

13:00–14:00, Thursday 6 October 2016
Centre for Digital Scholarship
Weston Library



Simon Baker
Jonathan Blaney
Marty Steer

INSTITUTE OF HISTORICAL RESEARCH
UNIVERSITY OF LONDON

Peaches and lemons are foodstuffs.
Trying to classify historical texts
using the BBIH Thesaurus.

The Institute of Historical Research collates and classifies a large variety of material relating to History in UK Higher Education. With the rise of open access publishing and the proliferation of web resources, keeping up has become more and more difficult. Funding from the AHRC has allowed the IHR's digital team to explore ways of automatically classifying works based on features such as book blurbs, article abstracts, titles and metadata.

This talk will cover the challenges the AHRC-funded TOBIAS project has met in classifying this material, using the Royal Historical Society's fine-grained vocabulary of British and Irish history. Techniques the team is trying range from “bag of words” string matching to machine learning.

Simon Baker is the editor of the Bibliography of British and Irish History (BBIH).

Jonathan Blaney is the Project Editor for British History Online.

Marty Steer is the Website Manager at the Institute of Historical Research.

All three speakers are engaged with the TOBIAS project, funded under the AHRC's Follow-on Funding Impact & Engagement Scheme.

Register for the talk via: <http://www.bodleian.ox.ac.uk/whatson/whats-on>.

DIGITAL
METHODS