

Digital humanities methods in comparative law.
Quantitative analysis on a plain text corpus of books to trace the
diffusion of legal concepts in public spheres.

Thomas Favre-Bulle*¹

¹EPFL ENAC IA ALICE, École Polytechnique Fédérale de Lausanne

February 28, 2014

Tracing the diffusion of legal concepts between different countries is at the core of comparative law. However, this diffusion is not limited to statutes and cases, but pervades through public spheres. This is visible in published materials of all sorts. How to trace this diffusion in public sphere? If cases are important to reconstruct the logics of diffusion in legislative and judicial spheres, one cannot rely on a limited number of sources to assess its importance in general public spheres.

In contrast with highly abstracted library databases, where searches have to be conducted by title and indexed topics, the Google books corpus gives access to plain-text search in a significant share of all the material ever published since the 17th century. I will present in this paper how to conduct analysis on serialized quantitative time series from the Google Books corpus, using GCorpusAnalytics, a tool I developed at the Swiss Federal Institute of Technology in Lausanne (EPFL).

I will present an application of this tool in the exploration of the diffusion of direct democracy instruments from Switzerland to the United States at the end of the 19th century. Specifically, I will show how a quantitative analysis of the google books corpus reveals that Switzerland was indeed the reference, and not New England, and how we can trace the influence of key literature in the public debate at the time. Additionally, I will present the possibility of querying the Google Case Law corpus in time series.

*thomas.favre-bulle@epfl.ch