



Persistent Identification for Conferences

RESEARCH PAPER

JULIAN FRANKEN

ALIAKSANDR BIRUKOU

KAI ECKERT

WOLFGANG FAHL

CHRISTIAN HAUSCHKE

CHRISTOPH LANGE

**Author affiliations can be found in the back matter of this article*

]u[ubiquity press

ABSTRACT

Persistent identification of entities plays a major role in the progress of digitization of many fields. In the scholarly publishing realm there are already persistent identifiers (PID) for papers (DOI), people (ORCID), organisation (GRID, ROR), books (ISBN) but there is no generally accepted PID system for scholarly events such as conferences or workshops yet. This article describes the relevant use cases that motivate the introduction of persistent identifiers for conferences. The use cases were mainly derived from interviews, discussions with experts and their previous work. As primary stakeholders who are involved in the typical conference event life cycle researchers, conference organizers, and data consumers were identified. The resulting list of use cases illustrates how PIDs for conference events will improve the current situation for these stakeholders and help with problems they are facing today.

CORRESPONDING AUTHOR:

Julian Franken

TIB – Leibniz Information
Centre for Science and
Technology, Hannover, DE
julian.franken@tib.eu

KEYWORDS:

PID; persistent identifier;
conference; academic event;
use case

TO CITE THIS ARTICLE:

Franken, J, Birukou, A, Eckert, K, Fahl, W, Hauschke, C and Lange, C. 2022. Persistent Identification for Conferences. *Data Science Journal*, 21: 11, pp. 1–18. DOI: <https://doi.org/10.5334/dsj-2022-011>

Conferences are an essential part of the scholarly system. They take on different functions: on the social level, they are an instrument for networking within the scientific community but also between academics and practitioners from (e.g., industry, politics). Furthermore, they allow the fast and efficient communication of research results including the provision of feedback mechanisms. Other documented functions of conferences are: learning effects in regards to presentation skills and academic conduct, room for informal communication, for example, about job opportunities and collaborations, providing an overview of current research and the possibility to increase visibility of researchers (Haus 2018). Depending on the scientific community, conferences also take on other fundamental functions. In particular in computer science, conferences and their proceedings are the preferred venues for publishing original research results (Hansen & Budtz Pedersen 2018; Meho 2019). Their position in citation-based rankings has been studied (Kochetkov, Birukou, & Ermolayeva 2020), confirming the importance of conferences also in engineering and to a lesser extent in mathematics, physics, materials sciences and social sciences.

Although conferences play such an important role in the science system in general, compared to other important entities in Science, information about them is not digitized in a similarly useful fashion. In contrast to formally published journal articles and conference proceedings for the conference event itself, it is not (yet) common practice to issue persistent identifiers (PIDs). Neither are metadata about the conference event captured as widely and systematically as it is done for journal articles or many formally published conference proceedings. Similarly, other important entities of the science system are increasingly receiving their own PIDs such as, for example, researchers (ORCID) or organizations (ROR) (Cousijn et al. 2021). This paper's intention is to give a broad overview about the many use cases of PIDs for conference events and associated metadata. Thereby, we show the usefulness and potential benefit for the whole science system. We argue that conference events should be the entity next in line to receive PIDs more widely.

This paper builds on the experience of the authors of organizing, participating in, and publishing proceedings of various academic conferences and workshops, as well as the participation of the authors in the Crossref/DataCite group on persistent identifiers for conferences (Ackermann et al. 2019). The group gathers relevant scientific communities and publishers, librarians, researchers, and infrastructure providers. Further experience stems from conducting and establishing projects around conference data, namely ConfRef¹ and ConfIDent² – in which interviews with potential stakeholder groups were conducted (for detailed information about the method employed see Section 6), as well as collaborations and data exchange with DBLP³ and other sources for event metadata.

2 STRUCTURE OF THE PAPER

The paper is structured as follows: at first, key terms with regards to conference events will be defined that will be used in this paper (Section 3). Utilizing the previously introduced terms, a general conference event life cycle is described in Section 3.1. Here, all major steps, mostly from the perspective of an event organizer or publisher, that lead up to and follow after a conference event are defined. The event life cycle later serves as the structure and context in which the use cases are framed. After having introduced persistent identifiers (PIDs) and their use in science in Section 4, a short overview is given about the history of PIDs with regard to conference events, referencing previous discussions and initiatives. Tying into this, other related work is summarized in Section 5. Next, the methods by which the use cases were produced are stated in Section 6, followed by the detailed presentation of use cases for PIDs for conference events in Section 7. Concluding, the results and their implications are discussed in Section 8, with an emphasis on ways for widespread adoption of PIDs for conference events and further digitization of conference event information.

1 <http://confref.org>.

2 <https://projects.tib.eu/confident/>.

3 <https://dblp.org/>.

We use the following terminology:

Conference event: The single conference event that was held at a certain place (or virtually) and time, where people met and presented ideas, results or publications. The output of an event may be published in conference proceedings. Example: The 18th International Semantic Web Conference (ISWC) in Auckland, New Zealand, held from October 26–30 2019. With the term conference event we mean to encompass also other types of academic events such as workshops, symposia, colloquia, and the like. Since there is no shared definition between scientific disciplines what exactly constitutes a conference event, we intend to use the term broadly.

Conference series: A series, such as the ‘International Semantic Web Conference’, collates multiple conference events belonging to this series. A conference event can belong to more than one series (co-located or joint conferences); the series name can be informal and can change over time, often when the organizers of the conference events change.

Conference outputs: The output of conferences is mainly published as conference proceedings, journal issues or books⁴ with appropriate identifiers at least on the proceedings level (ISBN, ISSN, DOI). Ideally, single contributions (articles, papers) are also identified using digital object identifiers (DOIs). Next to, or instead of such formal proceedings the conference might produce a number of videos, discussions on social media, blog posts, presentation slides, and so on. Whether archived, linked from the conference website or not, such outputs constitute an important part of the scholarly discussion that took place at the conference.

Conference reference: A conference reference is used to uniquely identify a conference. As long as no PID for conferences is available, the reference often consists of an acronym/year combination such as ‘ISWC 2019’. Consider a citation such as ‘Proceedings of The Semantic Web – ISWC 2019 18th International Semantic Web Conference, Auckland, New Zealand, October 26–30, 2019, Proceedings, Part I’. This reference points to the proceedings of a conference but not necessarily to the conference itself. Nevertheless, such references contain valuable information about the conference, which also makes indirect identification possible as intended by the citation style being used. We therefore identify the following elements commonly used in proceedings references:

Acronym: A short name for the conference often consisting of 3 to 8 upper case letters trying to be unique but actually often being ambiguous. For instance, ISWC may refer to the International Semantic Web Conference or to the International Symposium on Wearable Computing.

Frequency: Annual, biennial, triennial – most events have an annual frequency, and this is mostly not stated explicitly (not stated explicitly in this example).

Event reach: Target reach of the conference such as ‘International’, ‘European’, ‘East Asian’ (International).

Event type: Such as Conference, Workshop, Symposium (Conference).

Year: A two or four-digit reference to the year in which the event took place – not to be confused with the year of publication of the proceedings, which might be different (2019).

Ordinal: Often used to enumerate the conference series instances (18th).

Date: Start date and end date or date range of the conference (October 26–30).

Location: Description of the location of the conference often consisting of country, region, and city – sometimes with details about the exact venue. (Auckland, New Zealand).

Title: The title often contains scope, type, and subject of the conference (International Semantic Web Conference).

Subject: Description of what the conference is about often prefixed with ‘on’ (Semantic Web).

Delimiters: A variety of syntactic delimiters such as blanks, commas, colons, brackets are used depending on the citation style.

⁴ Introduction to Crossref content types <https://www.crossref.org/education/content-registration/content-types-intro/>.

Conference organizer: An entity or entities (person(s)/organization(s)) involved in the organization or planning (cf. Section 3.1) of one or more conference events or conference series. This definition is supposed to include the local chair, who is involved only once, as well as an entity that has been leading the organization of the conference series since its inception (e.g., a steering committee), which could be described more specifically as ‘conference series maintainer’ (Ackermann et al. 2019).

Researcher: Researchers are the group conferences are conducted for primarily. Every person who is participating in the scientific endeavor falls into this group. Usually already early on in their career most researchers make experiences with conferences.

Data consumer/data reuser: Data consumers include everyone concerned with gathering, harvesting, or providing metadata about conferences. To name a few examples, those include: Publishers, indexing services, current research information systems (CRISs), funding organizations, and universities. These data consumers already handle a lot of PIDs and metadata about other entities of the science system like publications (e.g., DOIs), researchers (e.g., ORCIDs) or institutions (e.g., RORs). They usually benefit from data that conforms to the FAIR principles, which state that data should be findable, accessible, interoperable and reusable (Wilkinson et al. 2016).

3.1 EVENT LIFE CYCLE

For the discussion of persistent identifiers for conferences, a solid understanding of the life cycle of such a *conference event* is crucial. It is particularly important to understand at which points identifiers can reasonably be generated and at which points they could be used, which will be further explored in the use cases. In the following, we describe such a life cycle by listing its main activities and events using the Unified Modeling Language state diagram (Object Management Group 2017) semantics (see [Figure 1](#)). The rounded rectangle nodes in the diagram depict the state (phase of activities) of the scholarly conference while the arrow edges show the transitions between these states. A typical life cycle⁵ for an academic event includes the following steps:

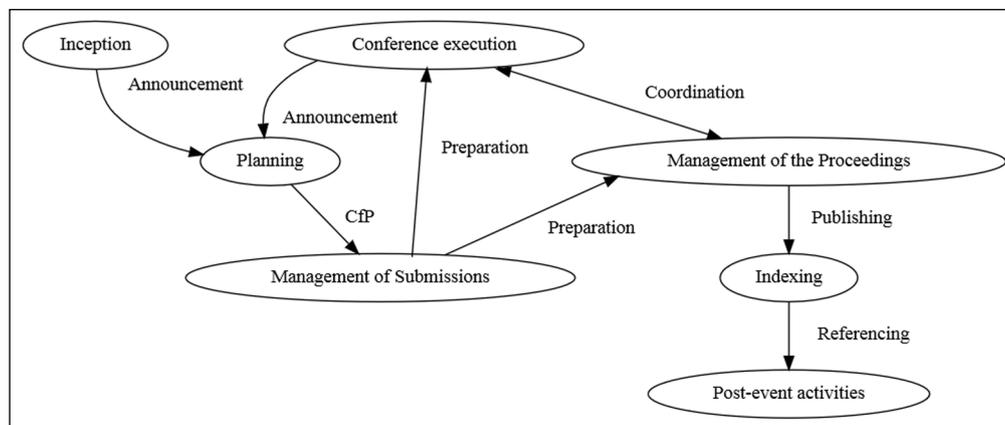


Figure 1 Scholarly Conference Event Life Cycle.

- *Inception of the conference series.* It is rarely conscious, as most of the long-running events typically were started from a one-time workshop on a topic, which then was repeated, and repeated again, thus resulting in a series of events. The conference event (series) is created by a group of people who act as the *conference organizer* (e.g., the *steering committee*). This activity ends with an *announcement* of the next event. This could be done at the closing session of the previous conference event or on the series’ website or via mailing lists in case of the first conference event.
- *Planning of a specific conference event.* At this stage, the steering committee decides on the venue and appoints the event chairs, program committee (PC) chairs, who define the topics, PC members, and the deadlines for paper and review submission for the specific

⁵ Note that the comprehensive life cycle would include also people (authors, PC chairs, and members) and topics covered by the event. Describing those are out of scope for this paper, but one can find such definitions in the SPAR ontologies (Peroni & Shotton 2018).

event. They also decide on the process of handling submission of the papers, that is, the type of peer review and the conference management system. If the event will have formal or informal proceedings, the publisher is selected also at this stage. Anecdotally, the publisher is often selected based on previous experiences of the organizers, their perceived prestige, and the indexing services proceedings will be submitted to. It is rare that this decision is done based on the careful evaluation of the alternatives, such as w.r.t. value for money. This activity ends with issuing a *call for papers (CfP)*, which announces the topics, program committee members, deadlines, peer review process of the *conference events*. CfPs are generally done digitally nowadays, for example, by sending out an e-mail to mailing lists, by using a CfP system such as WikiCFP or EasyChair or by adding a CfP link to the event's home page.

- *Management of submissions*. This includes the submission of contributions as abstracts or full papers and some kind of their selection. In computer science, electrical engineering and some related disciplines the submissions are often reviewed by several PC members or external reviewers and then the PC selects the papers to present at the conference. Depending on the disciplines and event format the selection can range from light (removing obviously bogus contributions) to highly selective, that is, with a 10–15% acceptance rate. This activity produces a *conference schedule*, which outlines how and when the *conference event* takes place, who participates, and so on. If the proceedings are already published at the conference (pre-proceedings), the creation of the proceedings also starts here. Some conferences publish only papers for selected presentations or have another selection process after the conference (post-proceedings).
- *Conference execution* is the next activity, during which participants enjoy the keynotes, plenary, talks, poster and demo sessions where authors present their work. The activity leads to the end of the conference and, often, the *announcement* of the next event, see above.
- *Management of the conference outputs*. This activity involves the dissemination of the scientific outputs of the conference and depends on the discipline. For instance, in the computer science community, conference proceedings contain peer reviewed formal publications that count towards one's publication record. They include peer-reviewed papers presented at the conference and can be published as conference proceedings, books, or journal issues. The publication happens before (pre-proceedings) or after (post-proceedings) the event. Traditionally, the proceedings are the citable results of a conference.
- *Indexing*. If the proceedings of the event constitute a formal archival publication, they would normally be included in a number of abstracting and indexing services, such as DBLP, Google Scholar, Microsoft Academic Search, Scopus, Web of Science, Dimensions, ConfRef. The metadata about the conference proceedings is submitted to (or automatically harvested by) such services. The indexing service makes a decision on indexing specific proceedings. The decision can be based on indexing proceedings of previous conferences in the series, on indexing the publication outlet (i.e., the journal or book series such as LNCS or [CEUR-WS.org](http://www.ceur-ws.org)), or on the quality of the standalone proceedings.
- *Post-event activities* include the time after the event when conference proceedings became part of the scholarly record and accumulate metrics related to their usage, citations, altmetrics (alternative metrics). Such metrics are of interest to the conference event authors, participants and organizers and can be used for research evaluation in some countries or organizations. They are also an important part of the community culture, for example, for determining the influential papers or conference events decades later.⁶

4 PERSISTENT IDENTIFIERS

An identifier is a unique string of characters that is assigned to anything, for example, a concept, or any kind of physical or digital object. Identifiers for scholarly metadata have been available for a long time. An early and popular example is the above-mentioned International Standard

⁶ For instance the SWSA Ten Year Award, <http://swsa.semanticweb.org/content/swsa-ten-year-award>.

Book Number (ISBN), which identifies a specific manifestation of a book. Some identifiers transport semantics, but this is not necessary for them to work. Assigning IDs to objects of any kind is a way of clearly distinguishing between these objects and specifically naming or addressing each one. An object with a certain ID is exactly this object and no other. Confusion and ambiguity are thus a thing of the past – an object with an ID is distinct from others (Lynch 2005).

With the Web and especially with the advent of the Semantic Web, resolvable IDs became widespread, which not only identify things but also link to their digital representation. A key success factor is the persistence of an identifier that serves as a reference pointer for a long period of time. Amazon's success in the book market was boosted by the fact that Amazon is using the ISBN of a book to create a persistent identifier for its offering. The first book ever sold on Amazon in July 1995, 'Fluid Concepts and Creative Analogies: Computer Models Of The Fundamental Mechanisms Of Thought' by Douglas Hofstadter is still available at <https://www.amazon.com/dp/0465024750> 26 years later. Even now books and other products still have the URL format /dp/{product-PID}.

Persistence as a property of identifiers became more important as the consequences of 'link rot' became clearer (Klump & Huber 2017; Lawrence et al. 2001). This means that web resources are no longer accessible because, for example, they were moved to a different domain. An important feature of web-based PIDs is that they are resolvable. This means that their technical function is to redirect to another, the actual resource. If PIDs are used correctly, moving a resource from one domain to another is not a problem that leads to the destruction of references. If the PID of the resource is redirected to the new location, all references are still valid and functional.

Persistent identifiers for entities in the science system. There are several new or already established initiatives for persistent identification of a range of scholarly activities and outputs other than conferences, some of which are already well established and some of which are still less 'mature' (Cousijn et al. 2021). The most widespread is probably the digital object identifier (DOI), which serves as a PID for various types of research output. While the DOI provider Crossref assigns DOIs primarily for publications such as journal articles, conference papers or book chapters, Datacite primarily targets DOIs for emerging research outputs such as research data and software (Paskin 2009). DataCite and Crossref are both DOI registration agencies. It is also becoming more and more common practice to use PIDs to identify other important entities or objects.⁷ ORCIDs are PIDs for authors; they have been assigned since 2012. Their main purpose is the unique identification of authors, especially in the context of scientific publications (L. L. Haak et al. 2012). A rather new PID is the ROR ID, being issued by the Research Organization Registry. Its purpose is to serve as a reference to research institutions (Lammey 2020). PIDs for physical research-related objects are being developed in recent years, with the Research Resource ID (RRID) (Bandrowski et al. 2015) and the PID for Instruments (PIDINST)⁸ being two recent examples.

Persistent identifiers for conferences. The first trace of conference identifiers we managed to find through numerous discussions within the conference PIDs community goes back to the 1970s and the use of conferences in InSPIRE, a leading information platform for high-energy physics (HEP) literature. DBLP, a major computer science bibliography website, was launched in 1993 and included conference series (Ley 2009). Springer launched a pilot on making structured metadata on conference proceedings available freely as linked open data (LOD) in 2015 in collaboration with the University of Mannheim within the LOD2 project (Bryl et al. 2014). The initial release included metadata of roughly 10,000 conferences published in the five-book series publishing proceedings in computer science. Since then, the concept of conference PIDs has been discussed with various researchers, publishers, abstracting, and indexing services and led to the launch of the Crossref/DataCite group on PIDs for conferences,⁹

⁷ The FREYA project even recommends that every entity of the science system receives a PID (Cousijn et al. 2021, p.6).

⁸ <https://www.rd-alliance.org/group/persistent-identification-instruments/case-statement/persistent-identification-instruments>.

⁹ <https://www.crossref.org/working-groups/conferences-projects/>.

and the technical group for implementing conference PIDs¹⁰ while the Springer pilot was used as basis for ConfRef.¹¹

Identifiers used in popular conference-related web platforms. Since there is no standardized conference PID system in place, different methods of identification have been applied by conference-related web platforms. A simple approach, for example, is to increment a number to be used as an identifier. **Table 1** shows the style of identifiers being used by popular conference-related web platforms. Some platforms try to use the conference or series acronym as an identifier but then have to disambiguate when the acronyms are not unique.

NAME	IDENTIFIER	ON DATE	#CONFERENCES	EXAMPLE ISWC 2008	#SERIES	EXAMPLE ISWC
DBLP (Ley 2009)	mostly acronym ^a	2021-12	48517	semweb/iswc2008	5308	semweb
ConfRef	alphanumeric	2021-12	37945	semweb2008	4857	semweb
OpenResearch	acronym	2021-12	9678 ^b	ISWC 2008	1106	ISWC
Microsoft Academic (Färber 2019)	counter	2020-05	~50000 ^c	^d	4467	1155608529 ^e
GND	counter	2021-12	736899 ^f	1168594898	122825	1092290281
WikiCFP	counter	2021-12	89874	1974	6019	1769
Wikidata	counter	2022-02	7470	Q48026643	4232	Q6053150

5 RELATED WORK

Specifying the benefits and use cases of PIDs has been done in the past for several entities of the science system. Regarding ORCID IDs, several institutions and projects contributed to a collection of use cases by formulating detailed descriptions of how they integrated ORCID IDs into their systems.¹² Some commonalities between all individually framed use cases were identified that can be understood as ‘core’ use cases for the integration of ORCID IDs. Those were ‘creation and capture of ORCID IDs for individuals’, authentication in general and for users that previously lacked an identifier, linking ORCID IDs to other or local identifiers and ‘the improvement of attribution’ (Brown, Wilmers, & L. Haak 2015, p. 7). Stocker et al. also involved institutions and projects that each individually worked out use cases for PIDs for instruments (PIDINST) that were specific to them and their needs. Based on these use cases the authors specified metadata properties that were common among them and would be required for enabling them (Stocker et al. 2020). For organization identifiers, Demeranville et al. identified several use cases that described solutions to problems concerning the PID infrastructure (during 2016 and before). Instead of generalized user groups, the use cases targeted concrete agents, namely ORCID and DataCite, as their main beneficiaries (Demeranville et al. 2016). Dappert et al. provided several user stories addressing different stakeholders from academia, thereby highlighting the benefits that are to be expected from a mature PID-landscape in general (Dappert et al. 2017).

Use cases for conference events were most prominently first described by the Crossref/DataCite working group on PIDs for Conferences and presented in the FREYA project. ‘Abstracting and indexing systems’, ‘academic organizer’, entities like learned societies or companies that are involved in the organizer, publishers of conference proceedings and ‘data consumer’ were all user groups which are said to benefit from conference event PIDs. Several use cases were found for each of those user groups.¹³ The use cases elaborated by the working group were fundamental to most of the use cases described in this paper. The FREYA project also illustrated with the help of user stories that having a large database (DataCite Commons) with PIDs of

Table 1 Identifiers used by popular conference related web platforms.

^aOccasional deviations are usually due to historical reasons, e.g., series changing their name.

^bCurated state at RWTH Aachen i5 Server.

^cEstimate based on the assumption that the data is mostly imported from dblp.

^dNo distinct record.

^eSite is down as of 2022-02 – OpenAlex is offering the content now (<https://openalex.org/>).

^fThis includes cultural and commercial events as well as events from other domains. The exact number of academic conference events cannot be determined.

¹⁰ <https://www.project-freya.eu/en/blogs/blogs/towards-persistent-identification-of-conferences>.

¹¹ <https://www.confref.org>.

¹² <https://info.orcid.org/member-use-cases/>.

¹³ <https://www.project-freya.eu/en/blogs/blogs/towards-persistent-identification-of-conferences>.

many entities, which are linked with each other, will provide great benefit to a wide range of users (Fenner & Petryszak 2020).

6 USE CASE CONSTRUCTION

The use cases were constructed based on the experiences of the authors, discussions with colleagues and curators and the work done by the Crossref/DataCite working group on PIDs for Conferences and in particular what this group has produced at a workshop in 2019 at CERN. The use cases described here also heavily build on the research that was conducted in the DFG funded ConfIDent project.

In the context of this project, expert interviews (Bogner, Littig, & Menz 2009; Helfferich 2019) were conducted to get a better understanding of potential users for the service to be built and their needs. The supporting interview guidelines were intentionally kept as open as possible since they stood at the beginning of the project and were thus supposed to be of an exploratory nature. Their purpose was to construct and extend theories (Przyborski & Wohlrab-Sahr 2019, p. 116) about

1. the process of how to arrive at the decision to participate in a conference,
2. decision-making criteria concerning the participation,
3. evaluation of conferences, and
4. actions after the conference.

Those were the main topics of the interview guidelines, which were used to interview primarily researchers. Interviewees that were knowingly experienced with the organization of conferences were asked about

1. the type of work needed to organize a conference,
2. how they advertise and make the conference better findable,
3. the evaluation of a conference from the organizer's perspective, and
4. gathering feedback from attendees afterwards.

The sample selection was based on dimensions that were thought to produce the greatest contrast between interviewees' responses (Przyborski & Wohlrab-Sahr 2019, p. 114). The career stage and the scientific field researchers were active in were considered most important. Additionally, attention was paid to include both male and female researchers.

In total 14 researchers were interviewed, of which eight had a computer science background. The other six interviewees' academic background can broadly be defined as 'mobility research'. Mobility research is an interdisciplinary field. The original academic upbringing of those participants ranged from social science, over logistics to engineering and was thus quite diverse and heterogeneous compared to the group of computer scientists. In addition to their role as researchers, five also had direct experience with the organization of a conference and therefore received different questions. Besides those 14 formally conducted interviews several more informal conversations were held with people that can be considered potential 'data consumers' of conference metadata and PIDs. They were asked to describe how academic events play a role in their system and what problems they encounter.

After being transcribed, the interviews were examined (Mayring & Fenzl 2019) for what needs and pains stakeholders have in the current environment. Finally use cases were constructed describing in detail how PIDs and metadata about conference events would lead to a multiplicity of different benefits for various stakeholder groups.

7 RESULTING USE CASES

In this section the use cases for PIDs and metadata about conference events are described. A use case is understood as a brief description of needs or pains stakeholders have and how PIDs and metadata about conferences can help with those needs or alleviate some of the pain associated with for example a task or situation. They are not meant to be a detailed description of the interaction between a system and its users. Rather they are supposed to illustrate and

give an impression of the many different areas and situations where PIDs may provide real benefits. The use cases follow the order laid out by the event life cycle (3.1).

Target groups/stakeholder groups. The people that might benefit from assigning PIDs to events and conference metadata can be roughly categorized into three groups that were introduced in Section 3: researchers, conference organizers and data consumers. The use cases in the following section were mainly constructed around these three broad user groups. Each group is involved with conferences in different ways and consequently has a different perspective on the topic. By clearly distinguishing these three groups we can become more aware of the plurality of needs and pains that are associated with conferences in a large context.

7.1 PLANNING

7.1.1 Advertising an event

Objective: An event organizer wants to increase his reach. He wants to inform more people about his event than by using only his traditional channels.

Organizers have different means of advertising their events. Those are, for example, newsletters, the conference website, learned societies, word of mouth, or online platforms that display calls for papers, such as WikiCFP.¹⁴ Advertising events is important for several reasons. Firstly, their economic viability usually relies on fees paid by participants. Secondly, for conference organizers having a high number of submissions increases the pool to select high-quality contributions from. Having many researchers submit to a conference or register for it also signals interest from the respective community. Thirdly, the number of paper or abstract submissions has an additional significance in some disciplines. For instance, in computer science the acceptance rate (relation between submitted and accepted papers) of a conference event is oftentimes seen as an indicator for the quality of the event and of its published output (a low acceptance rate is usually perceived as indicative of high quality) (Laplante et al. 2009, p. 192).

Now, conferences that are assigned PIDs and are described with adequate metadata can much more easily be listed by indexing systems (see Section 7.4.1). Once conferences have been assigned a PID, they and their associated metadata can be more easily incorporated into other information systems. This PID-related interoperability thus leads to potentially greater visibility and thus triggers positive attention effects for the conference and its organizations.

Academic events that do not intend to attract as many submissions as possible (outside of computer science) and have a very narrow topical focus can be discovered more easily by those few researchers that are looking for a conference with this exact focus but were not yet aware of it. Hence, PIDs for conferences and the consequential indexing of them would help organizers to achieve a better match between their events and participants that are interested in exactly these events' topics.

7.1.2 Track event series history and handling of name changes

Objective: An event organizer wants to find out how old an event series is. For that he needs to know which events belong to the series and when the first event was held.

Organizers of an event oftentimes do not stay the same. It is custom for people in the scientific community to only be periodically responsible for the organization of a conference until somebody else takes over. It is not unheard of that all of the original people that were involved with a conference series at the time of its inception are not around anymore after some time. The likelihood of this to happen increases the older the conference becomes. Certainly, human memory is not the only thing that can be used to reconstruct the history of a conference series. Especially conferences organized by learned societies such as IEEE or ACM or publishing with major publishers can be expected to have some documentation of their history. However, with smaller events that are not under the umbrella of such an organization, websites frequently are only installed to inform about the next upcoming event and do not contain much information about past events. Shortly after the occurrence of the advertised event they are shut down.

¹⁴ <http://www.wikicfp.com/cfp/>.

Publications that emerge from conferences can help, but those need to be tracked down first and are not always easily findable and accessible. Maybe the conference even grew out of a smaller workshop where publishing proceedings was not deemed necessary at first. Of course one could argue that if a reoccurring event changes so considerably over time it is not the same event series anymore. Nevertheless, people are and always will be interested to find out more about their shared history. It constitutes part of their shared identity as a (scientific) community – regardless of whether the definition of an event series applies to a certain selection of events or not.

This matter is complicated even further when conference series change their names over time. Thus, the property by which we normally try to identify a conference series – its name – cannot be trusted completely to be persistent.¹⁵

By persistently identifying a conference series with an unchanging and persistent identifier, these issues would mostly be resolved. As soon as the second iteration of a conference is announced, this should constitute a conference series, which would receive a PID. Linking a conference to the corresponding series by documenting the PID of the corresponding series in its metadata makes remembering this connection or relying on other (less structured) documentation about this link obsolete, provided that this metadata is kept permanently accessible. If the conference series has received a PID, even a name change will not affect the traceability of its history anymore. On the contrary, when every conference is properly linked with its series, people will be able to find out more easily that name changes occurred in the past.

7.2 MANAGEMENT OF SUBMISSIONS

7.2.1 Support decision-making

Objective: Researchers want to have as much information as possible about conferences to make a well-informed decision whether to participate or not.

To illustrate why PIDs and well-structured metadata about conferences greatly support researchers in making a well-informed decision which conference to participate in, it is helpful to remind ourselves what kind of needs researchers have when it comes to attending a conference. Aktas and Demirel group researchers' needs into four categories: 'Academic development',¹⁶ 'networking', 'organizational issues' as well as 'location and social aspects' (Aktas and Demirel 2019, p. 24). Similarly, Hauss describes two of the main benefits that researchers get from participating in conferences as 'learning effects' and 'access to knowledge' (Hauss 2020, p. 5). Now, for a researcher to ex-ante evaluate whether it is likely for an event to fulfill his or her academic development needs, he or she first of all requires information about what the broader subject of the conference is going to be. The more detailed the information about specific topics, the better. The question of who presents or even participates at a conference needs to be answered to give researchers the chance to estimate how likely their networking needs will be fulfilled if they choose to go to that conference. Researchers expect to be able to develop new contacts for academic cooperation, meet important peers in their area of research, gain reputation among peers, or simply socialize with peers (Aktas & Demirel 2019, p. 24; Hauss 2020, p. 5).

Admittedly, a well-maintained conference website should make all this information publicly available. However, storing all this information in a well-structured and machine-readable manner would enable researchers to specifically look for conferences that fit the criteria by which a conference would fulfill their needs, only by specifying a few keywords such as subject or the name of a colleague. Currently, you have to visit each conference website individually to check whether it would match your criteria. Navigating through different websites in search of the (structurally) same information is often tedious and would be unnecessary if indexed

¹⁵ For example the WACV was called 'Workshop on Applications of Computer Vision' until 2013, until it probably had grown enough to justify a renaming to 'Winter conference on Applications of Computer Vision'.

¹⁶ Academic development summarizes needs that revolve around improving one's own expertise and reputation as well as furthering scientific progress in general (Aktas & Demirel 2019, p. 24).

centrally in the same format. Narrowing down a selection of potentially interesting conferences and comparing them with each other would be substantially easier. Not only would it be easier to find the ‘right’ conference, a well-sustained (complete) database would also solve the problem of researchers not taking conferences into consideration although they would fit well, simply because they are unaware they exist. In particular, early career researchers or researchers that are, due to interdisciplinary interests, new to a field do not have the best overview of their fields yet. They run the risk of missing a conference because they never heard of it. A query of conferences that identify with a certain subject would also bring up those that one is not aware of yet.

7.2.2 Help recognize fraudulent conferences

Objective: Researchers want to be able to identify fraudulent or fake conferences.

Even well-established researchers can fall victim or even knowingly participate at fraudulent or questionable conferences.¹⁷ In this context PIDs could be used as some kind of quality sign, only assigning PIDs to trustworthy conference events. However, we have to assume that organizers of fake conferences are able to mint PIDs for their events too. Cases in the past where publishers of predatory journals registered DOIs for their journals^{18,19} prove that this is likely to happen with conference events too. Using an issued PID as a quality indicator therefore is not possible and is at least not what DOIs are intended for.²⁰ Still, having well-structured, easy-to-comprehend metadata about a conference can help to identify fake conferences better. For example, to increase their conference’s prestige, attract more researchers, and appear more trustworthy, organizers of fake conferences regularly advertise accomplished researchers as contributors to the conference without that actually being true. Once it becomes common practice that speakers at a conference get referenced with their ORCID ID in the conference’s metadata, those referenced researchers would get notified by the ORCID services and could reject being associated with this conference. There are numerous other common features of fraudulent conference events (Asadi et al. 2018; Eaton 2018), which could be detected from the conference’s metadata²¹: for example, a suspiciously short time span between paper submission and review or a name that is confusingly similar to a well-established conference series.²² The probably most secure way for researchers to figure out the quality and sincerity of a conference is to review what research was actually presented, that is, to measure the quality of a conference by the quality of the associated research. This is admittedly the most labor-intensive approach, but also the best way to protect from ‘false positives’ that may arise from deriving the trustworthiness of a conference from other ‘softer’ indicators. By having a direct link to any research output that came out of the conference, researchers are spared the oftentimes tedious search for papers, posters, presentation slides or recordings of the actual talks.

7.3 MANAGEMENT OF THE CONFERENCE OUTPUTS

7.3.1 Persistently link event with research outputs

Objective: An indexing system/service (Scopus, WOS, DBLP, etc.) wants to persistently link a conference to its research output. Researchers want to find research output that was presented at a conference.

Different research outputs such as presentation slides, posters, recordings of the presentations, papers, abstracts, and the like may be of interest to researchers after the execution of the

17 <https://www.uni-hannover.de/en/universitaet/aktuelles/online-aktuell/details/news/statement-hochschulleitung-raubverleger-scheinkonferenzen0/>.

18 <https://thegeyser.substack.com/p/crossref-terminated-omics-membership>.

19 <https://www.crossref.org/board-and-governance>.

20 <https://predatory-publishing.com/is-a-doi-an-indicator-of-quality/>.

21 It should be noted that defining what constitutes a fraudulent conference is a complicated endeavor. Similar issues to defining predatory journals probably would arise (Grudniewicz et al. 2019).

22 <https://thinkcheckattend.org/>, https://www.fz-juelich.de/zb/EN/Expertise/open_access/predatory_publishers/predatory_conferences/predatory_conferences_node.html.

conference. Finding this research output can be at least tedious. If the research output was persistently linked to its conference, looking for it would become much more convenient. However, this implies that in the best-case scenario the research output itself receives its own PID.²³ Moreover, researchers might encounter a conference that they were previously unaware of by reading a paper that was published in its proceedings. If the conference were referenced in the metadata of the paper with its PID, inquiring about it would also become more convenient.

To make the search for scientific output of a conference or the conference itself easier when one or the other is known, indexing systems would need to be able to link one with the other. For this to be done automatically and on a wide scale, the scientific output's descriptive metadata needs to contain a PID, which identifies the conference it was presented at. Additionally, the metadata of the conference needs to link all the scientific output that was produced at the conference.

7.4 INDEXING

7.4.1 Support conference indexing

Objective: An indexing service such as (Scopus, WOS, DBLP, etc.) wants to automatically index a lot of conferences at once and have the metadata change in real time when an organizer changes it.

If at some point it becomes common practice to register a DOI for conferences, the availability of well-structured conference metadata will increase significantly. Currently, there are only a handful of sources providing openly accessible metadata about conferences (WikiCFP, GND, etc.). These are rarely as FAIR (findable, accessible, interoperable, reusable) (Wilkinson et al. 2016) as one would like them to be and usually cover only a small portion of all scholarly conferences. In contrast to the current situation, having only a few centralized sources such as Crossref, where organizers register DOIs and enter conference metadata, would make gathering this data as easy for indexing systems as it is now for them to gather metadata about published articles. The effort that is otherwise needed to harvest metadata from a variety of sources would be much higher.

As with DOIs and their corresponding metadata in general, indexing systems would also benefit from the mandate that the institution that minted a DOI is responsible for keeping its metadata up-to-date.²⁴ Metadata about conferences and in particular conference series will have to be amended regularly to stay current because at every stage of the event life cycle information about the conference will change or new information will be available. Keeping track of all the changes themselves would be an unfeasible task for indexing systems.

Another reason why PIDs would make indexing conferences more feasible is that disambiguation is easier. As long as each conference has only one PID assigned, indexing systems can easily make sure not to include any duplicates in their database.

7.5 POST-EVENT STAGE

7.5.1 DOI claiming

Objective: A CRIS operator wants to enter an event and its metadata in his or her system.

Data consumers such as CRIS operators are currently facing difficulties when it comes to integrating information about conferences in their systems. Having no centralized source of conference metadata, they often have to rely on researchers from their institution who participated in the conference to enter information about the conference – most likely by hand. Or administrative personnel has to research relevant information and enter it. This oftentimes results in incomplete and messy data. On top of that, they have to rely primarily on the name of a conference to unambiguously identify it, which again might lead to duplicates in their database.

²³ Besides recommending to 'use PIDs for all entities' the FREYA project also recommends to 'track and record connections between PIDs' (Cousijn et al. 2021, p. 6).

²⁴ <https://support.datacite.org/docs/doi-registration-policy>.

Suppose that DOIs for conferences are being registered with, for example, Crossref en mass, CRIS operators and any other data reusers then have the possibility to access the Crossref API²⁵ and import the desired metadata into his or her system by referring to the conference's DOI. Moreover, redundancy and duplicates can be prevented better because it could be checked whether any of the new additions already exist in the database. Similar to the use case discussed in Section 7.4.1, any changes to the original metadata could be automatically adopted.

7.5.2 Acknowledgment for contributing to a conference

Objective: Researchers want the effort they spend organizing conferences to be more acknowledged by science's reward system.

Acting as a reviewer, setting up a website, coordinating a caterer, planning the schedule and many other activities can all be considered contributions to the organization of a conference. As different as these tasks may seem, they all have one thing in common: They demand time and effort from scientists that they cannot spend on other activities any more. This would not be a problem if the reward system in science was not dominated as much by metrics based on publication and citation (Nassi-Calò 2017; Steele, Butler, & Kingsley 2006). Hence, effort that is not spent on publishing may not further a scientist's career as much. Although activities such as contributing to the organization of an event are appreciated by peers, most resource distribution mechanisms in science tend to not take this type of effort into account.

Looking back at how publication-based metrics in science have become as influential as they are now, one has to wonder whether this development had something to do with the availability and readiness of publication metadata. One reason for the strong influence publication-based metrics have might be that there historically were few other 'outputs' of science, which were as widely and as diligently measured as publications and citations. When in the wake of the 'new public management paradigm' the call for ways to make science more accountable for the tax money it consumes and to measure its productivity better became louder (Bleiklie 2018), the most extensive and accurate data about scientific produce that was available was of bibliographic nature. Using the data that was already generated and curated by librarians and publishers was practical. To suspect a causal relationship between the availability of such metadata and the importance of the work it describes for Science's reward system is not far-fetched. In any case, having this kind of data about conference contributions will make it much easier for institutions to systematically take these kinds of efforts into consideration when distributing resources. A cultural shift towards a more multidimensional reward system in science requires good data about more than just publications.

Not only scientists might have an incentive that their contributions to conferences are better and more systematically recorded. Well-received conferences improve the institutions' reputations that are associated with the scientists that organize them. It is also in the best interest of such institutions to have a good overview of the events that are held in their name. The assumption is that, as soon as conferences can be better identified via PIDs and have 'FAIRer' metadata, many agents from inside science would want to make use of them to convey their excellence.

7.5.3 Monitor conference activity and performance

Objective: Researchers and agents concerned with the evaluation of science want to have a convenient and reliable way to compare conferences with each other in terms of impact and other measurements of performance. This enables organizers, participants, funders and research evaluators to evaluate specific conferences.

With regard to researchers, this use case is similar to the one discussed in Section 7.2.1, only that this use case emphasizes a particular category of conference metadata – those related to the conference's 'scientific quality'. In computer science, two of the most important indicators used to measure scientific quality are the acceptance rate²⁶ and the CORE ranking.²⁷ These two indicators are generally well accepted for measuring a conference's scientific quality

25 <https://www.crossref.org/documentation/retrieve-metadata/rest-api/>.

26 Number of accepted submissions in relation to the total number of submissions.

27 Computing Research and Education Association of Australasia <https://www.core.edu.au/home>.

and its quality or prestige more generally. Other conference attributes from which to infer its quality or prestige are for example the number of attendees, the series' age, who gives keynote speeches, and so on.²⁸ Since one of the main motivations for researchers to participate in conferences is to further their academic development (Aktas & Demirel 2019), they seek out 'high-quality' conferences. Hence, this type of metadata would be used by researchers to make informed decisions about which conferences to attend best. Based on this observation, we can assume that organizers and data consumers also have an interest in making this information easily findable, accessible and reusable. Especially organizers of conferences that are already performing well should have an incentive to have this information unambiguously attributed to their conference.

Indexing systems could enable their users to compare conferences based on these indicators. Not only researchers and organizers (with good indicator values) would benefit from that. Other agents of the science system wanting to produce or consume evaluations (funding agencies, universities, etc.) would benefit as well. As such, the whole science system would receive additional performance indicators by making the scientific quality of an important entity of the science system more measurable. If such metrics are included in the conference metadata and then combined with information about who contributed to the conference (cf. Section 7.5.2), the way researchers and academic institutions are currently evaluated could be improved. Performance measurements would be more diverse and less single-dimensional.

7.5.4 Preserving information

Objective: Researchers want to be able to get information about conferences long after they occurred. Organizers want to document the history of their conference to prove its legitimacy and increase its reputation.

Currently, information about a conference can be found best on the conference's website. Researchers can reasonably expect to find websites of even smaller academic events. As long as researchers know the name of the conference they are looking for, they are usually able to find the respective conference website. This situation, however, changes as time passes. After a conference event is over, advertising it is no longer needed (at least the individual event, advertising the series remains a concern). Hence, a major incentive for organizers to maintain the conference website, that is, to advertise it (cf. Section 7.1.1), disappears. In particular organizers of smaller academic events with a tight budget will think twice whether to keep the website online or not. Information about events that belongs to well-established series admittedly has a lower risk to be lost. Still, event series being discontinued or becoming absorbed by larger series is a real possibility in which cases the information that was previously stored on its website(s) can become either partially or completely lost.

Instead of relying on the fragile existence of event websites for preserving information about conferences, we had better make use of the already existing infrastructure dedicated to preserving and providing metadata about a multitude of different resources. The only requirements of course are: PIDs and well-structured metadata.

7.5.5 Find event via a DOI of a referencing paper

Objective: A researcher wants to find out what the associated conference event of a paper he read is.

One way for researchers to encounter previously unknown conferences is to look for the conference a paper was presented at. Especially in computer science, this is a common way for researchers to get acquainted with the relevant conferences in their field of study. Having at least a rough overview of which conferences exist in their field is a necessary condition for researchers to be able to make a decision about which one to attend (see Section 7.2.1). Currently, researchers usually find more information about the associated conference of a paper by looking at the name of the corresponding proceedings, which normally contains the name and/ or acronym of a conference in its title, and then searching for it with a search engine, resulting in them hopefully accessing the conference website.

²⁸ For an extensive overview about possible quality indicators for conferences, see Vahdati et al. 2020. Hansen and Budtz Pedersen identify areas where conferences have a positive impact (Hansen & Budtz Pedersen 2018).

Having a machine-readable connection between paper, proceedings, and the conference itself in the form of PIDs identifying a conference event in the metadata of a paper and/ or proceedings (see Section 7.3.1) would lead to an easier and more efficient search. From the perspective of an author, referencing conferences would become a more feasible option, which in turn would lead to more machine-readable connections between papers and conferences.

8 DISCUSSION

This article is meant as a first starting point to structurally develop persistent conference identifiers and associated metadata. We highlighted the need for PIDs and metadata about conference events by describing several use cases for them, thus emphasizing the benefits for researchers, event organizers, and data users. The use cases presented here do not claim to be exhaustive, but rather are seen as the most relevant ones in terms of how much they benefit the community once enabled. The results clearly show that a considerable amount of the use cases depend on a large percentage of conference events having a PID assigned and being described by at least basic metadata. To make conference PIDs useful at least the following information should be provided with it: name, name of associated series, place of venue, time, and scientific discipline. The ConfIDent project has suggested the Academic Event Ontology AEON²⁹ and a consequential metadata schema³⁰ for this purpose. Obviously, the richer the metadata is, the more use cases are covered.

However, PIDs need to be minted and metadata needs to be generated first. Looking at the amount of conference events that are held every year it is reasonable to expect that only a joint effort by large parts of the scientific community can increase the amount of PIDs minted and metadata generated. Moreover, information about future conference events is very susceptible to change. For instance, dates and deadlines get moved, keynote speakers or venues change. Because of these challenges that will come with attempting to crowdsource this task, it is important to think of ways to support the scientific community. The conference life cycle offers several opportunities where PIDs could be minted automatically. In future endeavors it will be worthwhile to identify points where the generation of PIDs can be triggered. In particular for the use cases 7.1.1, 7.2.1, and 7.2.2, it is important to create a PID early on in the life cycle.

In addition to the community effort that is needed, technical and organizational infrastructures have to be financed. Establishing a viable business model is an important matter for PID providers like DataCite (Cousijn 2020), ORCID³¹, ROR³², or IGSN (Lehnert et al. 2021). Given the fact that all of these examples at least partly rely on some kind of membership model, founding a PID initiative for conference events on a membership model might be a promising approach as well. Yet, for this approach to be a viable option, PIDs for conferences need to become more popular first. Similar to the IGSN 2040 initiative (ibid), reusing an already existing infrastructure and partnering with DataCite could reduce a lot of the otherwise needed initial investment. The DataCite metadata schema can indeed be reused for describing conference events without needing a lot of far-fetched interpretations.³³ Still, the landing page for the DOI needs to be provided persistently, which is a task the entity creating the DOI usually is responsible for. As described in 7.5.4 conference websites tend to vanish after some time and therefore are less suitable as landing pages.³⁴ Since DataCite already has set up a fee structure for registering DOIs in addition to its membership fees in 2020, some of the costs needed to register a DOI for a conference event could be externalised to the conference event's organizer. Registering a DOI for a conference event and providing a landing page could become a service offered by DataCite member organizations. In contrast to, for example, registering DOIs for data sets, versions, or collections thereof, where the amount of needed DOIs could potentially be quite

29 <https://github.com/tibonto/aeon>.

30 <https://stroemphi.github.io/ConfIDent-schema/>.

31 <https://info.orcid.org/orcid-we-wont-be-sold/>.

32 <https://scholarlykitchen.sspnet.org/2019/12/04/are-you-ready-to-ror-an-inside-look-at-this-new-organization-identifier-registry/>.

33 <https://blogs.tib.eu/wp/tib/2021/05/31/finally-persistently-identifying-academic-events-and-event-series/>.

34 The Confident project aims to provide persistent landing pages for conference event DOIs.

high (Klump, Huber, and Diepenbroek 2016), the average organizer only needs to have one DOI registered per year. Registration costs thus should not become prohibitive.

As was proposed by the Crossref/Datacite conference PID Workinggroup, issuing DOIs using the DataCite schema could be applied in particular to the ‘long tail’ of conference events that do not have formal proceedings. For those that collaborate with a publisher, the publishers would mint DOIs when registering proceedings or reuse already minted ones (in case of several publishers publishing proceedings from the same conference). As most of the publishers of the conference proceedings are already Crossref members, the very same DOI infrastructure can be used to support DOIs for conferences and conference series. Although this last described method is a promising one in terms of governance, reduced financial costs, and convenience for the organizer, minting a DOI when the conference event is already over would be less helpful for some use cases (7.1.1, 7.2.1 and 7.2.2).

All in all, it is obvious that conference PIDs with rich metadata are potentially a strong asset for the scholarly PID graph. Linking events to organizations, projects, conference outputs and persons will lead to a more efficiently and fairly operating science system and ultimately will foster faster scientific progress.

COMPETING INTERESTS

Julian Franken, Christian Hauschke, Wolfgang Fahl and Christoph Lange all work in the DFG funded ConfIDent project (<https://projects.tib.eu/en/confident/>) on the development of a service platform for scientific events. It aims to make the descriptive metadata on conferences and other formats of scientific events permanently accessible in a high quality through automated processes and scientific data curating. The project plans to issue DOIs for conference descriptions.

AUTHOR CONTRIBUTIONS

All authors contributed equally to this work.

AUTHOR AFFILIATIONS

Julian Franken  orcid.org/0000-0003-3999-253X

TIB – Leibniz Information Centre for Science and Technology, Hannover, DE

Aliaksandr Birukou  orcid.org/0000-0002-4925-9131

Springer Nature, Heidelberg, DE; Peoples’ Friendship University of Russia (RUDN University), Moscow, RU

Kai Eckert  orcid.org/0000-0002-5423-561X

Stuttgart Media University, Institute of Applied Artificial Intelligence, DE

Wolfgang Fahl  orcid.org/0000-0002-0821-6995

RWTH Aachen, i5, Aachen, DE

Christian Hauschke  orcid.org/0000-0003-2499-7741

TIB – Leibniz Information Centre for Science and Technology, Hannover, DE

Christoph Lange  orcid.org/0000-0001-9879-3827

RWTH Aachen, i5, Aachen, DE; Fraunhofer FIT, Sankt Augustin, DE

REFERENCES

- Ackermann, MR**, et al. 2019. *Global persistent identifiers for conferences and Crossmark for conference proceedings - publishing in progress*. URL: <https://docs.google.com/document/d/1URIVkUpzcfjSd2YFISr dRIrOyrKSbFfhkdpGPRTAFI/edit#> (visited on 03/18/2020).
- Aktas, G** and **Demirel, B**. 2019. The genuine needs of conference attendees: An analysis by the modern quality function deployment. *International Journal for Quality Research*, 13(1): 13–32. ISSN: 18006450. DOI: <https://doi.org/10.24874/IJQR13.01-02>
- Asadi, A**, et al. 2018. Fake/Bogus conferences: Their features and some subtle ways to differentiate them from real ones. *Science and Engineering Ethics*, 24(2): 779–784. DOI: <https://doi.org/10.1007/s11948-017-9906-2>
- Bandrowski, A**, et al. 2015. The resource identification initiative: A cultural shift in publishing. *F1000Research*, 4: 134. ISSN: 2046-1402. DOI: <https://doi.org/10.12688/f1000research.6555.2>

- Bleiklie, I.** 2018. New Public Management or Neoliberalism, Higher Education. In: *Encyclopedia of International Higher Education Systems and Institutions*, 1–6. Dordrecht: Springer Netherlands. ISBN: 978-94-017-9553-1. DOI: https://doi.org/10.1007/978-94-017-9553-1_143-1
- Bogner, A, Littig, B and Menz, W** (eds.), 2009. *Experteninterviews: Theorien, Methoden, Anwendungsfelder*. 3., grundlegend überarbeitete Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften. ISBN: 9783531162591.
- Brown, J, Wilmers, C and Haak, L.** 2015. *Final Report: Sloan ORCID Adoption and Integration Program 2013–2014*. URL: https://figshare.com/articles/journal_contribution/Final_Report_Sloan_ORCID_Adoption_and_Integration_Program_2013_2014/1290632. DOI: <https://doi.org/10.6084/m9.figshare.1290632.v1>
- Bryl, V,** et al. 2014. What is in the proceedings? Combining publisher's and researcher's perspectives. In: Garcia Castro, A, et al. (eds.), *Proceedings of the 4th Workshop on Semantic Publishing co-located with the 11th Extended Semantic Web Conference (ESWC 2014)*. Anissaras, Greece, May 25th, 2014.
- Cousijn, H.** 2020. *The new DataCite Membership and Fees Model*. DOI: <https://doi.org/10.5438/xn0c-m298>
- Cousijn, H,** et al. 2021. Connected research: The potential of the PID graph. *Patterns* (New York, N.Y.), 2(1): 100180. DOI: <https://doi.org/10.1016/j.patter.2020.100180>
- Dappert, A,** et al. 2017. Connecting the persistent identifier ecosystem: Building the technical and human infrastructure for open research. *Data Science Journal* 16. DOI: <https://doi.org/10.5334/dsj-2017-028>
- Demeranville, T,** et al. 2016. *Organisation Identifiers - Minimum viable product requirements*. URL: https://figshare.com/articles/dataset/Organisation_Identifiers_-_Minimum_viable_product_requirements/3479141. DOI: <https://doi.org/10.6084/m9.figshare.3479141.v1>
- Eaton, SE.** 2018. *Avoiding Predatory Journals and Questionable Conferences: A Resource Guide*. DOI: <https://doi.org/10.11575/PRISM/20>
- Färber, M.** 2019. The Microsoft academic knowledge graph: A linked data source with 8 billion triples of scholarly data. *Lecture Notes in Computer Science*, 113–129. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-030-30796-7_8
- Fenner, M and Petryszak, R** 2020. *FREYA webinar - The PID Graph in practice - Jupyter Notebook demonstration*. DOI: <https://doi.org/10.5281/zenodo.4004426>
- Grudniewicz, A,** et al. 2019. Predatory journals: no definition, no defence. *Nature*, 576(7786): 210–212. DOI: <https://doi.org/10.1038/d41586-019-03759-y>
- Haak, LL,** et al. 2012. ORCID: A system to uniquely identify researchers. *Learned Publishing*, 25(4): 259–264. ISSN: 09531513. DOI: <https://doi.org/10.1087/20120404>
- Hansen, TT and David, BP.** 2018. The impact of academic events—A literature review. *Research Evaluation*, 27(4): 358–366. ISSN: 0958-2029. DOI: <https://doi.org/10.1093/reseval/rvy025>
- Hauss, K.** 2018. *Der Nutzen wissenschaftlicher Konferenzen in der Nachwuchsausbildung*. Wiesbaden: Springer Fachmedien. ISBN: 978-3-658-19625-7. DOI: <https://doi.org/10.1007/978-3-658-19626-4>
- . 2020. What are the social and scientific benefits of participating at academic conferences? Insights from a survey among doctoral students and postdocs in Germany. *Research Evaluation*, 24: 21. ISSN: 0958-2029. DOI: <https://doi.org/10.1093/reseval/rvaa018>
- Helfferich, C.** 2019. Leitfaden- und Experteninterviews. In Baur, N and Blasius, J (eds.), *Handbuch Methoden der empirischen Sozialforschung*, 669–686. Wiesbaden: Springer Fachmedien Wiesbaden, ISBN: 978-3-658-21307-7. DOI: https://doi.org/10.1007/978-3-658-21308-4_44
- Klump, J and Huber, R.** 2017. 20 years of persistent identifiers – Which systems are here to stay? *Data Science Journal*, 16. DOI: <https://doi.org/10.5334/dsj-2017-009>
- Klump, J, Huber, R and Diepenbroek, M.** 2016. DOI for geoscience data - How early practices shape present perceptions. *Earth Science Informatics*, 9(1): 123–136. ISSN: 1865-0473. DOI: <https://doi.org/10.1007/s12145-015-0231-5>
- Kochetkov, D, Birukou, A and Ermolayeva, A.** 2020. *The Importance of Conference Proceedings in Research Evaluation: a Methodology for Assessing Conference Impact*. arXiv: 2010.01540 [cs.DL].
- Lammy, R.** 2020. Solutions for identification problems: A look at the Research Organization Registry. *Science Editing*, 7(1): 65–69. ISSN: 2288-8063. DOI: <https://doi.org/10.6087/kcse.192>
- Laplante, P,** et al. 2009. Quality in Conference Publishing. *IEEE Transactions on Professional Communication*, 52(2): 183–196. ISSN: 0361-1434. DOI: <https://doi.org/10.1109/TPC.2009.2017989>
- Lawrence, S,** et al. 2001. Persistence of Web references in scientific research. *Computer*, 34(3): 26–31. ISSN: 00189162. DOI: <https://doi.org/10.1109/2.901164>
- Lehnert, K,** et al. 2021. *IGSN 2040 Summary Report: Defining the Future of the IGSN as a Global Persistent Identifier for Material Samples*. DOI: <https://doi.org/10.5281/zenodo.5118289>
- Ley, M.** Aug. 2009. DBLP: Some lessons learned. *Proc. VLDB Endow*, 2(2): 1493–1500. ISSN: 2150-8097. URL. DOI: <https://doi.org/10.14778/1687553.1687577>
- Lynch, C.** 2005. Identifiers and their role in networked information applications. *Bulletin of the American Society for Information Science and Technology*, 24(2): 17–20. ISSN: 00954403. DOI: <https://doi.org/10.1002/bult.80>
- Mayring, P and Fenzl, T.** 2019. Qualitative Inhaltsanalyse. In Baur, N and Blasius, J (eds.), *Handbuch*

- Methoden der empirischen Sozialforschung*, 3: 633–648. Wiesbaden: Springer Fachmedien Wiesbaden. ISBN: 978-3-658-21307-7. DOI: https://doi.org/10.1007/978-3-658-21308-4_42
- Meho, LI.** 2019. Using Scopus's CiteScore for assessing the quality of computer science conferences. *Journal of Informetrics*, 13(1): 419–433. DOI: <https://doi.org/10.1016/j.joi.2019.02.006>
- Nassi-Calò, L.** 2017. Evaluation metrics in science: Current status and prospects. *Revista latino-americana de enfermagem*, 25: e2865. DOI: <https://doi.org/10.1590/1518-8345.0000.2865>
- Object Management Group.** 2017. *Unified Modeling Language. Version 2.5.1.* <https://www.omg.org/spec/UML/2.5.1/PDF>.
- Paskin, N.** Dec. 2009. Digital Object Identifier (DOI®) System. *Encyclopedia of Library and Information Sciences*, 1586–1592. Third Edition. CRC Press. DOI: <https://doi.org/10.1081/E-ELIS3-120044418>
- Peroni, S and Shotton, D.** 2018. The SPAR Ontologies. In: Vrandeĭ, D et al. (ed.), *The Semantic Web – ISWC 2018*, 119–136. Cham: Springer International Publishing. ISBN: 978-3-030-00668-6. DOI: https://doi.org/10.1007/978-3-030-00668-6_8
- Przyborski, A and Wohlrab-Sahr, M.** 2019. Forschungsdesigns für die qualitative Sozialforschung. In Baur, N and Blasius, J (eds.), *Handbuch Methoden der empirischen Sozialforschung*, 41: 105–123. Wiesbaden: Springer Fachmedien Wiesbaden. ISBN: 978-3-658-21307-7. DOI: https://doi.org/10.1007/978-3-658-21308-4_7
- Steele, C, Butler, L and Kingsley, D.** 2006. The publishing imperative: The pervasive influence of publication metrics. *Learned Publishing*, 19(4): 277–290. ISSN: 09531513. DOI: <https://doi.org/10.1087/095315106778690751>
- Stocker, M,** et al. 2020. Persistent identification of instruments. *Data Science Journal*, 19. DOI: <https://doi.org/10.5334/dsj-2020-018>
- Vahdati, S,** et al. 2020. A comprehensive quality assessment framework for scientific events. *Scientometrics*, 126(1): 641–682. ISSN: 0138-9130. DOI: <https://doi.org/10.1007/s11192-020-03758-1>
- Wilkinson, MD,** et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>

TO CITE THIS ARTICLE:

Franken, J, Birukou, A, Eckert, K, Fahl, W, Hauschke, C and Lange, C. 2022. Persistent Identification for Conferences. *Data Science Journal*, 21: 11, pp. 1–18. DOI: <https://doi.org/10.5334/dsj-2022-011>

Submitted: 01 October 2021

Accepted: 18 February 2022

Published: 05 April 2022

COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.