

DH2013 Abstract:

A key challenge facing professional researchers in the domain of cultural heritage across Europe and worldwide is the interrogation of growing digital humanities collections. However, the full value of these heritage treasures is not being realised. After digitisation, these collections are typically monolithic, difficult to navigate and can contain text which is highly variable in terms of language, spelling, punctuation, and consistency of terminology. These difficulties are compounded by a lack of normalised spelling in most European languages before the eighteenth century. This means that search across these digital collections tends to return sub-par results as multiple spellings for many common words are treated as independent document keywords. CULTURA is a corpus agnostic environment with a suite of services, including personalisation, annotation, and recommendation, providing necessary supports and features for a diverse range of professional researchers