

**SIB - SNSF MARK16 VREs conference - Lausanne, 10-11 September 2020 - mark16.sib.swiss**

W. Andrew Smith, University of Edinburgh

### **Mining Manuscript Data in the New Testament Virtual Manuscript Room**

In the late 1980s, businesses were collecting customer data because they knew it was valuable to what would later be termed business intelligence, but until the 1990s there was no clear path to how to mine and usefully exploit those data. The New Testament Virtual Manuscript Room is in the business world's 1980s: we have collected great amounts of textual data, but those data could be mined for information much broader than critical editions of the New Testament. Reconceptualizing the use of that massive database would prompt changes in data entry, as the encoding for data points such as codicological and paratextual features are woefully under-developed in TEI-XML. This reconceptualization would also require new means to query and post-process the data. This paper will provide a proposal for how this extremely valuable VRE could be developed into a much more powerful resource.