Purdue University

Purdue e-Pubs

TIB AV-Portal: A Trusted Home for Conference Recordings

Bastian Drees
German National Library of Science and Technology, bastian.drees@tib.eu

Margret Plank
German National Library of Science and Technology, margret.plank@tib.edu

TIB AV-PORTAL: A TRUSTED HOME FOR CONFERENCE RECORDINGS

Bastian Drees and Margret Plank

Technische Informationsbibliothek (TIB) – German National Library of Science and Technology,
Germany
E-mail: Bastian.Drees@tib.eu and Margret.Plank@tib.eu

Abstract
Research results presented at scientific conferences are valuable information resources for scientific communities. Traditionally, the results are published in conference proceedings, documenting the current state of research. Additionally, an increasing number of conferences are recorded and the videos are subsequently published. While libraries have well-established procedures for collecting textual conference reports as part of the difficult-to-obtain grey literature, comparable procedures for audio-visual conference recordings have not yet been established. In most cases these videos are published by the organisers on commercial platforms like YouTube and Vimeo or directly on the conference website. A systematic search for conference recordings is difficult e.g. because the URL changes or external links lead to nowhere. Usually, conference websites are not maintained on a long-term basis and commercial platforms may remove videos or change the conditions for access for a variety of reasons.

In order to prevent the loss of conference recordings the German National Library of Science and Technology (TIB) has developed the AV-Portal. The AV-Portal provides the ideal infrastructure to host, find and reuse scientific videos including conference recordings. It's a single access point for videos from different conferences and years. All videos are assigned a Digital Object Identifier (DOI). These persistent identifiers allow for reliable referencing both as stable online links and as correct citations in scholarly work.

This paper describes how sharing scientific results via audio-visual media has become an important part of scientific communication. Further the paper introduces the TIB AV-Portal as a trusted home for conference recordings. It describes the workflow for linking the recordings to the corresponding proceedings in TIB's extensive collection of conference reports and vice versa via DOIs, thereby fulfilling a key demand of the Pisa Declaration on Policy Development for Grey Literature Resources.

Keywords: grey literature, non-textual material, conference recordings, conference reports

1. Introduction
Scientific conferences play an important role within the scientific landscape. Research results and scientific work in progress are often presented long before the actual publication. However, most work presented at conferences is, if at all, published as grey literature and therefore difficult to obtain both for libraries and scientists. Additionally, more and more conference proceedings are published in digital form (a recent survey among organizers of conferences, meetings and events showed that "approximately 70% of organisers (68.8%/) and suppliers (66.8%) believe that in the future paper-free events will provide content and information in digital form only" [GCB, 2016]) thereby raising the same questions as in other areas of digital publishing, e.g. questions concerning copyright and digital preservation. Finally, the types of conference related publications are more diverse than before, leading to a combination of presentation slides, research data, software, conference proceedings, audio recordings, video recordings, graphics, tweets, blog postings, etc. Libraries are therefore confronted with the challenge of how to treat this variety of information sources. Locating, hosting, preserving, indexing, linking to related material, enabling scientific citation are but a few issues concerning these materials.

The Technische Informationsbibliothek (TIB) – German National Library of Science and Technology – Leibniz Information Centre for Science and Technology and University Library[1] is one of the largest specialised libraries worldwide for the realms of technology, architecture, chemistry, information technology, mathematics and physics. TIB is a member of the Leibniz Association and a national infrastructure facility for the provision of scientific information.

In addition to its large holdings, the library has also a number of unique collections characterised by their special content or regional orientation. These include PhD theses, patents and standards, and difficult-to-obtain grey literature from the fields of science and technology, as well as research and conference reports. The latter are often strategically ground-breaking, and a recognised, valuable source of information in many specialist disciplines. TIB holds more than 9 million items almost 60% of which being grey literature [TIB, 2015]. Along these lines, TIB supports the Pisa Declaration on Policy Development for Grey Literature Resources [Greynet, 2014] that demands a commitment to Open Access, a greater recognition to high-quality grey literature, and support for the collection development and digital preservation of grey literature.

Moreover, TIB has established a competence centre for non-textual materials (KNM), which has been funded by the Leibniz Association since 2011. Besides developing innovative software and infrastructures, the KNM focuses on collecting, indexing, providing and (digitally) archiving non-textual materials such as audiovisual media, 3D objects and research data. In order to improve the accessibility, citability and the sustainable use of scientific videos, the TIB developed the TIB AV-Portal[2] in cooperation with the Hasso Plattner Institute[3]. The portal provides access to more than 10.000 scientific videos including about 3.000 conference recordings.

In this paper we present the ongoing efforts of TIB to extend this unique collection of conference recordings and to establish a tight connection to related materials such as research data, presentation slides and especially conference proceedings.

2. The AV-Portal
The explosive growth in scientific video content has triggered the development of a video portal at TIB. The AV-Portal makes videos freely available on a permanent basis and focusses especially on videos from the areas of technology, architecture, chemistry, computer science, mathematics and physics. The bilingual portal (english / german) links the current state of the art of relevant multimedia retrieval technologies (text-, speech and image recognition) with semantic analysis in order to improve the discoverability of videos and video segments [Strobel, 2015]. All videos are assigned a Digital Object Identifier (DOI), i.e. a permanent and explicit identifier that also remains valid even if the URL of the video changes. That way, it can be ensured that digital objects can be linked and retrieved permanently. By combining the DOI with a "media fragment identifier" (MFID), the AV-Portal even offers the possibility to cite a video down to the second. As a result, videos can not only be linked reliably, but also cited correctly in scholarly work.

3. Trusted home for conference recordings
A recent survey [engage, 2017] among 30 conference hosts and organisers commissioned by TIB showed that 47% of the respondents had already produced conference recordings (40%) or plan to do so (7%) in the near future. However, almost all respondents (92%) publish the recordings on the corresponding conference website and almost half of them (46%) on YouTube while none of them is using persistent identifiers. This raises many problems concerning the scientific use and reuse of conference recordings as valuable information sources in the scientific landscape.

On the one hand, long-term accessibility is neither guaranteed on commercial platforms like YouTube nor on the conference website, which is often no longer maintained after the

---

[1] https://www.tib.eu/
[2] https://av.tib.eu/
[3] https://hpi.de/

conference is finished [Drees, 2016]. On the other hand, scientific reuse is hindered since a systematic search is not possible and citation according to scientific standards is difficult. Not to mention insufficient information on legal licences.

Fig. 1: Detailed display of a conference video in the AV-Portal. Keywords from the automated media analyses are shown on the right. The conference proceedings (print and e-Book) are linked in the "Related Material" section. The video is published under an open license (CC BY 4.0) and assigned a DOI (both shown in the "Formal Metadata" section). The citation link of the recent video segment, i.e. the combination of DOI and MFID, is dynamically changing right below the video player.

To solve these problems TIB acquires, hosts, preserves and indexes conference recordings and links them to related material such as conference proceedings. Digital Object Identifiers (DOI) are assigned to each video via DataCite[4]. These persistent identifiers prevent link rot and enable scientific citations. Moreover, by combining DOI and MFID, every video may be cited to the second just like books can be cited to the page. For instance, the recording of the opening session of the FOSSGIS conference 2015 in Münster, Germany may be cited using the DOI:10.5446/17570 [Helle, 2015]. In that session, three speakers gave a welcome talk. The use of MFID allows to exactly cite, e.g. the second talk, held bei Prof. Edzer Pebesma using https://doi.org/10.5446/17570#t=05:44,14:06. Likewise, it is also possible to cite only the four seconds in which he states that open source software is needed in computational science, "because we need complete reproducibility there", by using https://doi.org/10.5446/17570#t=08:43,08:47. Furthermore, to enable reuse of the video collection in other applications, the metadata of all videos is freely available at https://av.tib.eu/opendata in the public domain[5] as linked open data in the RDF format [Waitelonis, 2016]. Thus, the AV-Portal fulfills a key demand of the Pisa Declaration, namely the "use of persistent identifiers and open metadata standards for grey literature" [Greynet, 2014].
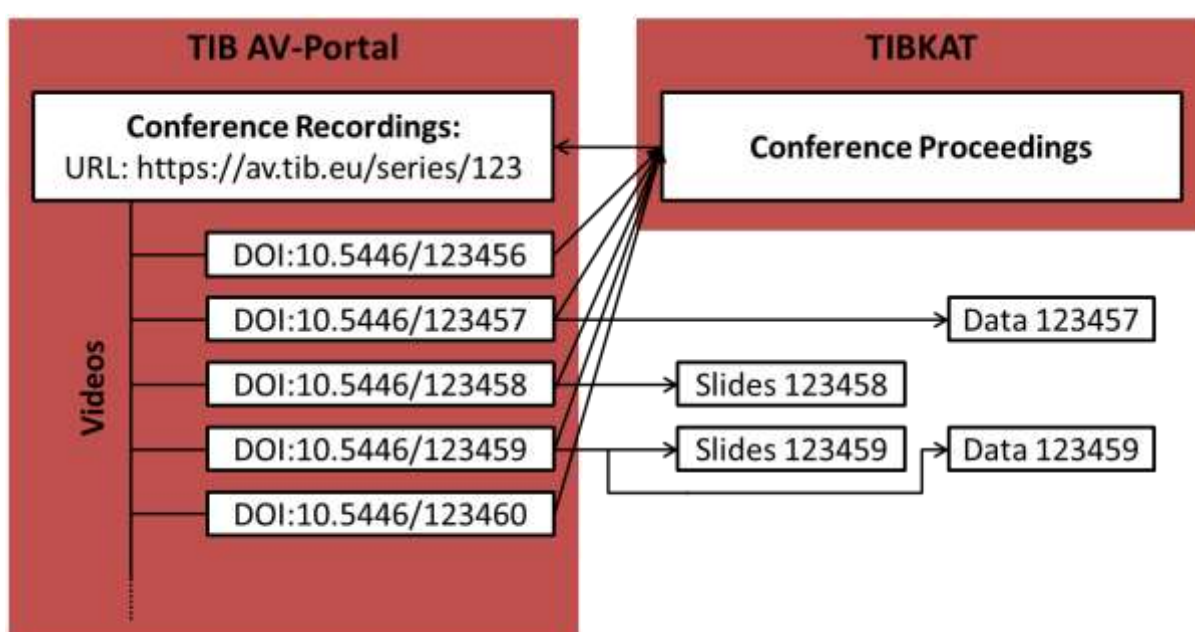

Fig. 2: All recordings of one conference are accessible via a series-URL, while each individual video is assigned a DOI. The metadata field "related material" of each video links to the conference proceedings in the TIBKAT (available via TIB-Portal at www.tib.eu) and if applicable to corresponding slides, data or publications. In the metadata of the conference proceedings, the series-URL is available in the field "additional information".

The value of the recordings is further increased by linking to related material such as presentation slides, research data or conference proceedings (cf. Fig. 2). Bidirectional links are especially desirable between conference proceedings and conference recordings. However, this raises some questions on how to implement those links in the respective metadata. While conference proceedings are in general not catalogued on article level, there is, of course, an individual entry for every video. On the other hand there is no individual entry for a video series in the AV-Portal, i.e. for all videos of one conference, which would correspond to the whole volume of the conference proceedings. There exists however a series link, i.e. a URL that contains all recordings of a given conference. As it is not desirable to place links to all related recordings in the metadata of the conference proceedings - this would inflate the metadata disproportionately - TIB decided on a solution as schematically shown in Fig. 2. The conference proceedings are catalogued in the catalogue of TIB and link to the URL of the series link in the AV-Portal (cf. Fig. 3). The other way around, there is a link in the metadata of every video to the conference proceedings (cf. Fig. 1). In this way there is a bidirectional connection between

---

[4] http://www.datacite.org/
[5] CC0 1.0 universal; http://creativecommons.org/publicdomain/zero/1.0/deed.de

proceedings and recordings along the lines of the Pisa Declaration that demands "systems for linking data and other non-textual content to their grey literature publications" [Greynet, 2014].

Document information

Conference:	FOSSGIS; 2016; Salzburg
	Anwenderkonferenz für Freie und Open Source Software für Geoinformationssysteme; 2016; Salzburg
Size:	101 Seiten
Remarks:	Illustrationen
ISBN:	9783000534379
Type of media:	Conference Proceedings
Type of material:	Print
Language:	German
Keywords:	Kongress, Open Source, Zweitausendsechzehn
Classification:	BKL: 74.48 Geoinformationssysteme / 54.52 Software engineering
Table of contents:	http://www.gbv.de/dms/tib-ub-hannover/873586530.pdf
Additional information:	https://av.tib.eu/series/230/fossgis+konferenz+2016+salzburg+04+06+juli+2016

Fig. 3: Catalogue entry of FOSSGIS 2016 conference proceedings as displayed in the TIB-Portal. The series-URL of the related conference recordings in the AV-Portal is given in the additional information field.

4. Cooperations

Most organisers or hosts of scientific conferences do not have a sustainable infrastructure to publish their conference recordings or to guarantee their long-term accessibility. On the other hand it is almost impossible to collect the relevant metadata and necessary legal licences from all speakers of a conference after the event. Therefore it is beneficial both for the organisers of a conference as for TIB to arrange a cooperation prior to the event. TIB has established cooperations with a number of conference organisations accordingly. These cooperations range from simple agreements on hosting, reuse and licensing to media partnership with semi-automated exchange of metadata. Examples are FOSS4G 2016[6], EuroPython[7] and the Free and Open Source Software Conference (FrOSCon)[8]. Media partnerships have been established, e.g. with FOSS4G 2017 in Boston and the German OSGeo chapter FOSSGIS.

5. Conclusion

Today's researchers move beyond text as their publications such as conference proceedings are enhanced by a variety of digital assets, e.g. recordings, conference slides and research data. However, only a negligible proportion of those digital assets are accessible at present, whilst text based conference proceedings are, in principle, sufficiently well-documented and available.

This paper summarizes the current status from the efforts undertaken by TIB to ensure the long-term provision of conference recordings and additional material in a legally sound manner. The video hosting arrangement includes indexing according to international standards, semantically enhancement, transcription, digital preservation and finally the allocation of a DOI name. In order to optimise the material's discoverability, the TIB AV-Portal - a web-based platform featuring state-of-the-art multimedia retrieval technology and semantic video analysis - was developed. Thus this library-operated service infrastructure underlines the way that the output of researchers is understood today: a combination of journal article, datasets, and source code, as well as video, such as conference recordings, video abstracts or supplements – all linked to each other through persistent identifiers like DOIs.

---

[6] https://av.tib.eu/series/253/foss4g+2016+bonn
[7] https://av.tib.eu/search?f=publisher%3Bhttp://av.tib.eu/resource/EuroPython
[8] https://av.tib.eu/series/231/froscon+2015

5. References

Drees, B. (2016). *404: Video not found - The TIB AV-Portal stops scientific videos from disappearing.* Retrieved from TIB Blog: https://blogs.tib.eu/wp/tib/2016/11/16/404-video-not-found-the-tib-av-portal-stops-scientific-videos-from-disappearing/ . Retrieved on 2017-06-15

engage AG, & Technische Informationsbibliothek (TIB). (2017). Marktpotentialanalyse "Konferenzaufzeichnungen". unpublished.

GCB German Convention Bureau e.V. (2016). *Meeting- & Event-Barometer 2016: Trends - Modern technologies are in demand in the meetings industry.* https://www.gcb.de/fileadmin/GCB/News/Press_Releases/Downloads/Meeting_EventBa EventB2016_presskit.pdf . Retrieved on 2017-06-15

Greynet. (2014). *Pisa Declaration on Policy Development for Grey Literature Resources.* http://www.greynet.org/images/Pisa_Declaration,_May_2014.pdf . Retrieved on 2017-06-15

Helle, D., Pebesma, E., & Knopp, O. (2015). FOSSGIS Konferenz 2015 - Eröffnung. (FOSSGIS, Hrsg.) Münster, Germany. DOI:10.5446/17570

Strobel, S., & Marín-Arraiza, P. (2015). *Metadata for Scientific Audiovisual Media: Current Practices and Perspectives of the TIB|AV-Portal.* In: Garoufallou E., Hartley R., & Gaitanou P. (eds). *Metadata and Semantics Research. Communications in Computer and Information Science*, vol 544. 159-170. Springer, Cham. DOI:10.1007/978-3-319-24129-6_14

Technische Informationsbibliothek (TIB). (2015). *Jahresbericht 2015.* https://www.tib.eu/fileadmin/Daten/presse/dokumente/tibub-jahresbericht-2015.pdf . Retrieved on 2017-06-15

Waitelonis J., Plank M., & Sack H. (2016). *TIB|AV-Portal: Integrating Automatically Generated Video Annotations into the Web of Data.* In: Fuhr N., Kovács L., Risse T., & Nejdl W. (eds). *Research and Advanced Technology for Digital Libraries. TPDL 2016. Lecture Notes in Computer Science*, vol 9819. 429-433, Springer, Cham. DOI:10.1007/978-3-319-43997-6_37