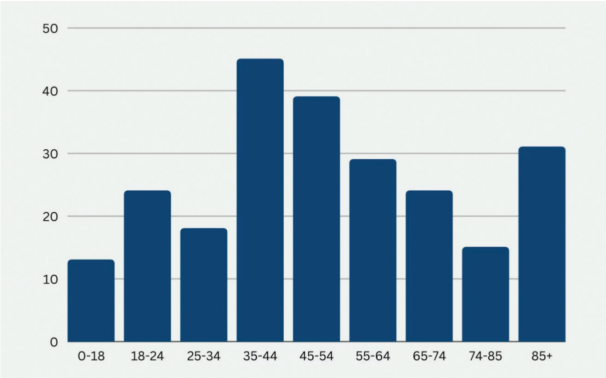

Fig. 7: Example of a histogram
(DATA BASE CAMP, 2023)

A second type of visualisation allows a direct comparison between the six discipline clusters. For this, the mean value of the six clusters is determined for each online attention source (Twitter, news, etc.) and displayed as a bar in a diagram, which also shows the standard error and the median. The median is the point in the range of data where 50 % of the data is above and 50 % is below. The standard error helps to determine how well the true mean value reflects the actual mean value, or, in other words, how much the mean value of the analysed sample deviates from the mean value of the total sample (Lang, 2017).

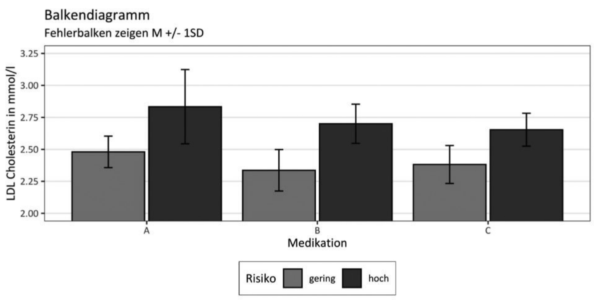

Fig. 8: Example of a bar chart with error bars,
(Ortmann, 2022)

Research question 4:

The data sets from the last enrichment were used for research question 4, and all entries with an attention score (absolute and normalised values) of zero were excluded. Visualisation is in the form of scatterplots.

A linear regression and a parabola have been added to the diagrams to analyse a linear or quadratic relationship.

Linear regression is a model that allows analysis of the relationship between an independent and a dependent variable. For this purpose, a linear function is placed in the scatter plot, keeping the distance between the data points and the straight line to a minimum (Astrid Schneider, 2010), (Weisberg, 2013). For a quadratic relationship, this results in a parabola, which is, again, embedded in the data. The function should thus visually cover as many data points as possible. The R2 values for both models are shown in the diagram. These values are statistical indicators that describe the model's success and are also known as the coefficient of determination. The more closely the determined model matches the data, the closer the value is to 1, but it is never less than 0 (Weisberg, 2013). This type of diagram can quickly capture the relationship between the data. Figure 9 shows an example.

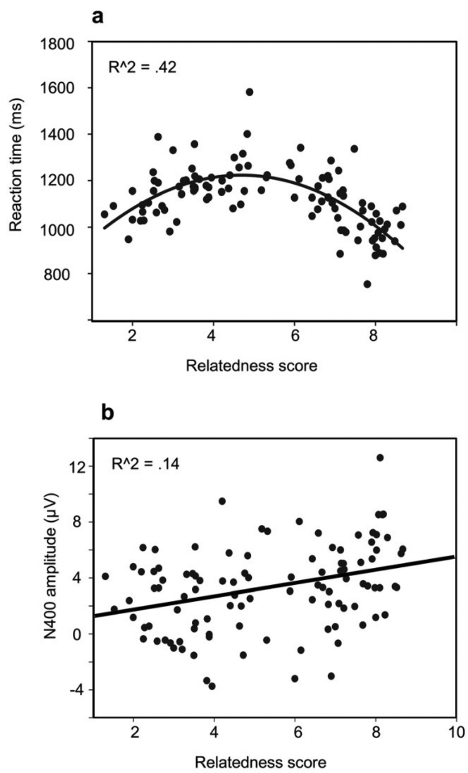

Fig. 9: (a) Parabola, (b) linear regression
(Jan Rouke Kuipers, 2018)

The data analysed in research question 4 is extensive, with many points to present in both the lower and the very high range of the scales. A logarithmic presentation is particularly suitable for this type of data.

The so-called log-log diagram proved most suitable for the present study. In this diagram, both the x-axis and the y-axis are logarithmic. While the intervals on the axes are equidistant in standard diagrams, this changes with logarithmic scales (Healy, 2018).

The new X and Y values are calculated using the following two formulae. As shown in Figure 8, in contrast to conventional scales, the first interval corresponds to a factor of 10 and the next, double, interval to 10^2, i.e. 100 (GeoGebra, 2024).

$x_{neu} = \log(x)$ $\qquad$ $y_{neu} = \log(y)$

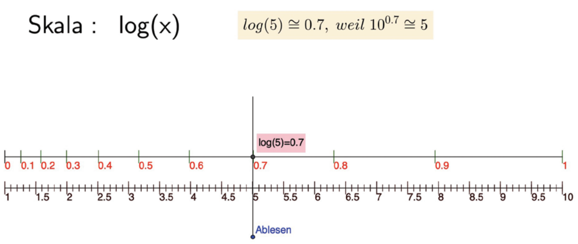

Fig. 10: Scale with x and the log(x), (Wengler, o. J.)

Statistical correlation coefficient

The Spearman rank correlation coefficient was used for research questions 2 and 3. This is suitable for ranks and is independent of the distribution of the data, which is why it is used in this study to analyse information on Open Access status. The coefficient can be used to determine the strength and direction of the relationship between the variables. A coefficient of 1 or -1 indicates a monotonic positive/negative relationship and 0 indicates no correlation (Gregory W. Corder, 2014).

Data types and formats

The programming languages "Python" and "R" were used for further processing. "R" was used mainly for the graphics, while "Python" was used to calculate the statistical tests and characteristics. The graphics were converted into .pdf files for the analysis.

Methodology (Research question 6):

A Crossref Event Data analysis was carried out for the first time in 2022 (in the course of which, among other things, no statistically significant difference was found between Crossref Event Data and Altmetric with regard to Tweets).

Despite various differences between the two altmetrics sources (including data types), metadata from Crossref Event Data was finally collected in August and September 2023 for all publications with a DOI, which also belonged to the AT2OA2 dataset already enriched by Altmetric. The "events" recorded in CrossRef Event Data were requested via https://api.eventdata.crossref.org/v1/events?obj-id=[DOI].

The collected data is available in the API in a branching JSON format but was transformed into a "flat" table for presentation. The aim was to aggregate the information in such a way that it could be compared with the AT2OA2 data. The "events" considered by Crossref Event Data were reduced to the actual number of events. Only the overlapping events occurring in both the Crossref Event Data and the enriched Altmetric data sets were taken into account. As a result, individual pieces of information were lost as they were only found in one source and, therefore, could not be compared (see the left and middle columns in Table 2).[9]

This left only a few source types for further consideration as comparative values (see the third column of Table 2). These included mentions on Wikipedia, although Crossref Event Data excluded duplicates as well as any articles with a colon in the title (e.g. "Talk:XYZ", "User:XYZ", or "Draft:XYZ"). Citations in patents, which are referred to as source type "cambia-lens" in Crossref Event Data, were still taken into account (as "patents" to retain the Altmetric terminology). Conventional scientific citation events were treated in the same way. Here, the source types labelled as "crossref" or "datacite" in Crossref Event Data were converted to the Altmetric source type "citations". The same applied to all those events that had either the value "cites" or "references" in the "relation_type_id" column (but not, for example, the value "is-preprint-of"). Incidentally, each DOI-DOI combination was only counted

9 Not all source types recorded by Altmetric were found in the AT2OA2 dataset. For example, Altmetric records mentions of scientific publications on StackExchange (as so-called Q&A sources), but these were not found in the AT2OA2 data.

Table 2: Source types for Crossref Event Data and Altmetric: discrepancies and similarities

Source types („source_id") for Crossref Event Data that did not have a clear equivalent in Altmetric (and were therefore not included)	Source types in Altmetric that did not have a clear equivalent in Crossref Event Data (and were therefore not included)	Source types that could be found in both data sets and were used for further analysis
• hypothesis • reddit-links • web (partially) • f1000 • facultyopinions • stackexchange • plaudit	• twitter • policy • facebook • all columns ending in „_index"	• news • blogs • patent • wikipedia • citations • score

once for these "citations". Duplicates were deleted. These presumably occurred when one work cited another twice, and the corresponding reference appeared twice in the metadata. However, it should be mentioned that Crossref Event Data only began recording conventional (scientific) citation data relatively late. A work's total number of citations may, therefore, be lower in Crossref Event Data (especially for older works) than is actually the case.

Crossref Event Data also had "wordpressdotcom" as the source type ("source_id"). This contained mentions of scientific publications in blogs hosted on Wordpress.com. For this reason, this source type was translated into the Altmetric terminology of "blogs". The source type "web", which comprised Internet addresses from highly heterogeneous sources, was coded manually and treated differently in each case. If the URL contained the character string "blog" (this mainly concerned the Blogspot platform), then the mention was assigned to the source type "blogs". This applied to 112 different domains and subdomains. The pages "phys.org", "arstechnica.com" and "arstechnica.co.uk" were assigned to the source type "news". Events on the page "citeulike.org" were omitted, as it is a bookmark platform with no equivalent in the Altmetric dataset. Those web sources that had "philpapers.org" and "stackexchange" in the URL were not included in the further analysis either due to the lack of a clear Altmetric counterpart.

Finally, weighting was integrated into the Crossref Event Data dataset based on the score weighting Altmetric attributes to each source type (Altmetric, 2023). Each scientific citation was multiplied by 0; each Wikipedia mention by 3; each news report by 8; each reference in a patent by 3; each link in a blog by 5. The total is listed in the "score" column. For example, if a work was scientifically cited ten times, mentioned in six news sources and linked on Wikipedia once, the addition of (10*0)+(6*8)+(1*3) resulted in a score of 51.

With regard to the use of alternative data sources for determining any altmetrics, random entries from a simple list of all DOIs assigned to date by the DOI Service Austria were examined in Altmetric at the beginning of February 2024. The relevant findings are included in the "Results" section of this report.

V. Results

Research question 1: Ascertaining the absolute and percentage shares of records with altmetrics in comparison to the total data volume (incl. breakdowns by publication year and discipline)

For ease of understanding, the results of the analyses carried out are presented in tabular form.

Table 3 shows a significant increase in the number of publications over the years. In 2015, the number of publications was 25,008; in 2021, it rose to 42,340. This equals a significant increase in output of almost 70 % within seven years.

The proportion of publications recognised by Altmetric remained relatively constant from 2015 to 2018, between 67 % and 73 %. It then rose to 92 % in 2019 and declined only slightly to 89 % in 2020 and 88 % in 2021.

A similar trend can be seen in the successfully enriched records. A publication is considered successfully enriched if the normalised Attention Score is greater than 0. In 2015 and 2016, the proportion was 34 %, which only rose by a few percentage points in the following two years. In the years 2019 to 2021, however, there was a larger increase, with the proportion rising to as much as 48 %.

Table 3: Overview of the publications recognised and enriched by Altmetric

	2015	2016	2017	2018	2019	2020	2021	
Total number of publications	25008	28163	28133	28921	30827	37528	42340	220920
Sum of publications recognized by Altmetric	17559	18759	19427	21077	28461	33398	37345	176026
Percentage of publications recognized by Altmetric	70%	67%	69%	73%	92%	89%	88%	
	2015	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Total number of publications	25008	28163	28133	28921	30827	37528	42340	220920
Sum of publications enriched by Altmetric (Attention Score Index >0)	8564	9516	9969	11122	14643	17711	20384	91909
Percentage of publications that were successfully enriched	34%	34%	35%	38%	48%	47%	48%	

The breakdown by "Field of Research" (FOR) is shown in Table 4. It is important to note that some publications are assigned to several FOR categories, meaning that the total number of publications in Table 4 is higher than in Table 3. The disciplines "Medical & Health Sciences" and "Physical & Mathematical Sciences" have the highest proportions of successfully enriched data. The disciplines "Social Sciences" and "Humanities" are the most underrepresented with 9 % and 2 % respectively.

Table 4: Overview of publications enriched by Altmetric by discipline

	Life Sciences	Engineering & Technology	Medical & Health Sciences	Physical & Mathematical Sciences	Social Sciences	Humanities		
Sum of publications enriched by Altmetric (Attention Score Index >0)	23094	11400	38402	28585	9636	2097		113214
Percentage of publications enriched by Altmetric	20%	10%	34%	25%	9%	2%		
Life Sciences	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	2250	2519	2526	2854	3753	4180	5012	23094
Percentage of publications enriched by Altmetric	10%	11%	11%	12%	16%	18%	22%	
Engineering & Technology	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	1068	1152	1250	1391	1846	2206	2487	11400
Percentage of publications enriched by Altmetric	9%	10%	11%	12%	16%	19%	22%	
Medical & Health Sciences	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	3564	3820	4022	4263	6068	7570	9095	38402
Percentage of publications enriched by Altmetric	9%	10%	10%	11%	16%	20%	24%	
Physical & Mathematical Sciences	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	2905	3303	3448	3892	4576	5135	5326	28585
Percentage of publications enriched by Altmetric	10%	12%	12%	14%	16%	18%	19%	
Social Sciences	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	616	765	798	1018	1623	2176	2640	9636
Percentage of publications enriched by Altmetric	6%	8%	8%	11%	17%	23%	27%	
Humanities	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	
Sum of publications enriched by Altmetric (Attention Score Index >0)	121	149	240	271	299	455	562	2097
Percentage of publications enriched by Altmetric	6%	7%	11%	13%	14%	22%	27%	

A trend over time can also be seen within the disciplines. The proportion of successfully enriched publications is growing continuously. Data-enriched publications have at least doubled between 2015 and 2021 in all disciplines. In 2015, in the discipline of "Humanities," only 6 % of all publications were enriched with data, whereas in 2021, this figure had already reached 29 %.

The trend can be seen clearly in Figure 9.

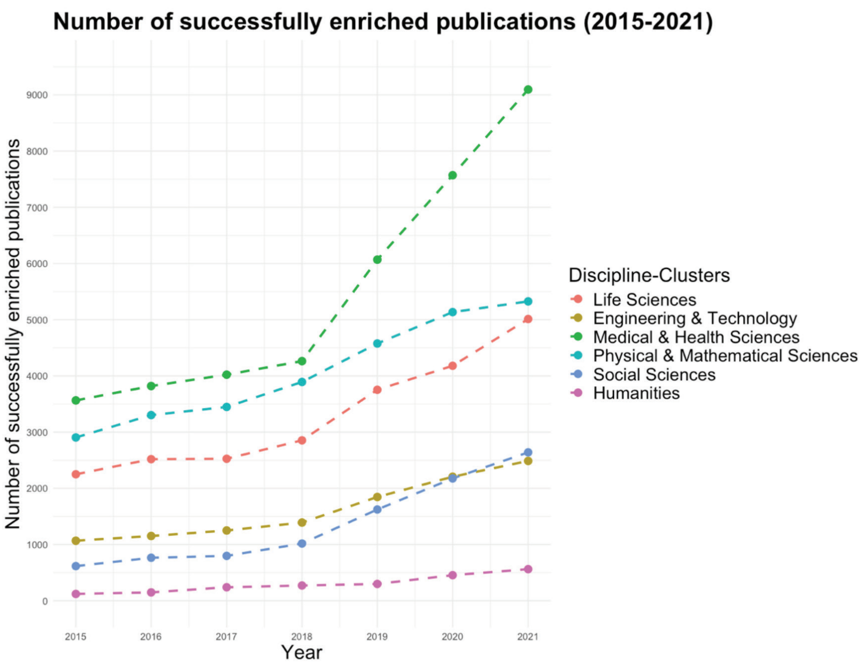

Fig. 9: Trend in successfully enriched publications by discipline, 2015–2021

Conclusion

Overall, the recognition of publications and data enrichment can both be considered successful. A continuous increase can be seen in the overall publication data and in the analysis by discipline.

Research question 2: Does Open Access correlate with increased online attention in Altmetric (and thus increased visibility)? Can differences be identified between the various Open Access statuses in relation to altmetrics?

Our in-depth analyses show no significant difference in online attention, verifiable via Altmetric, between Open Access and closed-access publications in the available publication data. There is also no correlation between the different Open Access statuses and the different forms of online attention.

Thus, at least for the purposes of this sub-project, we cannot assume, based on existing altmetrics, that Open Access publications are more visible per se than closed access publications.

For this research question, we created a large number of scatter plots (strip plots) and violin plots, all of which are available in the appendix to interested readers.

To illustrate the above statements, Figures 10, 11, and 12 show examples of three different forms of online attention relative to the respective Open Access status. These strip plots show each publication as a data point with a line marking the median. The number of data points varies according to the five Open Access statuses.

Figure 10 shows the news indices relative to the Open Access status. The median for Green OA is noticeable, indicating, as it does, that this status has more publications with a higher normalised news score. However, the Spearman rank correlation coefficient of -0.0002 and the P-value of 0.985 show that there is no demonstrable correlation between Open Access status and the news indices.

Figures 11 and 12 show something similar. These graphs show the normalised Wikipedia and Facebook scores as a function of the five Open Access statuses. There are no significant differences with regard to the median. The statistical tests confirm this, with coefficients of 0.014 for Wikipedia and 0.0174 for Facebook. The P-values do not contradict the null hypothesis either.

Finally, Figure 13 shows a general comparison between closed and Open Access using an example. This graph combines the previously separated Open Access statuses into one.

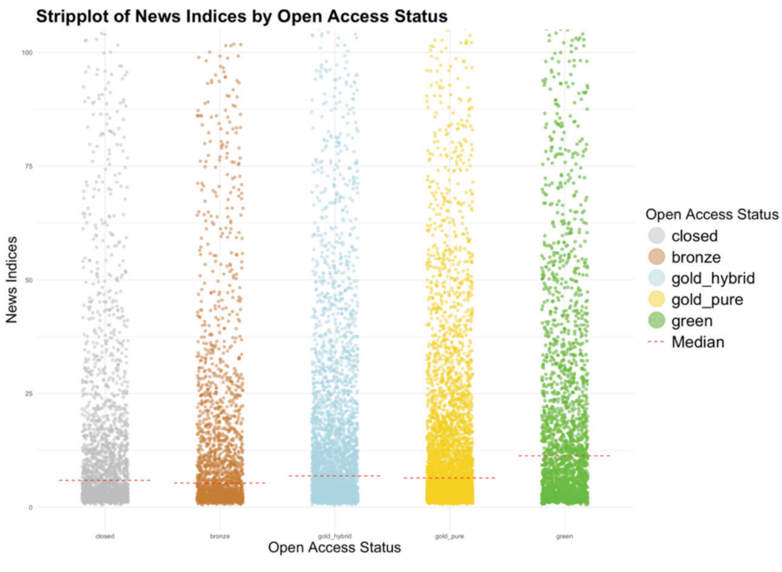

Fig. 10: Strip plot of the news indices relative to Open Access status (Spearman's rank correlation coefficient: -0.0002, P-value: 0.985)

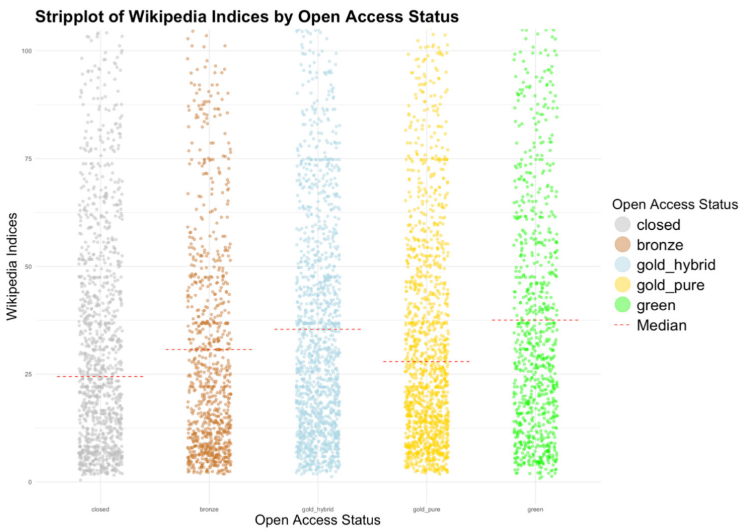

Fig. 11: Strip plot of Wikipedia indices relative to Open Access status (Spearman's rank correlation coefficient: 0.014, P-value: 0.23)

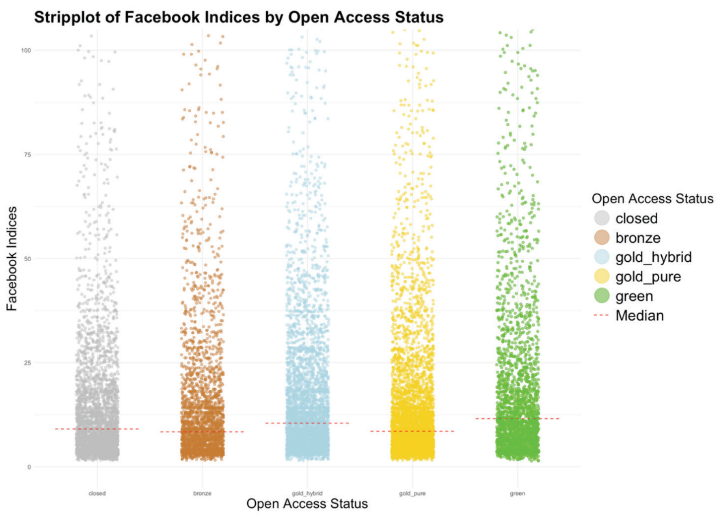

Fig. 12: Strip plot showing the Facebook index relative to Open Access status (Spearman's rank correlation coefficient: 0.0174, P-value: 0.03)

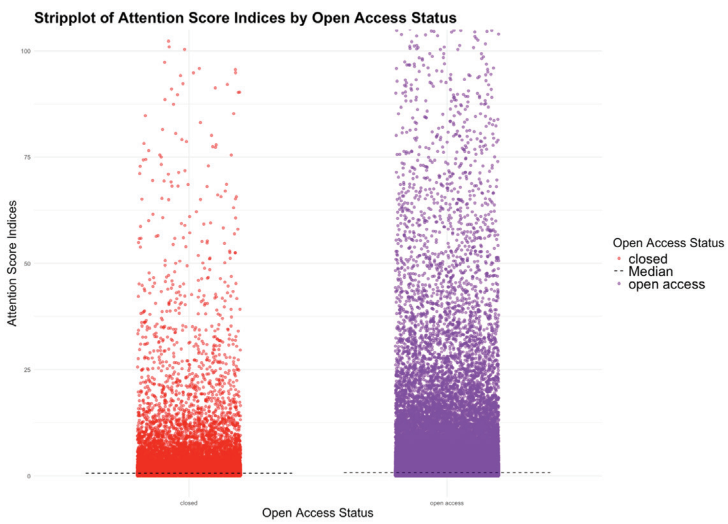

Fig. 13: Scatter plot showing the attention score indices relative to Open Access status

In contrast to the previous figures, the violin plots in Figures 14 and 15 also provide information about the density distribution of the data.

In Figure 14, the "Closed" and "Gold-pure" statuses show a similar distribution, with a density maximum at an index of 0.4-0.5.

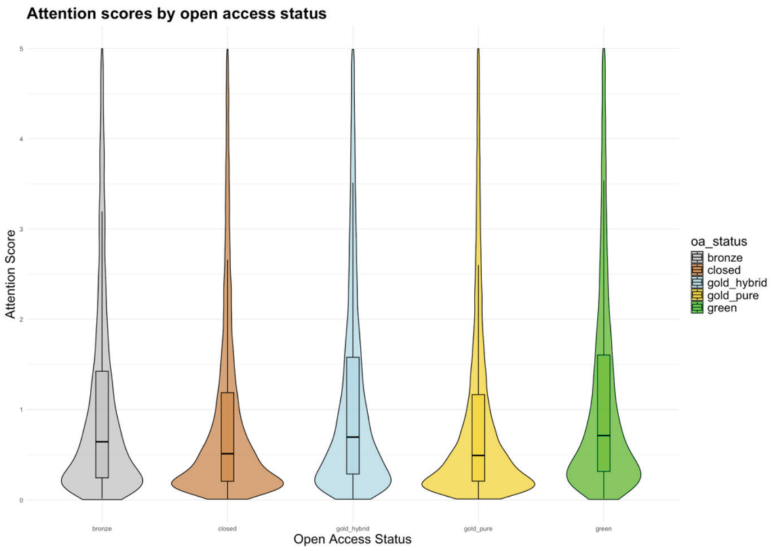

Fig. 14: Violin diagram of Attention Score indices relative to Open Access status

The blogs indices in Figure 15 show a different density distribution: the "green" status has a maximum density in the index range of 5–20 but is less concentrated. In contrast, the "bronze" status reaches its maximum just below an index of 5.

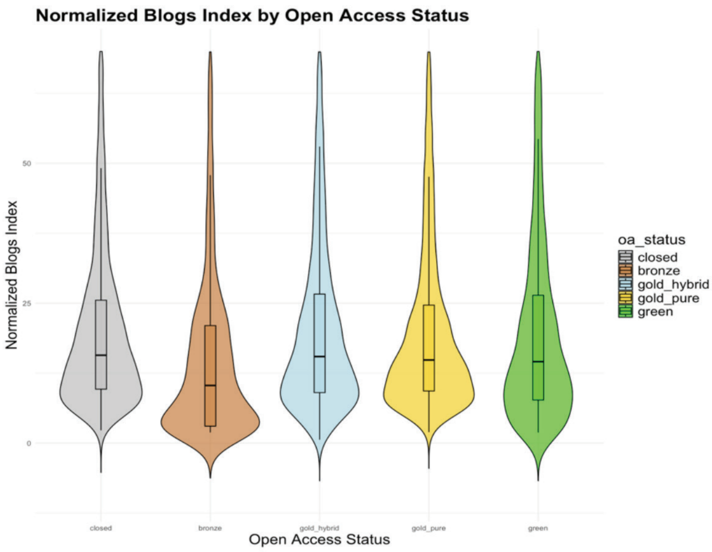

Fig. 15: Violin diagram of the blogs indices relative to Open Access status

Figure 16 shows a direct comparison between Open Access and closed-access publications in a violin diagram. The distribution is similar for both values. Both diagrams also show the highest density just below an index of 0.5, although this is more pronounced in the closed access diagram.

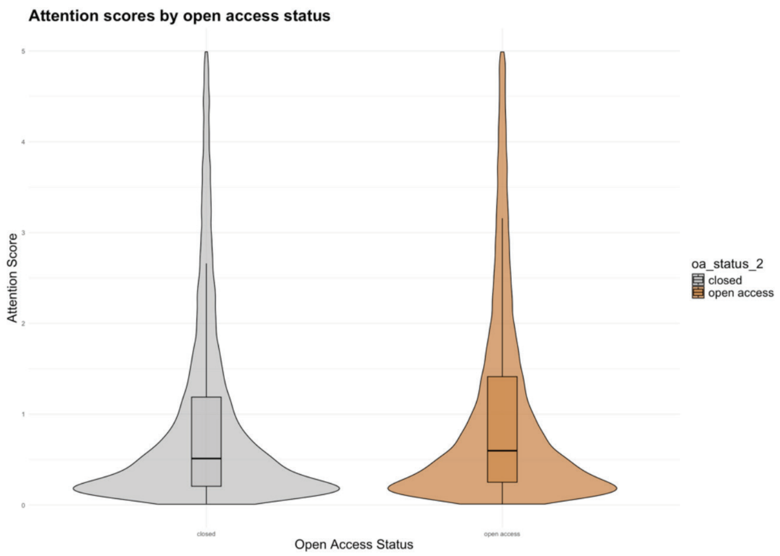

Fig. 16: Violin diagram of Attention Score indices relative to Open Access status

Conclusion

To summarise, no clear correlation can be established between the various forms of online attention and Open Access status. Nevertheless, differences are visible in characteristics such as the median and the density distribution.

Research question 3: Are there any differences by discipline with regard to altmetrics, especially in the various forms of online attention?

Reducing the "Fields of Research" (FoR) to six clusters proved helpful in comparing the different forms of online attention across disciplines. However, it must be emphasised that, despite normalisation, the high standard deviations, the standard error, and the significant differences in the number of publications between disciplines must be taken into account when interpreting the results.

It can be observed that the "Humanities" recorded the highest values for citations and for most online-attention channels (Facebook, Twitter, news, blogs, Wikipedia). The "Social Sciences" achieved the second-best values for citations, blogs and policy and the third-best for Facebook, Twitter, news, and Wikipedia. "Engineering & Technology", on the other hand, was the leader in patents and policy. The results thus show that the numerically underrepresented disciplines in Web of Science and Scopus have the best normalised altmetrics values.

This research question is also extensively visualised through bar charts, which are available in full in the appendix. The following examples serve to illustrate the results described above.

Figure 17 shows the average normalised blog indices for the six FoR clusters. With an average index of just over 10, the "Humanities" are the most successful but also have the longest error bar. "Engineering & Technology" and "Social Sciences", with indices just below and just over 6, are ranked second and third, respectively. The low median in all categories shows that the blog index is low for most of the data.

Reducing the "Fields of Research" (FoR) to six clusters proved helpful in comparing the different forms of online attention across disciplines.

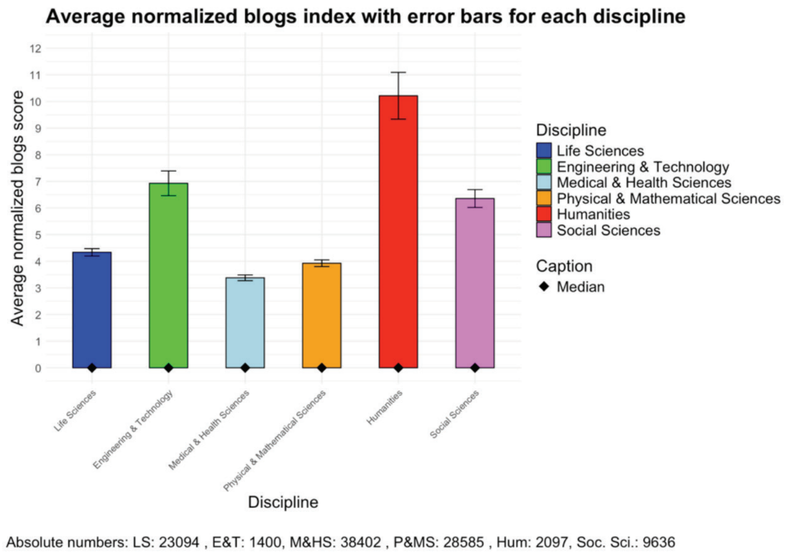

Fig. 17: Bar chart of the "Fields of Research" clusters relative to the normalised blogs indices

However, it must be emphasised that, despite normalisation, the high standard deviations, the standard error, and the significant differences in the number of publications between disciplines must be taken into account when interpreting the results.

It can be observed that the "Humanities" recorded the highest values for citations and for most online-attention channels (Facebook, Twitter, news, blogs, Wikipedia). The "Social Sciences" achieved the second-best values for citations, blogs and policy and the third-best for Facebook, Twitter, news, and Wikipedia. "Engineering & Technology", on the other hand, was the leader in patents and policy. The results thus show that the numerically underrepresented disciplines in Web of Science and Scopus have the best normalised altmetrics values.

This research question is also extensively visualised through bar charts, which are available in full in the appendix. The following examples serve to illustrate the results described above.

Figure 17 shows the average normalised blog indices for the six FoR clusters. With an average index of just over 10, the "Humanities" are the most successful but also have the longest error bar. "Engineering & Technology" and "Social Sciences", with indices just below and just over 6, are ranked second and third, respectively. The low median in all categories shows that the blog index is low for most of the data.

Figure 18 shows the normalised citation indices for the six discipline clusters. The "Humanities" category has the highest value, with an index of around 4.6. The "Social Sciences" and "Medical & Health Sciences" also perform well in terms of citations, with values of just over 3.5 and 3.4, respectively. The error bars are short for all categories. The median differs only slightly between the six clusters, so 50 % of the data is mostly below a similar value.

Figure 19 shows the patents indices. The "Engineering & Technology" discipline cluster has the highest value of slightly over 6, albeit with a comparatively long error bar. With an index of around 3, the "Medical & Health Sciences" cluster ranks second.

The last figure, Figure 20, shows the average Twitter indices. The chart clearly shows that three clusters have high scores for this channel. "Humanities" has an average normalised index of 7.5, "Engineering & Technology" has around 6.7, and "Social Sciences" has 5.5. However, "Humanities" and "Engineering & Technology" also have the longest error bars.

Conclusion

In summary, the normalised data shows that the disciplines less well-represented in Web of Science and Scopus ("Humanities", "Social Sciences" and "Engineering & Technology") have comparatively higher altmetrics values. This knowledge could be used to develop targeted further measures to increase visibility.

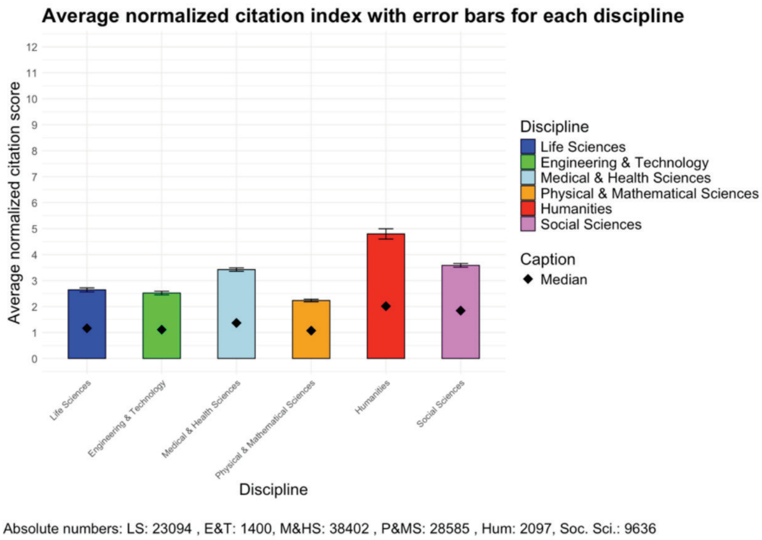

Fig. 18: Bar chart of the "Fields of Research" clusters relative to the normalised citation indices

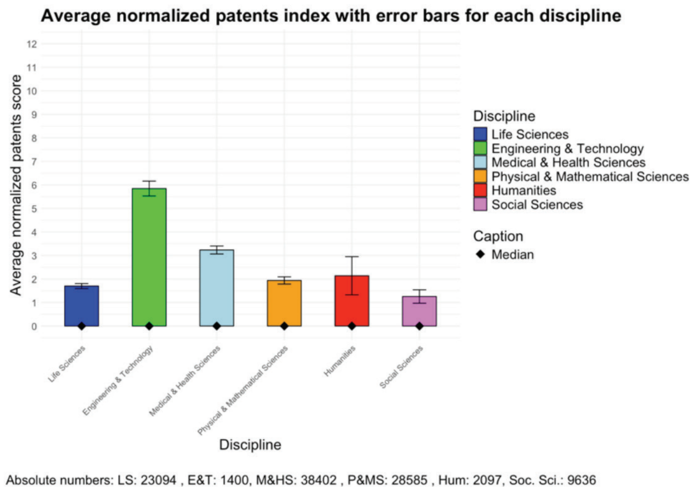

Fig. 19: Bar chart of the "Fields of Research" clusters relative to the normalised patents indices

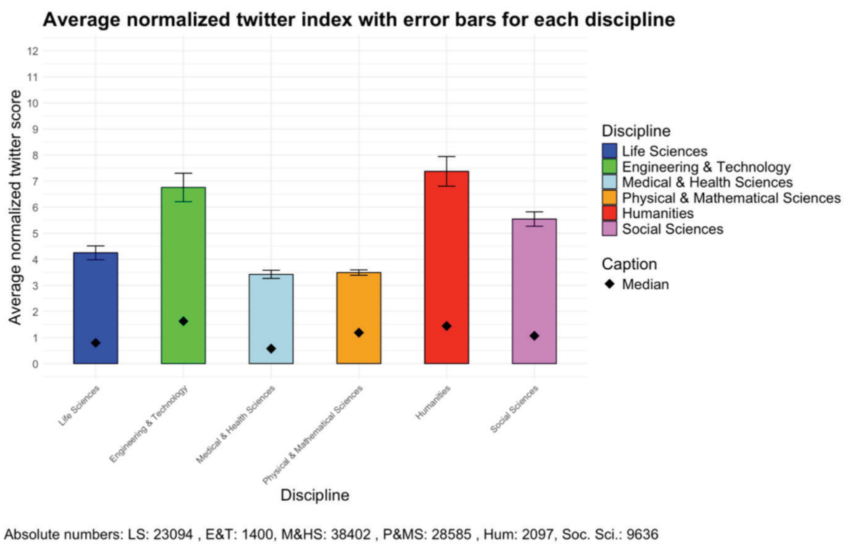

Fig. 20: Bar chart of the "Fields of Research" clusters relative to the normalised Twitter indices

Research question 4: Can correlations be identified between citations and online attention?

Overall, this sub-project could not identify any correlation in most cases and, at most, only a very slight correlation for a few of the cases analysed. The explanations for the scatter plots used can be found in the methodology. However, it is important here to draw attention again explicitly to the use of logarithmic scales.

In the graphs selected below, the "Medical & Health Sciences" cluster is visualised for various online-attention channels because this discipline has the highest correlation with a linear function.

In Figure 21, the normalised citation indices are plotted against those of the blogs as a scatter plot.

The R2 values show that approximately 19 % of the data points coincide with a linear graph, showing a direct correlation. The polynomial graph has a somewhat higher R2 value, which

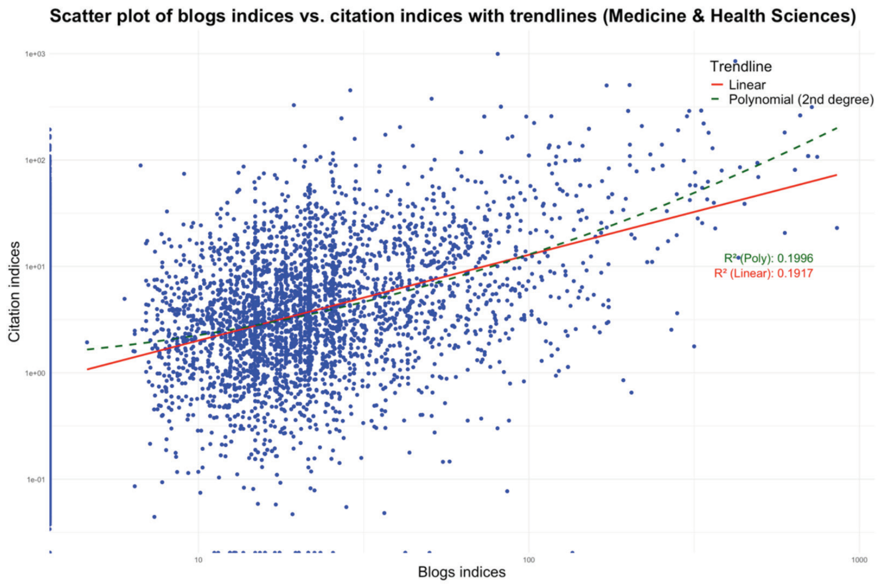

Fig. 21: Scatter plot of the normalised citation indices relative to the normalised blogs indices; trend line drawn for a linear function and 2nd degree polynomial

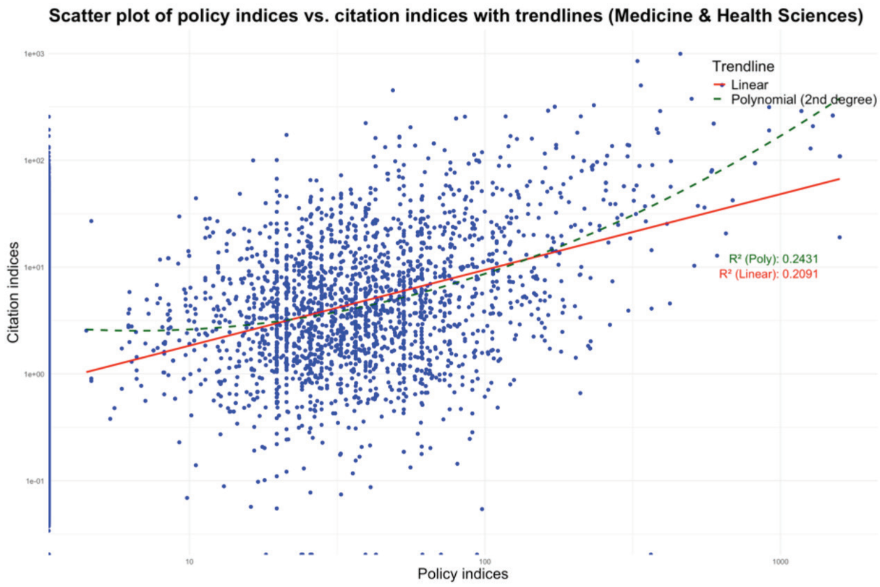

Fig. 22: Scatter plot of normalised citation indices relative to normalised policy indices, trend line drawn for a linear function and 2nd degree polynomial

can be explained by the trend of the data points in the higher x-axis range. However, as the R2 values are <0.3, the correlation found for "Medical & Health Sciences" cannot be considered significant.

Figure 22 shows the relationship between the normalised citation data and policy for "Medical & Health Sciences". The figure shows that 22 % of the data can be explained by a linear trend line and almost 25 % by a parabola. Nevertheless, the correlation found is not significant.

Conclusion

In summary, no significant correlation between the citations and the various forms of online attention can be established through the analyses carried out in our sub-project. The highest rates were found in the "Medical & Health Sciences" cluster between citations and news, and citations and policy indices.

Research question 5: Do the results become more meaningful through the use of normalised altmetrics values?

This research question can be answered in the affirmative. All analyses were carried out for both the absolute and the normalised values. However, only visualisations and findings from normalised data are included in the results section. The corresponding investigations based on the absolute values are available in full in the appendix.

Figure 23 shows the "policy" online attention channel as an example of how normalisation by discipline and publication year affects the results. The direct comparison clearly shows that the "Engineering & Technology" cluster performs poorly when using absolute values to compare disciplines but outperforms the "Social Sciences" with normalised data on publications from the same category. The effect is similar (albeit less pronounced) for the "Humanities" and "Physical & Mathematical Science".

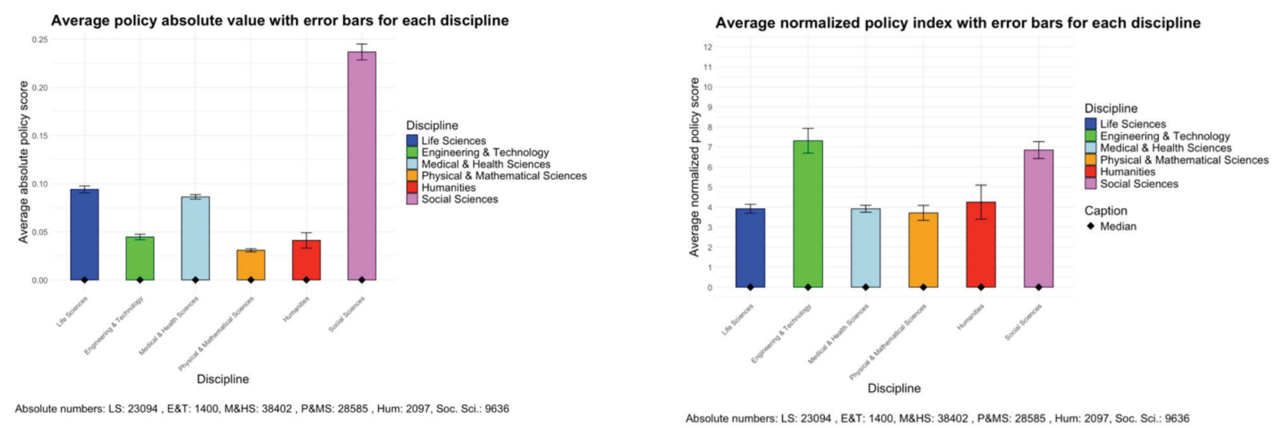

Fig. 23: Bar chart of the "Fields of Research" clusters relative to the absolute (a) and normalised (b) policy indices

Conclusion

In summary, there are significant differences between absolute and normalised altmetrics data. Similar to the citations, the significance of the analyses of the altmetrics also increases when all publications are considered in the context of a defined reference group.

Research question 6: How useful is Crossref Event Data as a free altmetrics provider?

Compared to Altmetric, Crossref Event Data offers some advantages but also disadvantages. One positive aspect is that Crossref Event Data is completely "open" in that all data is made available via an API, free of charge, machine-readable and provided with a CC licence.

However, the data captured is merely listed and neither quantitatively weighted nor aggregated into a specific "score", as Altmetric does. The information is also not displayed on a dashboard or visualised in any other form (although you could do this yourself with the appropriate programming skills).

Compared to Altmetric, there are also at least two other disadvantages. Firstly, Crossref Event Data looks at far fewer news sources than Altmetric, meaning that many links to scientific publications in the mass media are overlooked. Secondly, Crossref Event Data has been missing information about mentions of scientific publications on Twitter ("X" as of July 2023) since February 2023, as the previously free API access to Twitter/X was switched off that month (Rittman, 2023).

Conclusion

Due to the limitations described above, Crossref Event Data cannot be seen as a possible replacement product for Altmetric, but it may be seen as a complementary product. Not all institutions can/want to pay for an Altmetric licence, and Crossref Event Data offers the opportunity to carry out altmetrics analyses on a limited scale for free.

Research question 7: How useful is Crossref Event Data as a free altmetrics provider?

At the beginning of February 2024, a spot check was carried out in Altmetric on a list of all DOIs assigned to date by the DOI Service Austria. The result was sobering: a comparison with the Altmetric data revealed only 475 matches with the 132,458 DOIs submitted. Only 177 of those publications had an Altmetric Attention Score, and only 15 had a score above 10.

After careful consideration, it was therefore decided not to use Altmetric for further data enrichment due to the low number of matches.

Conclusion

Due to the limitations described above, Crossref Event Data cannot be seen as a possible replacement product for Altmetric, but it may be seen as a complementary product. Not all institutions can/want to pay for an Altmetric licence, and Crossref Event Data offers the opportunity to carry out altmetrics analyses on a limited scale for free.

VI. Summary of the results

Research question 1: Ascertaining the absolute and percentage shares of records with altmetrics in comparison to the total data volume (incl. breakdowns by publication year and discipline)

Publication output increased continuously from 2015-2021. The same trend can be seen in Altmetric's successful recognition of publications (from 67% to up to 92%) and enrichment of publications (from 34% to 48%).

This also applies to the analysis by discipline, where, as expected, the underrepresented disciplines in the data set correspondingly have less altmetrics data.

Table 5: Overview of the percentage of enriched publications per FoR discipline cluster

FoR discipline cluster	Altmetrics data in per cent
„Medical and Health Sciences"	34 %
„Physical and Mathematical Sciences"	25 %
„Life Sciences"	20 %
„Engineering & Technology"	10 %
„Social Sciences"	9 %
„Humanities"	2 %

Research question 2: Does Open Access correlate with increased online attention in altmetrics (and thus increased visibility)? Can differences be identified between the various Open Access statuses in relation to altmetrics?

No clear correlation can be established between the various forms of online attention and Open Access status in the available publication dataset. Spearman's correlation coefficients and the p-values are both below the significance threshold in all of the cases analysed.

Nevertheless, differences are detectable in characteristics such as the median and the density distribution.

Based on the altmetrics, we cannot assume that Open Access publications are necessarily more visible than closed-access publications, at least with regard to the data set we analysed.

Research question 3: Are there any differences between disciplines with regard to altmetrics, especially in the various forms of online attention?

Using normalised data shows that the disciplines less well represented in Web of Science and Scopus in particular ("Humanities", "Social Sciences", and "Engineering & Technology") have comparatively higher altmetrics values. This also applies to the individual analyses of the respective forms of online attention, where those FoR discipline clusters achieve a higher normalised score. "Medical & Health Sciences" only joins them in the top 3 for citations and patents.

Table 6: Overview of the respective top 3 score positions with regard to the different information channels

Online attention	Discipline	Top 3 score position for normalised altmetrics
Citations	Humanities	1
Citations	Social Sciences	2
Citations	Medical & Health Sciences	3
Blogs	Humanities	1
Blogs	Social Sciences	2
Blogs	Engineering & Technology	3
News	Humanities	1
News	Engineering & Technology	2
News	Social Sciences	3
Wikipedia	Humanities	1
Wikipedia	Engineering & Technology	2
Wikipedia	Social Sciences	3
Policy	Engineering & Technology	1
Policy	Social Sciences	2
Policy	Humanities	3
Patents	Engineering & Technology	1
Patents	Medical & Health Sciences	2
Patents	Humanities	3
Twitter	Humanities	1
Twitter	Engineering & Technology	2
Twitter	Social Sciences	3
Facebook	Humanities	1
Facebook	Engineering & Technology	2
Facebook	Social Sciences	3

Research question 4: Can correlations be identified between citations and online attention?

In summary, no significant correlation can be established between citations and the various forms of online attention for the analyses carried out in our sub-project. The highest values were found in the "Medical & Health Sciences" cluster between citations and news indices and citations and policy indices (R2 = 0.2).

Research question 5: Do the results become more meaningful through the use of normalised altmetrics values?

Clear differences between absolute and normalised altmetrics data were repeatedly evident during the sub-project. Similar to what is seen with citations, normalisation also increases the informative value of the analyses carried out with regard to altmetrics, as all publications are considered in the context of a defined reference group.

Research question 6: How useful is Crossref Event Data as a free altmetrics provider? How expedient are altmetrics analyses based on alternative data sources?

There are some differences between Crossref Event Data and Altmetric, which disqualify it as a possible replacement product. It may, however, be seen as a complementary product, as it is possible to carry out altmetrics analyses to a limited extent.

Although initial experiments with data from an alternative source (instead of Web of Science and Scopus) were disappointing and did not deliver any "positive" results at first glance, they raised numerous questions requiring further investigation.

Research question 7: How can visibility be increased outside of altmetrics?

Altmetrics served as a vehicle for addressing the issue of increasing the visibility of publication output throughout the project. However, they are only one of many options for increasing visibility and quality.

Other options are presented in the recommendations section and, in addition to altmetrics, include practical information on academic databases, journal rankings, journal metrics, peer review, Open Access, affiliations, FIS/CRIS systems, repositories, academic social networks, persistent identifiers (PIDs), EU projects and academic search engine optimisation (ASEO).

VII. Recommendations for researchers on how to use altmetrics to increase visibility

Altmetrics offer a valuable addition to traditional citation metrics by mapping the reach and resonance of academic work in real time and in a variety of digital contexts. They capture data from social media, news portals, academic blogs, Wikipedia, policy documents, patents and numerous online platforms.

Altmetrics paint a more comprehensive picture of the actual impact of scientific achievements. Academics can use them proactively for their publication strategy, and they also provide a valuable complementary approach to evaluating research performance. Crucial to this is that altmetrics, on the one hand, offer a broader perspective on the visibility of research results beyond the "publish or perish" paradigm and, on the other hand, provide evidence that attests to the involvement of a wide variety of target groups. Ideally, this would make it possible to recognise interesting research trends or identify relevant outputs that (could) significantly impact society.

Here are some practical recommendations for academics on using altmetrics effectively in their everyday activities:

Availability of altmetrics

Altmetrics encompass a wide range of online attention, including X-posts (formerly known as tweets), mentions in academic blogs and news articles, likes, shares, and citations on Wikipedia, in policy documents and in patents.

Academics should familiarise themselves with these different communication channels, the opportunities they offer, and the types of information they generate.

Altmetrics are now standard in most online academic journals, CRIS portals like Pure and the multidisciplinary databases Scopus (integrated as PlumX Metrics) and Dimensions. Tools such as the Altmetric Explorer, which requires a licence, and the freely accessible Crossref Event Data also offer ways of analysing altmetrics data.

Another useful tool is the freely accessible browser add-on Altmetric Bookmarklet, which allows users to look up altmetrics data on a specific academic work with a single click.

Research support services at many universities and research institutions offer assistance in science communication and publication strategies and help with questions relating to altmetrics. Academics should make active use of these resources, as well as the associated guidance and training on offer, to develop their careers. Adept use of altmetrics tools and an understanding of how to analyse and interpret altmetrics data are good distinguishing features in an increasingly competitive academic landscape.

Interpretation and use of altmetrics

When interpreting altmetrics, both quantity (the number of mentions) and quality (the source and context of the mentions) must be considered.

With altmetrics, as with citations, the absolute numbers and the associated calculated scores are only of limited significance. Normalisation (when output is viewed in the context of a defined reference group) certainly increases the informative value. This option is currently only offered in rudimentary form in altmetrics tools, but, at least in the Altmetric Explorer, it is due to be added as a standard feature in the foreseeable future.

Attention does NOT automatically indicate a positive effect or high quality. It is essential to look beyond absolute values and instead use altmetrics (together with citation metrics) in a narrative way to emphasise the qualitative characteristics of your research output:

When and where do the discussions take place? Who is involved? What is said? Who uses my output? Who benefits from my research? Has this research output informed or influenced other work in any way? If so, in what way?

Altmetrics, therefore, provide not only quantitative values but also a wealth of qualitative information that academics can use to good effect when writing CVs, applications, project proposals or research reports. The narrative form helps to illustrate broad interest in research output or social relevance, for example, positive mentions in major news portals or in the policy documents of renowned international organisations.

Strategic use of digital multipliers

An active presence on X (Twitter) and Facebook and intentional use of Wikipedia, academic blogs and mass media to disseminate academic content can significantly increase the visibility and reach of scientific achievements. Academics should make sure to present their work in an appealing and comprehensible way to reach a broad audience. In addition, when disseminating your output by these means, it is vital to always provide it with a persistent identifier (ideally a DOI) so that altmetrics providers can find and index the content.

Recommendations for increasing visibility, beyond altmetrics

Well-established qualitative and quantitative approaches

Peer review

Peer review is the central element in the self-regulation of science. It is a procedure for assessing and selecting academic work by scientists – peers from the same subject areas. Reviews are carried out for manuscripts of journal articles or books, as well as research proposals and applications for academic positions and prizes.

Peer review is intended to ensure professional and constructive quality control and assurance in science. Therefore, scientific papers that have undergone peer review are more reliable sources than texts without peer review. When selecting journals or publication platforms, authors should make sure to choose a peer-reviewed or refereed journal.

Databases

Publishing articles in sources that are indexed in the internationally recognised databases Web of Science and Scopus (both subject to licence), leads to increased visibility.

These multidisciplinary databases guarantee high quality standards through the selection criteria for indexing journals, one requirement being peer review, and regular evaluation processes.

Journal rankings (Scientific Journal Rankings SJRs)

Journal rankings represent the significance of a journal within its subject area and the prestige that derives from this.

These rankings can be lists compiled by academic institutions or by member voting. Because these evaluation methods are subject to distortion, many rankings are based on citation analyses.

The best-known annual journal rankings based on citation analyses are the Scimago Journal & Country Rank (freely accessible), which ranks the journals in the Scopus database, and the Journal Citation Reports (subject to licence), which rank the journals in the Web of Science Core Collection.

Journal metrics

Journal metrics are simple and readily available indicators for evaluating and comparing journals within a discipline. They are calculated using the ratio of a journal's citations to its publications.

The best-known metric is Clarivate's Journal Impact Factor, which is comparable to Elsevier's Cite Score. There are also journal metrics that take into account the typical publication and citation patterns in individual subject areas or particular aspects like the prestige of a journal.

Open Access

Open Access makes scientific research results freely accessible to researchers and a broader interested public worldwide. This promotes knowledge exchange worldwide and collaboration across geographical, institutional and disciplinary boundaries. Publishing in open access significantly increases the visibility of research achievements, whether through immediate Gold Open Access or subsequent Green Open Access publication.

Many research institutions have various Open Access agreements with publishers and Open Access publication funds in order to support researchers in publishing Gold Open Access. Most publishers have now adopted corresponding policies on Green Open Access, covering self-archiving or provision of so-called secondary publications (in repositories, for example), which are also easy to find with the help of tools such as Sherpa/RoMEO. This is an important measure to increase visibility, especially for publications that may originally only have appeared in print, like articles in anthologies.

Discoverability

Open Access publications are generally much easier to find, as, in the best case, search engines and other text and data mining tools index the full text of a publication and make it searchable.

Repositories

Repositories are servers, usually operated by universities or research institutions, on which scientific materials (publications, digitised objects/resources, research data, software, metadata, etc.) are archived and usually made freely accessible worldwide without a login barrier.

Standards such as unique identifiers, permanent links and certain long-term archiving formats ensure interoperability, sustainability and reusability.

There are various types of repository:

- Institutional repositories
  contain the output of an institution, such as a university, and can be actively used by all members of the institution.
- Subject-specific repositories
  offer content (publications, research data) on a specific subject area or discipline.
- Open Access repositories
  have freely-available content with an Open Access licence and no login requirements.
- Publication repositories
  contain full texts of publications and their associated metadata.
- Software repositories
  contain software (often open source).

Note

Specialised search engines can carry out comprehensive searches for repositories. Examples include OpenDOAR for Open Access repositories and re3data for research data repositories. The initiatives behind these repositories aim to promote the free movement of information to and within repositories across institutions and countries and to increase their visibility.

Current Research Information Systems (CRIS)

A CRIS (referred to in German as a FIS) is a database in which an institution continuously documents its research activities such as publications, (externally funded) projects and other output.

They are based on software that is either open-source, commercial or developed by the institution itself.

As a university bibliography, a CRIS provides an overview of an institution's academic achievements and supplies data for reporting and analysis purposes (intellectual capital statements, evaluations etc.).

An institution's own portal based on such a database helps to make both individual scientists and the entire institution more visible to the wider public through corresponding indexing in leading search engines. A CRIS can also serve as a data source for related websites (institute websites, for example) and other systems and platforms (like ORCID).

> **Note**
>
> A CRIS usually contains metadata that describes the existing content in a structured way. It can also be used as a kind of repository to store full texts and research data. However, a CRIS is not designed for long-term archiving.

Unambigious Affiliations

In the context of research, the term "affiliation" refers to researchers belonging to one or more research organisations.

Correctly specifying an affiliation ensures that academic publications and research achievements are attributed to the right researchers and their institutions and increases their visibility.

Academic institutions implement affiliation policies, setting out guidelines for the standardised specification of affiliations and also containing further recommendations for authors, such as ensuring a uniform spelling of their name and creating identifiers. Persistent identifiers, such as an ORCID iD (Open Researcher and Contributor ID), also enable research results to be correctly attributed to authors and should be stated during the publication process in addition to the name and affiliation.

Affiliation guidelines, which are usually in German and English, describe the procedures for correctly specifying the affiliation and are available on the research institutions' websites.

Determining an academic institution's research output is of great importance for the institution's internal research documentation. Correctly using a standardised spelling of the affiliation helps guarantee the documentation's accuracy.

Note

Standardised naming of the affiliation and unique identification (see Persistent Identifiers) is helpful for easier identification of publications when collating the research achievements of scholars and their institutions in databases, enabling complete citation analyses and evaluations.

Persistent Identifiers

Digital Object Identifier (DOI)

for digital objects ensure that they can be found in the long term even if their storage location changes

Digital Object Identifiers (DOI) have established themselves as the standard for scientific publications and research data, while Handle and Uniform Resolve Names (URN) offer technically equivalent solutions. They ensure that a publication is permanently accessible via a stable link and thus is still findable, even if its location (Uniform Resource Locator [URL]) has changed. Numerous platforms and publishers assign DOIs as standard for publications and datasets. Funding bodies also require their use for the publication of project results. DOIs are also frequently used for bibliometric analyses (publication analyses, altmetrics, etc).

The DOI should be included in every available form of a publication. Including the DOI when distributing an article via social media is particularly important to maximise the publication's visibility.

ORCID iD

The **ORCID iD** has become a standard identifier for researchers.

Accurate identification of researchers is critical, as the correct attribution of scientific publications to their authors is of central importance for researchers' academic careers. It is also vital for the identification of an institution's research output. The attribution of publications to their authors is prone to errors for various reasons, such as name similarities and ambiguity (especially in the case of common names), name changes (e.g. through marriage), name variations, different spellings (such as Meyer, Myer, Meyr, Meier, Meijer, Mayer, Maier, Mayr, Mair), or foreign language characters (such as å, ã, ç, ë, ł, ø, ô, š).

Creating and maintaining an ORCID profile contributes significantly to the correct attribution of publications and the visibility of their authors, as does adding the ORCID iD to every publication and entering it when uploading publications to platforms such as ResearchGate or Open Access repositories.

Research Organization Registry ID (ROR ID)

The ROR ID has become established as a unique identifier for institutions, such as universities and research establishments.

Using an ROR ID makes it easier to correctly identify an affiliation, even if an organisation has changed its name, has the names of benefactors in its title or has merged with or been taken over by another organisation. ROR is licenced under CC0 and used, amongst other things, to register DOIs with DataCite and Crossref.

If you are an author and need to indicate your affiliation, you can enter the ROR ID. Funding organisations, such as the FWF, can also be clearly identified by a ROR ID. The ID number can be quoted in the acknowledgements of third-party funded publications. Funding organisations can also use the Crossref Funder ID or the Crossref Grant ID.

Fig. 24: Erat, V., Fürst, E., & Puttinger, J. (2022). *RIS Synergy: Persistent Identifiers*. Zenodo. https://doi.org/10.5281/zenodo.7023003 (CC BY 4.0)

Academic Search Engine Optimization (ASEO)

Search engine optimisation (SEO) is a strategy used in online marketing to improve the findability of websites and documents by search engines. Commercial websites have been using SEO on a large scale for years. Academic search engine optimisation (ASEO) relates specifically to scholarly texts. It aims to provide researchers with the best possible support in finding results relevant to their search queries and help authors achieve a better ranking for their publications in search engines and databases. The elements to be "optimised" include the wording of the title and abstract, the choice of keywords and the provision of comprehensive metadata.

Title optimisation

DOs	DON'Ts
1. Use meaningful titles Make sure to include the most important terms and the conclusion of your research.	1. Avoid creative main titles Save catchphrases, funny remarks and quotes for the subtitle.
2. Important terms up front Position the most important words at the beginning of your main title.	2. Avoid special characters Hyphens, suspended hyphens, asterisks, slashes and the like impair the search function.
3. Think in search terms Would you use the terms in the title to search for your article?	3. Don’t use abbreviations If you use abbreviations, clarify them in the subtitle or abstract.
4. Make it succinct Short titles are easier to identify and more likely to be cited. Consider the display on mobile devices.	4. Don’t exaggerate Follow the guidelines for good academic practice, and don’t overstate your results.

Keywords and abstract optimisation

Keywords	Abstracts
1. Use thesauri MeSH, EMTREE or other subject-specific thesauri can help you choose a keyword.	1. Short sentences Write clearly, precisely and succinctly. Don't use overly flamboyant language.
2. Narrow vs. broad terms Alternate between specific and broader terms that make it easier to identify the topic.	2. Important terms up front Position the most important words at the beginning of your abstract.
3. Use the singular form When choosing a keyword, use it in its singular, uninflected form.	3. Use synonyms Improve your chances of being found by using a variety of terms.
4. Perspective of a searcher Would you use this keyword if searching yourself?	4. Write informatively State your claims, methodology and results.
5. Indicative terms Give information about the content, not the result.	5. Repeat keywords Repetition is a way to increase a publication's ranking and show the focus of the article.

Dissemination of EU project outputs

The European Commission offers a variety of free services to support your dissemination and exploitation activities:

Open Research Europe platform
An Open Access publication platform for academic papers for beneficiaries of Horizon 2020 and Horizon Europe, including open peer review and revision of articles.

Horizon Results Platform
A platform where you can present your research results, find collaboration opportunities and take inspiration from other people's results.

Horizon Results Booster
Free consultancy services, including a portfolio dissemination and exploitation strategy, business plan development and support for your market launch.

European Standardisation Booster Service for EU Projects
An initiative supported by the HORIZON-WIDERA-2021-ERA-01 European Research Area call for proposals and managed by REA).

Supports Horizon Europe and H2020 projects that contribute to standardisation in Europe and beyond.

Innovation Radar
An initiative that identifies high-potential innovations based on data-driven methodology and supports EU-funded researchers and innovators to reach the market with their innovations.

Source: https://rea.ec.europa.eu/horizon-europe-dissemination-and-exploitation_de

Academic social networks

In recent years, networks similar to social networks have been established for scientists, such as Academia.edu and ResearchGate. These networks offer the opportunity to network, interact with other professionals, and exchange publications if this is legally possible. As they are usually commercially orientated and require registration and the disclosure of personal data, they are not considered Open Access platforms or repositories.

Publications find their way into ASNs through automatic searches and aggregation of scientific literature and its metadata, as well as through co-authorship. This automatically creates online profiles for researchers, which are often incomplete and sometimes redundant as they are not curated. It is worthwhile to search for your own profile in these networks and, if necessary, to "claim" and curate it.

Note

One criticism of these networks is that by using them, you are supplying commercial operators with personal data, which they analyse for advertising purposes. Many of these networks finance themselves through personalised advertising, by advertising job offers and through sponsored posts.

VIII. Outlook

Sub-project 5 has proven to be an important national scheme. Its extensive results provide a solid basis for further activities beyond the project period.

The following measures are currently planned or have already been implemented:

1. Visualisation of the recommendations as a science comic

- The aim is to translate the findings and recommendations into an appealing, innovative and easy-to-understand format.
- The visualisation as a science comic will make our recommendations accessible to a wider audience, especially young scientists and the general public.

2. Archiving

- In accordance with our data management plan, all project data will be stored in the University of Vienna's digital archive PHAIDRA.
- This makes it available to all partner institutions to gain further specific insights and develop measures for their own institutions.

3. Conference papers and publications in scientific journals

- The extensive data material available at the national level will now be placed in an international context for discussion.
- We will continue collaborating with our scientific advisor, Mike Taylor, for this purpose, which will ideally result in further output, such as conference papers and publications in renowned journals.

Long-term perspectives

The project aims to have a lasting effect through the wide dissemination of the results and the use of innovative communication formats, such as science comics. It is intended to make a long-term contribution to expanding the state of knowledge on increasing visibility at an individual, institutional, and national level and to initiate concrete improvements in practice.

IX. Acknowledgements

As leaders of this sub-project, we would like to express our sincere gratitude and appreciation for the successful realisation of our ambitious plans within the national AT2OA2 project.

Firstly, we would like to thank the entire project management team, especially Brigitte Kromp, whose visionary leadership has been instrumental to the success of the project. Her continuous support and tireless commitment have motivated us throughout the various phases of the project.

Special thanks go to the entire project office, namely Laura Still, Melanie Stummvoll, Ursula Ulrych, Lothar Hölbling and Tobias Zarka, whose efficient organisation ensured that the project ran smoothly. The consistently professional communication, coordination and support with administrative tasks were invaluable.

We also express our deep gratitude to all those involved in TP5 who, in addition to their day-to-day work, contributed to meetings and working groups and thus made a decisive contribution to our joint success. Above all, special thanks are due to those who actively contributed to the project report. Their names and contributions are highlighted at the beginning of the report.

Martin Rösel, who was mainly involved in TP2, deserves special mention for converting extensive Excel data into an SQL database for our TP5. This database formed the basis for our data analyses.

Last but not least, we would like to thank our colleagues from Digital Science. Ben McLeish provided us with free access to Altmetric and the use of their data throughout the project, while Mike Taylor and Carlos Areia provided us with regular data enrichment and scientific advice. Without their support and co-operation, this success would not have been possible.

Thank you once again to everyone involved for your exceptional performance and commitment. As always, at the end of the project, we are left with many possibilities and ideas that could not be realised within the timeframe. On that note, we look forward to further possible collaborations and future joint successes.

Christian Gumpenberger and Andreas Ferus

X. References

Altmetric (2023, 29. September). *Altmetric attention score*. Retrieved August 6, 2024, https://help.altmetric.com/support/solutions/articles/6000233311-how-is-the-altmetric-attention-score-calculated-

Austrian Transition to Open Access (2017, 01. Jänner). *AT2OA – Austrian Transition to Open Access*. https://at2oa.univie.ac.at/home.html

Basson, I., Simard, M. A., Ouangré, Z. A., Sugimoto, C. R., & Larivière, V. (2022). The effect of data sources on the measurement of open access: A comparison of Dimensions and the Web of Science. *PLoS ONE*, *17*(3), e0265545. https://doi.org/10.1371/journal.pone.0265545

Bethesda Statement on Open Access Publishing (2003, 11. April). *Bethesda Statement on Open Access Publishing*. https://dash.harvard.edu/bitstream/handle/1/4725199/Suber_bethesda.htm?

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, *8*(4), 895–903. https://doi.org/10.1016/j.joi.2014.09.005

cOAlition S (2018, 04. September). *Why Plan S*. https://www.coalition-s.org/why-plan-s/

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric Statistics for Non-Statisticians. A Step-by-Step Approach* (2nd Edition). John Wiley & Sons.

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive compari son of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, *66*(10), 2003–2019. https://doi.org/10.1002/asi.23309

Elsevier (2024, 21. August). *SciVal*. Retrieved August 27, 2024, https://www.scival.com

Erdt, M., Nagarajan, A., Sin, S.-C. J., & Theng, Y.-L. (2016). Altmetrics: An analysis of the state-of-the art in measuring research impact on social media. *Scientometrics*, *109*(2), 1117–1166. https://doi.org/10.1007/s11192-016-2077-0

Erat, V., Fürst, E., & Puttinger, J. (2022). *RIS Synergy: Persistent Identifiers*. Zenodo. https://doi.org/10.5281/zenodo.7023003

European Commission (2019, 13. Dezember). *Trends for open access to publications. Trends for open access to publications*. https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/trends-open-access-publications_en

European Commission (o. J.). *Dissemination and exploitation*. https://rea.ec.europa.eu/dissemination-and-exploitation

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. *Scientometrics*, *124*(3), 2519–2549. https://doi.org/10.1007/s11192-020-03564-9

FWF Österreichischer Wissenschaftsfonds (o. J.). *Open-Access-Policy*. https://www.fwf.ac.at/ueber-uns/aufgaben-und-aktivitaeten/open-science/open-access-policy

Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics*, *108*(1), 413–423. https://doi.org/10.1007/s11192-016-1910-9

Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing Social Media Metrics of Scholarly Papers: The Effect of Document Properties and Collaboration Patterns. *PLoS ONE*, *10*(3), e0120495. https://doi.org/10.1371/journal.pone.0120495

Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014). Tweeting Biomedicine: An Analysis of Tweets and Citations in the Biomedical Literature. *Journal of the Association for Information Science and Technology*, *65*(4), 656–669. https://doi.org/10.1002/asi.23101

Healy, K. (2019). *Data Visualization: A Practical Introduction*. Princeton University Press.

Hölbling, L. (2019). Datenerhebung und Analyse des Publikationsoutputs von Forschenden an österreichischen Universitäten und außeruniversitären Forschungseinrichtungen 2015 bis 2017 im Rahmen von AT2OA – Werkstattbericht zu einer bibliometrischen Studie. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*, *72*(1), 50–58. https://doi.org/10.31263/voebm.v72i1.2290

Holmberg, K., Bowman, S., Bowman, T., Didegah, F., & Kortelainen, T. (2019). What Is Societal Impact and Where Do Altmetrics Fit into the Equation? *Journal of Altmetrics*, *2*(1), 1–11. https://doi.org/10.29024/joa.21

Holmberg, K., Hedman, J., Bowman, T. D., Didegah, F., & Laakso, M. (2020). Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics*, *122*(1), 645–659. https://doi.org/10.1007/s11192-019-03301-x

Holmberg, K., & Vainio, J. (2018). Why do some research articles receive more online attention and higher altmetrics? Reasons for online success according to the authors. *Scientometrics*, *116*(1), 435–447. https://doi.org/10.1007/s11192-018-2710-1

Kooperation E-Medien Österreich (o. J.). *Open-Access-Vereinbarungen in Österreich*. Retrieved August 29, 2024, https://www.kemoe.at/open-access/informationen-und-kontakte

Kuipers, J. R., Jones, M., & Thierry, G. (2018). Abstract images and words can convey the same meaning. *Scientific Reports*, *8*(1), 7190. https://doi.org/10.1038/s41598-018-25441-5

Laakso, M., & Björk, B.-C. (2022). Open access journal publishing in the business disciplines: A closer look at the low uptake and discipline-specific considerations. *Journal of Librarianship and Information Science*, *54*(2), 216–229. https://doi.org/10.1177/09610006211006769

Lang, S. (2017, 04. September). *Statistik: Standardabweichung oder Standardfehler nach Gusto?* Forschen-Schreiben-Publizieren. https://forschen-schreiben-publizieren.de/2017/09/sd-oder-sem-im-research-paper/

Max Planck Society (2003, 22. Oktober). *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanties*. https://openaccess.mpg.de/Berlin-Declaration

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology*, *66*(9), 1832–1846. https://doi.org/10.1002/asi.23286

Mohammadi, E., Thelwall, M., Kwasny, M., & Holmes, K. L. (2018). Academic information on Twitter: A user survey. *PLoS ONE*, *13*(5), e0197265. https://doi.org/10.1371/journal.pone.0197265

Nuzzo, R. L. (2019). Histograms: A useful data analysis visualization. *PM&R*, *11*(3), 309–312. https://doi.org/10.1002/pmrj.12145

Open Science Austria (2022, 13. Juni). *Open Science Austria. Die Open Science Plattform der Österreichischen Universitätskonferenz*. https://www.osa-openscienceaustria.at/

Open Science Policy Austria (2022, 23. Februar). *Österreichische Policy zu Open Science und der European Open Science Cloud*. https://www.bmbwf.gv.at/Themen/HS-Uni/Hochschulgovernance/Leitthemen/Digitalisierung/Open-Science/Open-Science-Policy-Austria.html

Open Society Foundations (2002, 14. Februar). *Read the Declaration. Budapest Open Access Initiative*. https://www.budapestopenaccessinitiative.org/read

Ortega, J. L. (2018). The life cycle of altmetric impact: A longitudinal study of six metrics from PlumX. *Journal of Informetrics*, *12*(3), 579–589. https://doi.org/10.1016/j.joi.2018.06.001

Ortmann, M. (2022, 05. Februar). *Statistik für Mediziner. So erstellst du schnell ein Balkendiagramm für Mittelwerte in R.* Dr. Magdalene Ortmann – Statistik für Mediziner. https://ortmann-statistik.de/so-erstellst-du-muehelos-ein-balkendiagramm-fuer-mittelwerte-in-r/

Pierian Training. (2023, 29. April). *Phyton Basics, Tutorials. Understanding the searborn stripplot in Phyton*. https://pieriantraining.com/understanding-the-seaborm-stripplot-in-python/

Piwowar, H. (2013). Value all research products. *Nature*, *493*(7431), 159. https://doi.org/10.1038/493159a

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, *6*, e4375. https://doi.org/10.7717/peerj.4375

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011, 18. November). *altmetrics: a manifesto*. https://digitalcommons.unl.edu/scholcom/185/

Rittman, M. (2023, 08. Februar). *Dropping Twitter as an event source. Crossref Community Forum*. https://community.crossref.org/t/dropping-twitter-as-an-event-source/3315

Schneider, A., Hommel, G. & Blettner, M. (2010). Lineare Regressionsanalyse. Teil 14 der Serie zur Bewertung wissenschaftlicher Publikationen. *Deutsches Ärzteblatt*, *107*(44), 776–782. https://doi.org/10.3238/arztebl.2010.0776

Seaborn (o. D.). *seaborn.stripplot*. https://seaborn.pydata.org/generated/seaborn.stripplot.html

Simard, M. A., Ghiasi, G., Mongeon, P., & Lariviere, V. (2022). National differences in dissemination and use of open access literature. *PLoS ONE*, *17*(8), e0272730. https://doi.org/10.1371/journal.pone.0272730

Stephen, D., & Stahlschmidt, S. (2024). Contrasting cross-correlation: Meta-analyses of the associations between citations and 13 altmetrics, incorporating moderating variables. *Scientometrics*. Online-Vorveröffentlichung. https://doi.org/10.1007/s11192-024-05006-2

Suber, P. (2012). *Open Access*. MIT Press. https://doi.org/10.7551/mitpress/9286.001.0001

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, *68*(9), 2037–2062. https://doi.org/10.1002/asi.23833

Taylor, M. (2023). Slow, slow, quick, quick, slow: five altmetric sources observed over a decade show evolving trends, by research age, attention source maturity and open access status. *Scientometrics*, *128*(4), 2175–2200. https://doi.org/10.1007/S11192-023-04653-1

Taylor, M. (2024). *Evaluating Open Access Advantages for Citations and Altmetrics (2011–21): A Dynamic and Evolving Relationship*. arXiv. https://doi.org/10.48550/arXiv.2406.10535

Thelwall, M. (2017). Three practical field normalised alternative indicator formulae for research evaluation. *Journal of Informetrics*, *11*(1), 128–151. https://doi.org/10.1016/j.joi.2016.12.002

Thelwall, M. (2020). The Pros and Cons of the Use of Altmetrics in Research Assessment. *Scholarly Assessment Reports*, *2*(1), 1–9. https://doi.org/10.29024/sar.10

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS ONE*, *8*(5), e64841. https://doi.org/10.1371/journal.pone.0064841

Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics*, *103*(2), 555–564. https://doi.org/10.1007/s11192-015-1547-0

Weisberg, S. (2013). *Applied Linear Regression* (4th Edition). John Wiley & Sons.

Wengler, G. (o. J.). *Logarithmische Skala*. GeoGebra. Retrieved September 6, 2024, https://www.geogebra.org/m/w2Fj6up9

Wouters, P., & Costas, R. (2012). *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*. SURFfoundation. https://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, *101*(2), 1491–1513. https://doi.org/10.1007/s11192-014-1264-0