

# The networked edition humboldt digital

## Dumont, Stefan

dumont@bbaw.de  
Berlin-Brandenburg Academy of Sciences and Humanities,  
Germany

## Kraft, Tobias

kraft@bbaw.de  
Berlin-Brandenburg Academy of Sciences and Humanities,  
Germany

## Thomas, Christian

thomas@bbaw.de  
Berlin-Brandenburg Academy of Sciences and Humanities,  
Germany

The scholarly edition project *Alexander von Humboldt auf Reisen - Wissenschaft aus der Bewegung* (AvH-R), based at the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW), edits the American, Russian-Siberian and European travel journals as well as thematically related letters and manuscripts of the Prussian naturalist and explorer Alexander von Humboldt.<sup>1</sup> Research articles, a chronology of Humboldt's life and extensive indexes supplement these texts. The publication strategy is "digital first", i. e. the documents are initially published online under a free licence, with a curated selection to appear later in a series of print editions. Taking into account the decade-long history of Humboldt-related research at the BBAW, AvH-R has been pioneering the continuous digitization of Alexander-von-Humboldt-Research, creating the first fully digital manuscript edition of his works.

Since its launch in 2016, eight versions of the project's *edition humboldt digital* (*ehd*) have been released.<sup>2</sup> AvH-R's editorial team works with *ediarum*,<sup>3</sup> a research software solution based on X-technologies, the free XML database eXist-db and the software Oxygen XML Author. With the current version 8 of the *ehd*, the project keeps its promise to make the digital, text-critical documentary edition<sup>4</sup> available *as data* (mainly in TEI-XML) on GitHub<sup>5</sup> and (from summer 2023) for long-term archiving and worldwide access in the Zenodo repository. This step significantly enhances the re-usability of the integral research data from the edition, e. g. for initiatives such as the National Research Data Infrastructure Germany (NFDI), where the BBAW is involved as a co-applicant and active partner in the Text+ consortium.<sup>6</sup>

In addition to accessing the digital edition via the web interface of the *ehd* or via the GitHub- and Zenodo-repositories, the data can also be retrieved via an application programming interface. The API itself offers not 'only' the annotated full text transcriptions, but also various metadata to enable comprehensive networking of the edition, e. g. the metadata for the edited texts and research articles in Dublin Core format via OAI-PMH (Lagoze et al. 2015). Thus all texts can be found in the open access search engine BASE.<sup>7</sup> In addition, *ehd* provides both a BEACON interface (Voß and Schindler 2017) and a CMIF interface for correspSearch (Dumont, Grabsch and Müller-Laackman 2021).

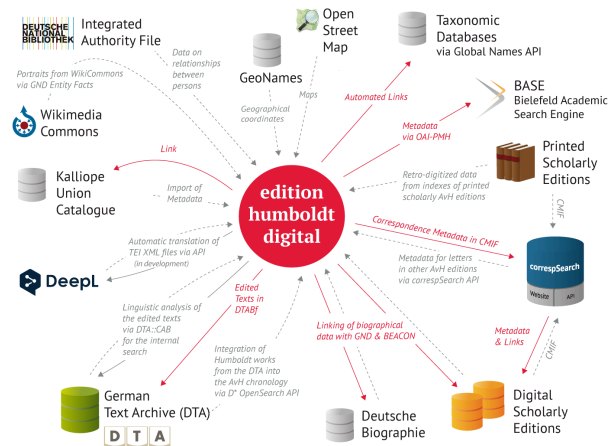


Fig. 1: *edition humboldt digital* uses or links to external data and web services (Cf. Kraft/Dumont 2020)

*edition humboldt digital* not only provides its own data, but also uses other data and external web services to enrich its functions (fig. 1). For example, the TEI-XML files are sent to the web service DTA::CAB<sup>8</sup> (Jurish 2011) for lemmatisation. This enables a lemma-based search that takes into account (historical) spelling variants and inflections.

Apart from the search function, the chronology of Alexander von Humboldt's life is a central networking and integration point for external resources and is exemplary for the project's attempt not only to create Humboldtian scholarship as digital data but also to restore and curate text-based research data for the digital age. This chronology was begun in the 1960s at the then German Academy of Sciences and Humanities in Berlin, revised in the 2000s and published as an HTML version on the web, converted to TEI-XML in 2015/16 and finally integrated into the *edition humboldt digital*. There, it is continuously maintained, expanded and linked to the edited texts and index entries. In addition, it integrates various external resources and services into the edition, such as the metadata of Alexander von Humboldt's published correspondence from correspSearch,<sup>9</sup> Humboldt's works from the German Text Archive<sup>10</sup> and entries from the court calendar of the Prussian monarchy<sup>11</sup> with reference to Humboldt's life. In this way, the chronology opens up the *edition humboldt digital* to numerous externally available materials and information. By intensive cross-referencing, users are automatically made aware of additional sources available from other projects or in earlier (printed) editions without having to investigate themselves and without leaving the *ehd* platform.

The indexes also use external data. On the one hand, via BEACON interfaces other related and featured digital publications, such as Alexander von Humboldt's Kosmos-Lectures in the German Text Archive, are linked automatically.<sup>12</sup> In addition, with the help of entire records of the GND (i. e. not just the URIs) the index entries of the *ehd* are automatically linked to *each other* – based on family and friendship relations noted in the GND. One index that makes extensive use of external services is the plant index. Based on the taxonomic names, various web services are queried and linked automatically. In this way, the edition opens up Humboldt's diaries to various disciplines such as biodiversity research. With the intensive use of external web services and data, the benefit of the digital edition increases significantly.

## Notes

1. *edition humboldt digital*: <https://edition-humboldt.de/>. Cf. Kraft/Dumont 2020.
2. For details on the quality and quantity of the published research data see the site's section "Versionen" at <https://edition-humboldt.de/about/index.xql?id=H0020382&l=en>.
3. <https://www.ediarum.org/>; see also Dumont et. al 2021.
4. For an orientation on the ehd edition model, see Sahle 2016 and in particular the concept of the 'documentary edition', for example Pierazzo 2011.
5. <https://github.com/telota/edition-humboldt-digital>.
6. *NFDI*: <https://www.nfdi.de/?lang=en>; *Text+*: <https://www.text-plus.org/en/home/>. Team members of the *ehd* and *Telota* have contributed a "User Story" to motivate the development of more *multilingual* query services within the consortium, available at <https://www.text-plus.org/en/research-data/user-story-408/>.
7. <https://www.base-search.net/>
8. <https://www.deutschestextarchiv.de/demo/cab/>.
9. <https://correspsearch.net/de/suche.html?s=http://d-nb.info/gnd/118554700>.
10. <https://edition-humboldt.de/chronologie/index.xql?jahr=1827&dta=on>.
11. Cf. Introduction Court Calendar, <https://actaborussica.bbaw.de/v5/P0006298>, in: Akademienvorhaben "Anpassungsstrategien [...]" 2021.
12. See, for example, the register entry on August Böckh: <https://edition-humboldt.de/H0003413>.

## Bibliography

Akademienvorhaben "Anpassungsstrategien der späten mitteleuropäischen Monarchie am preußischen Beispiel (1786-1918)" (Ed.) (2021): *Die Sprache der Monarchie* (Version 5). Berlin: Berlin-Brandenburgische Akademie der Wissenschaften. URL: <https://actaborussica.bbaw.de/>

**Dumont, Stefan / Nadine Arndt / Sascha Grabsch / Lou Klappenbach** (2021): *ediarum.BASE.edit* (Version 2.0.0) [Computer software]. <https://doi.org/10.5281/zenodo.5897100>

**Dumont, Stefan und Sascha Grabsch und Jonas Müller-Laackman** (Ed.) (2021): *correspSearch – Briefeditionen vernetzen* (2.0.0) [Webservice]. Berlin: Berlin-Brandenburgische Akademie der Wissenschaften. <https://correspSearch.net>.

**Ette, Ottmar** (Ed.) (2022): *edition humboldt digital* (Version 8). Berlin: Berlin-Brandenburgische Akademie der Wissenschaften. <https://edition-humboldt.de/>.

**Jurish, Bryan** (2011): *Finite-State Canonicalization Techniques for Historical German*. Potsdam: Universität Potsdam. urn:nbn:de:kobv:517-opus-55789.

**Kraft, Tobias und Stefan Dumont** (2020): "The Humboldt Code. On creating a hybrid digital scholarly edition of a 19<sup>th</sup> century globetrotter." In *Wiener Digitale Revue* 1. <https://doi.org/10.25365/wdr-01-03-02>.

**Lagoze, Carl, Herbert Van de Sompel, Michael Nelson, and Simeon Warner** (2015): "The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0". <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>.

**Pierazzo, Elena** (2011): "A Rationale of Digital Documentary editions". In *Literary and Linguistic Computing*, 26,4, 463-477. <https://doi.org/10.1093/lc/fqr033>.

**Sahle, Patrick** (2016): "What is a Scholarly Digital Edition?" In *Digital Scholarly Editing: Theories and Practices*, Ed. von Matthew James Driscoll und Elena Pierazzo, 19-39. Cambridge, UK: Open Book Publishers. <https://books.openedition.org/obp/3397>

**Voß, Jakob, and Matthias Schindler** (2017): 'BEACON Link Dump Format'. Informational. Göttingen. <http://gbv.github.io/beaconspec/beacon.html>.