Converging cultures of open in language resources development

Alannah Fitzgerald

https://www.flickr.com/photos/wolfgangfoto/4705194692
MINING & LINKING OPEN CONTENT FOR DATA DRIVEN LEARNING
Data-Driven Learning

The metaphor that Johns evoked was one where language is treated as empirical data and “every student is a Sherlock Holmes”, investigating the uses of linguistic data directly to assist with language acquisition (Johns, 2002, p. 108).
FLAX Academic English Collections

PhD Abstracts Collections
These collections come from the E-theses Online Service (EThOS) Open Access Initiative managed by the British Library.

- Arts and Humanities
- Life Sciences
- Physical Sciences
- Social Sciences
- Useful words for academic writing

British Academic Written English Collections
These collections come from the British Academic Written English (BAWE) corpus, which was developed at the Universities of Warwick, Reading and Oxford Brookes.

- Social Sciences
- Arts and Humanities
- Physical Sciences
- Life Sciences
- Useful words for academic writing

Collections Created by Registered Users
If you would like to build your own collections and put them here, please email us.

Note: Note: we have moved some teacher created collections, which are currently under construction, to collections.flax.nzdl.org. However, if you would like your finalised collections to appear here, please email us.

- Financial Crisis Lectures

LAW Collections
The resources used in these law collections come from open podcasts, Massive Open Online Courses (MOOCs) and Open Access publications. They have been developed to support learners with Legal English and to demonstrate the types of domain-specific collections that can be built using the FLAX software.

- Age of Globalization MOOC (University of Texas at Austin with edX)
- English Common Law MOOC (University of London with Coursera)
- Environmental Politics and Law (Yale University with OpenYale)
- CopyrightX (Harvard University)
- Law Articles (Open Access Law Journal Publications)
- Legal Terms List
- British Law Report Corpus (BLaRC)
- ContractsX (Harvard University)
- Law PhD Theses Abstracts (EThOS at the British Library)

EThOS at the British Library

- Social Sciences PhD Thesis Abstracts
- Arts and Humanities PhD Thesis Abstracts
- Law PhD Thesis Abstracts
- PhD Thesis Abstracts on Water Politics and Tourism Studies
- STEM PhD Thesis Abstracts

FLAX Language Digital Library Project, University of Waikato, NZ

MANAGED OPEN ACCESS FOR REDISTRIBUTION IN EAP
Repurposed as Open EAP Corpora

• *Electronic Theses Online Service (EThOS)*, managed by the British Library
  – Abstract metadata of 400,000 PhD theses via the *EThOS* toolkit
  – “reuse by third parties for not-for-profit purposes” (British Library, n.d.).

• *British Academic Written English (BAWE)* corpus, managed by the Oxford Text Archive (OTA)
  – 2500 pieces of learner writing (3 UK universities)
  – A formal request must be registered with the OTA to develop the *BAWE* corpus for non-commercial “research use or educational purposes” (IT Services, University of Oxford, OTA, 2015)
FLAX ETHOS PHD ABSTRACTS
COLLECTIONS FOR ACADEMIC ENGLISH
EThOS PhD Abstracts Collections

PhD Abstracts (Social Sciences)

About this collection

This collection contains 8769 pieces of proficient assessed PhD student writing from disciplinary areas of the Social Science. It is a sub-collection of the PhD thesis abstracts collections in FLAX and is derived from the EThOS Open Access toolkit, which is made available by the British Library for reuse by third parties for not-for-profit purposes. The copyright remains with the authors of the PhD theses unless otherwise stated in the licensing information provided by EThOS.

ETHOS at the British Library uses the Dewey Decimal Classification system, which organizes library materials by discipline or field of study. The main divisions of the Dewey Decimal Classification system include philosophy, social sciences, science, technology, and history. These main divisions fall within a scheme made up of ten classes, each divided into ten divisions, and each having ten sections. Here with the PhD Abstracts Collections in FLAX, abstracts are directed into one of four over-arching discipline divisions to make up the following sub-collections: Arts and Humanities, Social Sciences, Physical Sciences and Life Sciences. This Social Sciences sub-collection in FLAX includes abstracts that have been classified into discipline sub-divisions from the 000, 300 and 600 classes of the Dewey Decimal Classification system, specifically:

- 020 Library and information sciences
- 300 Social sciences, sociology and anthropology
- 310 Statistics
- 320 Political science
- 330 Economics
- 340 Law
- 350 Public administration and military science
- 360 Social problems and social services
- 370 Education
- 380 Commerce, communications and transportation
- 390 Customs, etiquette and folklore
- 600 Technology
- 640 Home and family management
- 650 Management and public relations

Using Dewey Decimal Classification

. 330 Economics
  . 330 Economics
  . 331 Labor economics
  . 332 Financial economics
  . 333 Economics of land & energy
  . 334 Cooperatives
  . 335 Socialism & related systems
  . 336 Public finance
  . 337 International economics
  . 338 Production
  . 339 Macroeconomics & related topics

. 340 Law
  . 340 Law
  . 341 Law of nations
  . 342 Constitutional & administrative law

Browse by Discipline

PhD Abstracts (Social Sciences)

- The adoption of e-government in the Kingdom of Bahrain
- Consumers saving behaviour: an empirical investigation of consumers attitudes towards commercial banking services with special reference to savings accounts in Iraq
- Rhetorical strategies of legitimation: the 9/11 Commission’s public inquiry process
- Multidisciplined individuals: defining the genre
- Regulation of foreign investment in Kenya, 1963-81: an empirical study
- Modelling, forecasting and riding credit risk in the Sterling Eurobond market
- Exploring working lives through the framework of the ‘psychological contract’: a study of clergy in the Church of England in the 21st century
- Process design in an information-intensive service delivery system: an empirical study
- Social discount rates and welfare weights for public investment decisions in Turkey
- Financial market efficiency: a study of the time series properties of the Jordanian stock market
- Essays on wage inequality: the role of composition, immigration and the cost-of-living
- An ontological approach for monitoring and surveillance systems in unregulated markets
- Improving the change management process: executive summary
- Statistical aspects of the portfolio construction programme
- Solar energy applications in the Yemen Arab Republic
- MNC overseas subsidiaries in Japan: in search of centres of excellence
- Model development for the adoption of technology: electronic commerce in the construction industry
- The usefulness of analysts’ cash flow forecasts and analysts’ earnings forecasts excluded by I/B/E/S

http://flax.nzdl.org/greenstone3/flax?a=b&rt=r&s=ClassifierBrowse&cl=CL1&c=PASS&if=flax
 Wikification of Specialised Terminology

**PhD Abstracts (Social Sciences)**

**Modelling, forecasting and riding credit risk in the Sterling Eurobond market**

This thesis aims to make a contribution to the understanding of credit risk dynamics in the Sterling Eurobond market. The background to the thesis is the increasing size, complexity and Volatility (finance) of all debt markets, where the tasks of measuring, understanding and forecasting credit risk are of central importance to Investment institutions and corporate and sovereign borrowers. We investigate the changes in the perceived Credit rating of bond issuers through three different approaches. First, we describe the evolution of Credit spread (bond) over time, exploring whether they reflect economic fundamental, or whether they represent self-generated force. This question is central to the fixed income literature in general, and to the pricing of risky debt and Credit derivative in particular. The Time series properties of our Credit spread (bond) provide strong evidence of Mean reversion (finance), Nonlinear system, and directional and persistent volatility. All these stylised facts are captured by time-varying Stochastic volatility. Second, we assess the information value of bond ratings, by examining the dynamics of bond spreads around rating revision dates. In contrast to standard event studies we apply a novel Autoregressive conditional heteroskedasticity to the panel data. This lets us examine the effects of the regrading event on the dynamics of bond spreads. The event upgrades are not. An asymmetric pattern is also observed. A third approach, the Estimator, allows us to disentangle the effects of default probability, and credit risk. The forecasting models that have been used in the above analysis have been found on forecasting the downgrades, and for the pricing of credit risk instruments.

**Source**

In econometrics, “Autoregressive Conditional Heteroskedasticity” (ARCH) models are used to characterize and model observed time series. They are used whenever there is reason to believe that, at any point in a series, the terms will have a characteristic time, or variance. In particular ARCH models assume the variance of the current Errors and residuals in statistics or innovation (signal processing). Innovation to be a function of the actual sizes of the previous time periods' error terms: the variance is related to the squares of the previous innovations.

Related topics in Wikipedia:
- Autoregressive moving average model
- Conditional variance
- Autoregressive model
- White test
- Full Bayesian estimation

Exercise 1

- regulatory
- legitimate
- law
- illegal
- governing
- constitutional
- legal
- criminal

_____ system

_____ advice

_____ aid

_____ action

face _____ action

considering _____ action against

take _____ advice

Type in your guessed word: legitimate legal

Check answer

play
download exercises
RESEARCH INTO MOOC LINGUISTIC SUPPORT
MOOC participants

“MOOC participants register for educational courses; they do not sign up as language learners.” (Wu, Fitzgerald & Witten, 2014)

• English Common Law MOOC (Coursera and University of London)
• CopyrightX (Harvard Law School, HarvardX and the Berkman Centre for Internet & Society)
• ContractsX (Harvard Law School and edX)
Q11 When you want to find out how to express something in English what resource(s) do you use? You may select more than one.

Answered: 103  Skipped: 9
You’ve got some gall

This behaviour is referred to as **GALL, Google-Assisted Language Learning** (Chinnery, 2008), and this activity mimics the majority of online search behaviour for information retrieval.
## Main reasons for using FLAX

<table>
<thead>
<tr>
<th>Reason</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>To assist me with finding specific subject information to help me prepare for e.g. tutorials, note-taking, forum discussions and exams in CopyrightX</td>
<td>68.48% 63</td>
</tr>
<tr>
<td>To assist me with modelling how to discuss (e.g. in forums, tutorials) in CopyrightX</td>
<td>15.22% 14</td>
</tr>
<tr>
<td>To assist me with modelling how to complete written assessments (e.g. exams) in CopyrightX</td>
<td>29.35% 27</td>
</tr>
<tr>
<td>To compare different resources in the FLAX collections to see and compare how specific legal terms and concepts are used</td>
<td>0.00% 0</td>
</tr>
<tr>
<td>To link to further resources (e.g. Wikipedia, other language corpora/collections) to learn how specific legal terms and concepts are used in different contexts other than the CopyrightX materials</td>
<td>59.78% 55</td>
</tr>
<tr>
<td>To save and learn specific legal terms and concepts in CopyrightX through the Cherry Picking Basket function in the FLAX collections</td>
<td>43.48% 40</td>
</tr>
<tr>
<td><strong>Other (please specify)</strong></td>
<td><strong>Responses</strong></td>
</tr>
</tbody>
</table>

Total Respondents: 92
Searchability rates highly

CopyrightX (Harvard University)

Search Words in Collection

You may be interested in create, creation, creator, creativity, recreate

Search Result: 142 sentences found

Group by patterns

- There remains a narrow category of works in which the creative spark is utterly lacking or so trivial as to be virtually nonexistent.
- By now, you should be familiar with the rules governing what types of creative works are subject to copyright protection, who owns the copyrights on those things, and how one secures and transfers a copyright.
- Surprisingly often, evidence survives of a defendant's conscious, intentional use of the plaintiff's work during the course of his own creative processes, and that evidence turns up in discovery.
- As you might expect, when expanding or revising copyright law, lawmakers frequently deploy arguments concerning how much legal protection creative works ought to get.

So now let's turn to the topic for today, which, as I've indicated, is copyright theory. As I just mentioned, by theory, I mean nothing more than arguments concerning when and why copyrights should be created and what the scope or limits of those rights should be. Before zeroing in on the specific theories that will occupy us during this lecture, I should say a bit more about what these theories do and why they matter. Here's the first and most obvious role of theory. As you might expect, when expanding or revising copyright law, lawmakers frequently deploy arguments concerning how much legal protection creative works ought to get. When engaged in debates of these sorts, lawmakers or their advisers sometimes advert directly to formal theories developed by economists, political theorists, and philosophers that address that issue, and even more often, to less formal variants of those theories in general circulation in popular discourse. The result is that, in order to understand how copyright law has assumed over time its current form and how it's likely to evolve in the future, you need to know, among other things, the content of the theories that the lawmakers in practice attend to.
- This description of Sun's creative process is consistent with the evidence presented at trial.
- 183 [14] Amici McNeal and Sutphin explain that "a quick examination of other programming environments shows that creators of other development platforms provide the same functions with wholly different creative choices."

http://tinyurl.com/h7sddkr
"thanks, Alannah, for this FLAX site for the Contract Law course. i discovered your work a bit late (one month after you posted your announcement of the FLAX site). at the beginning, i was jumping ahead, skipping lectures, and found myself not knowing certain terms (e.g., dead weight loss). i had to go back to the beginning, searching for where Prof. Fried defined this term. if i had known about your FLAX site, it would be of great help. in general, your FLAX site would be great for a review, or for someone searching for certain case. surely, in the future, in case i need to review some concepts of Contract Law, i would hit your FLAX site first. again, great work. thanks."
...as does the linking in of Wikipedia for glossing and expanding knowledge
...and the saveability of useful language patterns (collocations, lexical bundles)

http://tinyurl.com/zx4wix
References


Thank You

FLAX Language Project & Software Downloads: http://flax.nzdl.org/
The How-to eBook of FLAX:
FLAX Game-based Apps for Android via Google Play Store (free):
https://play.google.com/store/apps/developer?id=FLAX%20TEAM&hl=en

Ian Witten (FLAX Project Lead): ihw@cs.waikato.ac.nz
Shaoqun Wu (FLAX Research and Development): shaoqun@waikato.ac.nz
Alannah Fitzgerald (FLAX Open Language Research): a_fitzg@education.concordia.ca

TOETOE Technology for Open English Blog: www.alannahfitzgerald.org
Slideshare: http://www.slideshare.net/AlannahOpenEd/
Twitter: @AlannahFitz