

'Catalogue these books': Digital Editions and the Digital Library

Ronan Crowley
Universität Passau
crowle01@gw.uni-passau.de

With over twenty million volumes contained in Google Books and over ten million in HathiTrust – to take but two examples – LSDI scanning represents the dominant economy of book digitisation. This poster presentation will address the intersections of digital archives with digitised material produced by other means: the digital editions of authors' manuscripts encoded in markup languages (for example, in the TEI XML format). What are the points of contact and friction that obtain between Big Text Data and heavily curated, deeply encoded editions of archival materials? What are the challenges and possibilities for interpretation afforded by competing or complementary scales of digital materiality?

The case study animating the project centres on an historical practice of reading recoverable from the holograph notebooks that James Joyce compiled while writing *Ulysses* (1922), and which are now housed at the National Library of Ireland. Essentially commonplace books, these documents list hundreds of lone words and abbreviated phrases, typically copied verbatim during the course of Joyce's reading. This storehouse provided the raw material for his novel in progress. Assiduous LSDI browsing and coordinated search is recovering the sources so integrated into *Ulysses*, but to what degree can such material be included in digital editions that prioritise text encoding? What is the compromise, then, for a digital materiality reckoned at scale between, on the one hand, the impracticality of marking-up large-scale data collections and, on the other, the desired granularity of the digital edition?