

Semantic minimal retrospective digitization of edited correspondence

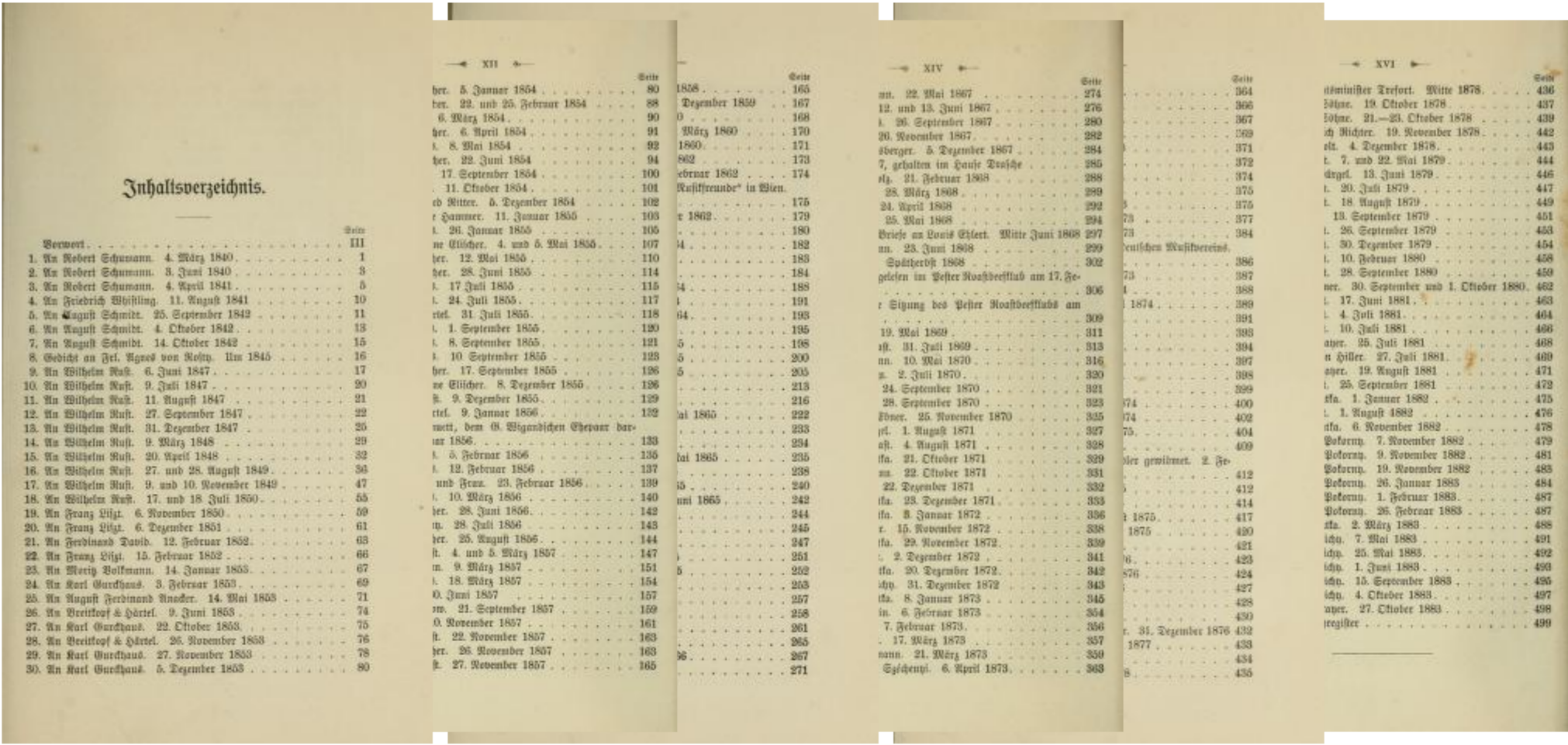
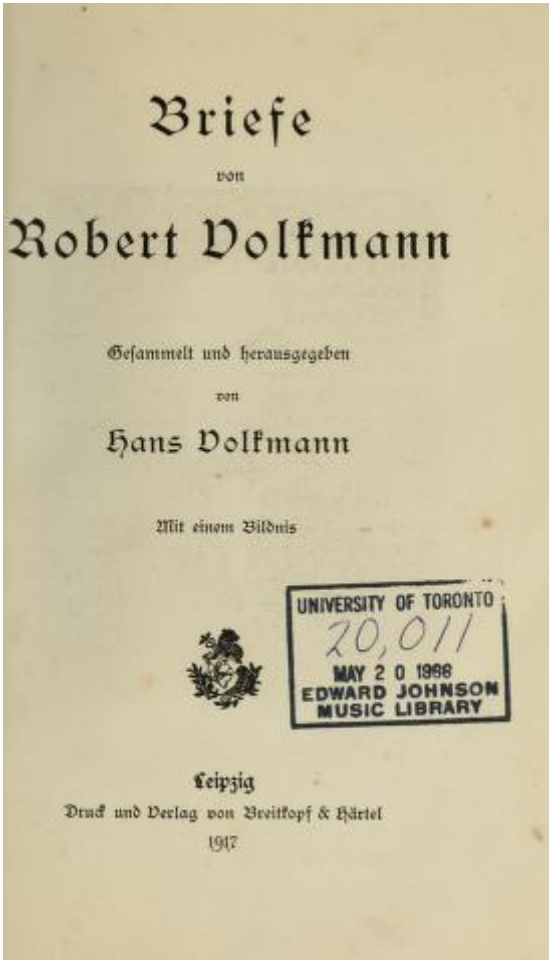
Klaus Rettinghaus

Correspondence is in general an essential key to research in many aspects. That is why around the world so much effort goes into editing corpora of letters. A problem is that most of the available edited correspondence is still imprisoned between cover boards and thus may stay hidden for the researchers area of interest.

- A complete retrospective digitization of a (critical) edition with OCR software basically is
- cost-intensive
 - time-consuming
 - prohibited by copyright laws
 - often of poor quality
 - nonsemantic.

And even then, it is highly unlikely, that it is to be found.

But the pure metadata for the edited correspondence however is not directly subject to copyright restrictions. And with the relatively new elements for **correspondence description**, introduced in the TEI Guidelines in April 2015, it is now easily possible to encode this correspondence metadata in a fully structured and standardized TEI file. On this basis, it is possible to create a so-called *semantic minimal retrospective digitization*, which only assembles the



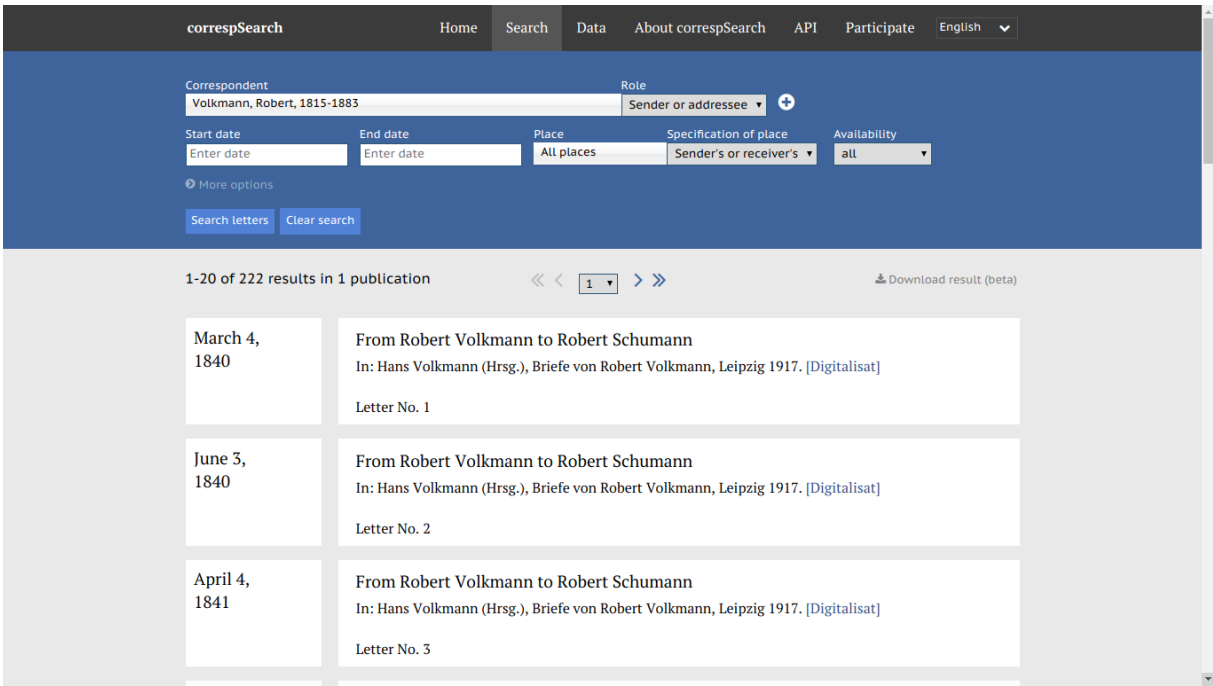
metadata from a printed edition in a digital index, enriched with authority control data, that is both, human-readable and machine-readable. That in turn makes this data interchangeable and thus easily searchable with external services like *correspSearch.net*.

For larger corpora it is not recommendable to try to compile a suitable XML index file by hand. A possible work flow that is way quicker and less prone to errors could be to enter the data in a simple table, which then gets converted into the desired *Correspondence Metadata Interchange format* with appropriate tools like *CSV2CMI*.

Nevertheless, it requires some effort. So by whom should this work be done? Libraries as information retrieval experts in the digital age come to mind first. Or scholars who e. g. have to work with scattered letters.

key	sender	senderID	addressee	addresseeID	senderDate	page
1	Robert Volkmann	11876912X	Robert Schumann	118611666	1840-03-04	
2	Robert Volkmann	11876912X	Robert Schumann	118611666	1840-06-03	3
3	Robert Volkmann	11876912X	Robert Schumann	118611666	1841-04-04	5
4	Robert Volkmann	11876912X	Friedrich Whistling		1841-08-11	10
5	Robert Volkmann	11876912X	August Schmidt	117501816	1842-09-25	11
6	Robert Volkmann	11876912X	August Schmidt	117501816	1842-10-04	13
7	Robert Volkmann	11876912X	August Schmidt	117501816	1842-10-14	15
8	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1847-06-06	17
9	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1847-07-09	20
10	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1847-08-11	21
11	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1847-09-27	22
12	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1847-12-31	25
13	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1848-03-09	29
14	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1848-04-20	32
15	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1848-06-28	36
16	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1848-07-18	65
17	Robert Volkmann	11876912X	Wilhelm Rust	116711728	1850-07-18	
18	Robert Volkmann	11876912X	Franz Liszt	1850-11-06	59	
19	Robert Volkmann	11876912X	Franz Liszt	1851-12-06	61	
20	Robert Volkmann	11876912X	Ferdinand David	1852-02-12	63	
21	Robert Volkmann	11876912X	Franz Liszt	1852-02-15	66	
22	Robert Volkmann	11876912X	Moritz Volkmann	1853-01-14	67	
23	Robert Volkmann	11876912X	Karl Guckhaus	1853-02-03	69	
24	Robert Volkmann	11876912X	August Ferdinand Anacker	1853-05-14	71	
25	Robert Volkmann	11876912X	August Ferdinand Anacker	1853-05-14	74	

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <!-- Generated from table of letters with csv2cmi 1.3.4 -->
  <fileDesc>
    <title>
      <title xml:id="title_m473gf7_x59F5dmd">cmi2csv example</title>
    </title>
    <editor>Klaus Rettinghaus</editor>
    </fileDesc>
    <publicationInfo>
      <publsher>Sächsische Akademie der Wissenschaften zu Leipzig</publsher>
      <idno type="url">http://www.leipzig.de/VolkmannBriefe.xml</idno>
      <date when="2017-11-30T11:43:15.799124">
        <availability>
          <license target="https://creativecommons.org/licenses/by/4.0/">This file is licensed under the terms
        </availability>
        </source>
        <dbtl type="print" xml:id="edition_010bbez4_r1512p9v">Briefe von Robert Volkmann, Leipzig 1917</dbtl>
      </source>
    </publicationInfo>
    <profileDesc>
      <correspDesc key="1" source="#edition_010bbez4_r1512p9v" xml:id="letter_d3f6b0v_l1p2zaccpw">
        <correspaction type="sent" xml:id="sender_4ykrBsnr_r2Bokkg">
          <persName ref="http://d-nb.info/gnd/11876912X">Robert Volkmann</persName>
          <date when="1840-03-04">
            </correspaction>
          <correspaction type="received" xml:id="addressee_cnzld92_5xt46nd">
          <persName ref="http://d-nb.info/gnd/118611666">Robert Schumann</persName>
          <correspaction>
            <correspdesc key="2" source="#edition_010bbez4_r1512p9v" xml:id="letter_373vywgh_lv1er5oo">
              <correspaction type="sent" xml:id="sender_5ng3qph_l609w2d">
                <persName ref="http://d-nb.info/gnd/11876912X">Robert Volkmann</persName>
                <date when="1840-06-03">
              </correspaction>
              <correspaction type="received" xml:id="addressee_qeb53wdz_luv1nj0">
```



CorrespSearch is a web service operated by the Berlin Brandenburg Academy of Sciences and Humanities.

CSV2CMI is free and available on **GitHub**

<https://github.com/saw-leipzig/csv2cmi>



This poster is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

DOI: 10.5281/zenodo.1410797



Sächsische Akademie der Wissenschaften zu Leipzig