

---

# Criteria for Recording and Categorizing Scholarly Digital Editions

Simon Sendler\*<sup>1</sup>, Daniela Schulz\*<sup>2</sup>, Hanna-Lena Meiners<sup>1</sup>, and Bernhard Fisseni<sup>3</sup>

<sup>1</sup>Göttingen State and University Library – Germany

<sup>2</sup>Herzog August Bibliothek Wolfenbüttel – Germany

<sup>3</sup>Leibniz-Institut für Deutsche Sprache Mannheim – Germany

## Abstract

CLARIAH-DE is a project funded by the German Ministry of Research and Education, merging two original infrastructure projects (CLARIN-D and DARIAH-DE) concerned with the digital investigation of textual and linguistic sources within the humanities and cultural sciences. An important goal of CLARIAH-DE is to evaluate which representations are adequate for different kinds of data and data processing. The native format of an edition must be able to represent all relevant aspects of the edition. However, as different fields and methods of data processing employ different data formats, possibilities of transforming scholarly digital editions into various digital representations must be explored. This paper describes a procedure to provide initial recommendations for considering transformation potential based on information on the edition goals and data collected through a detailed online questionnaire.

Upon completing the questionnaire, respondents receive initial recommendations on data conversion into pivot formats and on the conversion of already existing digital editions to different native formats. Users are also offered contact information for further counseling and feedback as well as literature and web references. The recommendations of the survey provide a starting point and will need to be reflected on and adapted according to individual requirements.

Before developing recommendations, the authors determined appropriate criteria for assessing the complexity of a digital edition according to seven categories based on an informal evaluation of editions curated in the authors' home institutions, and in the DARIAH and TextGrid communities. The questionnaire is based on these criteria, and is designed for users without much experience with the various data formats available to scholars. Users can consult this questionnaire to examine the convertibility of their edition into a simpler pivot format to facilitate data exchange.

The questionnaire comprises seven questions ranging from the overall aim of the edition to type and number of source documents and desired features and depth of data capturing. Questions are close-ended and multiple choice. They accommodate different scholarly objectives. Each answer is assigned a numerical score related to the scholarly objective and the use(fulness) of different formats. The total score will offer an indication of the relative complexity of a digital edition, where complexity refers to both the variety and depth of various modes of annotation.

---

\*Speaker

Questionnaire responses will be collected to document the use of the questionnaire and to assess the different needs of the community; collected data will also be used to improve the questionnaire. For this purpose, the open source software LimeSurvey hosted by the GWDG is being used.

The criteria as well as the mapping of criteria to recommendations in the questionnaire are reviewed, commented on, and documented by scholars involved in various current and recent scholarly digital editions, as well as scholars experienced in data conversion and annotation formats. Moreover, completed as well as ongoing edition projects will be explored and thus will help to validate the initial recommendations. Prospectively, the authors hope to extend the functionality of the questionnaire to also provide recommendations during the early stages of planning an edition.

**Keywords:** Digital Humanities, digital scholarly editions, interchangeability, standardization, interoperability, reuse, infrastructures