

# A Text Analytic Approach to Rural and Urban Legal Histories

*Adam Wyner* - School of Natural and Computing Sciences

*Jackson Armstrong and Andrew MacKillop* - School of Divinity, History, and Philosophy

*Phil Astley* – Aberdeen City & Aberdeenshire Archives

1 4 9 5



UNIVERSITY  
OF ABERDEEN

Royal Society of Edinburgh

February 14, 2014

Edinburgh, Scotland



ABERDEEN  
CITY COUNCIL

# Project Specifics

- Collaboration between Computing, History, and Aberdeen Council.
- Funded by University of Aberdeen's *dot.rural partnership resource for digital economy*
- About £53k for six months.

# Technical Objectives

## *Text Analytics*

- To apply automatic text analytic tools to annotate and enrich transcribed, machine-readable **Aberdeen Burgh Records from 1530-1531** (output of RISS Project).
- *Deep semantic, conceptual annotation.*
- To **query, extract, link, and reuse** the corpus content.
- Advantages: flexible, extensible, reapplicable, fast, efficient, multi-dimensional.

# Historical Objectives

## *Digital Humanities*

- To carry out deep **conceptual analysis** based on textual data:
  - Named entities, relations, legal roles, jurisdictions....
  - Grammatical information for Latin and Middle Scots.
  - Trace social/legal networks.
  - Follow the evolution of social/legal concepts.
  - Highlight historical resource management and conflict.
  - Who, what, when, where, why, how, what connections...?
- To address the questions using **semantic search within and across texts and in textual context.**

# Council and Archivists' Objectives

## *Public Engagement*

- To increase **public access** to archival materials for tourism, education, business, and research.
- To **draw out, augment, and link** the content of the materials.

# Text Analysis

- We focus on enriching the source text with annotations in order to facilitate semantic search.
- General Architecture for Text Engineering (GATE) - open source framework which supports plug-in NLP components to process a corpus of text.
- Human visible and machine-readable formats.
- Indexation and search with respect to annotations.

# Annotations - Human View

Printed Source: E. Gemmill, ed., **Aberdeen** **Guild Court** Records (2005), pp. 88-91

MS source: **Aberdeen** Council Register V(2), pages as numbered 689 – 691

[689]

**Curia** **gilde** huius **burgi** tenta per **Johannem de Marr**, prepositum, vijmo die mensis Februarii, anno Domini, etc., xlij [=7 Feb 1442/3, session of the **guild court**]

Quo die **Johannes Bertlotson** primo die vocatus ad intrandum coram **preposito** non comparauit, etc.

**Robertus Blyndseil** secundo die vocatus ad intrandum **Willelmum Bird**, hominem suum, etc.

**Gilbert Meignes**, **William Scherar**, **Thomas Blyndseil** and **Joh[n]n[e] Gray**, arbitra[irs] chosin betuex

**Williame Voket**, eldar' and **Gilbert Coly** apon the debate movit betuex thaim because of a certane pak of wal[d] thai have **ordanit** that how mykil is sald of the wald the mo[ne] of it sal be partit evynly amangis thaim, and the remanand of the wald to be partit right swa and at this be done under gude fai[th]

Item, because that **Matheu Fichet** said it was cum til his understand[ing] that **Maroiune of Crondane** has complengnit to the **lord of Erole** that he had al the wite and wilfully put hir fra hir' watt[eris] the quhilkis scho had of befor, quhar'for he askit a declaracion of the **counsaile** quethir he was cause of the forsaid mater or nocht, and than he removit and the **counsaile** habely decretit that he had made na sik cause na he had na wite tharof.

Thir' ar' the names

Type	Set	Start	End	Id	Features
Location		33	41	722	{rule=Location}
RegisterEntry		42	53	723	{rule=RegisterEntry}
Location		91	99	724	{rule=Location}
LegalBody		157	162	725	{rule=LegalBody}
LegalBody		163	168	726	{rule=LegalBody}

- IdeologicalTerm
- LegalBody
- LegalConcept
- LegalRole
- Location
- Lookup
- Name
- Offence
- Offender
- Office
- RegisterEntry
- ▶ Original markings

# Annotations - Machine View (XML)

```
<html><head><meta content="application/vnd.openxmlformats-officedocument.wordprocessingml.document"
  name="Content-Type"><meta content="Gemmill extract.docx" name="resourceName"><title></title></
  meta></meta></head><body><p>Printed Source: E. Gemmill, ed., <i><Location>Aberdeen</Location> <
  RegisterEntry>Guild Court</RegisterEntry> Records </i>(2005), pp. 88-91</p>
<p>MS source: <Location>Aberdeen</Location> Council Register V(2), pages as numbered 689 - 691</p>
<p><a name="_GoBack"></a></p><p>[689]</p>
<p><LegalBody>Curia</LegalBody> <LegalBody>gilde</LegalBody> huius <LegalBody>burgi</LegalBody>
  tenta per <Name>Johannem de Marr'</Name>, <Office>prepositum</Office>, vijmo die mensis
  Februarii, anno Domini, etc., xlij [=7 Feb 1442/3, session of the <RegisterEntry>guild court</
  RegisterEntry>]</p>
<p>Quo die <Name>Johannes Bertlotson'</Name> primo die vocatus ad intrandum coram
  <Office>preposito</Office> non comparavit, etc.</p>
<p><Name>Robertus Blyndseil</Name> secundo die vocatus ad intrandum <Name>Willelmum Bird</Name>,
  hominem suum, etc.</p>
<p><Name>Gilbert Meignes</Name>, <Name>William Scherar'</Name>, <Name>Thomas Blyndseil</Name> and <
  Name>Joh[n]n[e] Gray</Name>, arbitra[irs] chosin betuex <Name>Williame Voket</Name>, eldar' and
  <Name>Gilbert Coly</Name> upon the debate movit betuex thaim because of a certane pak of wal[d]
  thai have <RegisterEntry>ordanit</RegisterEntry> that how mykil is sald of the wald the mo[ne]
  of it sal be partit evynly amangis thaim, and the remanand of the wald to be partit right swa
  and at this be done under gude fai[th]</p>
<p>Item, because that <Name>Matheu Fichet</Name> said it was cum til his understand[ing] that
  <Name>Maroiune of Crondane</Name> has complengnit to the <Name>lord of Erole</Name> that he had
  al the wite and wilfully put hir fra hir' watt[eris] the quhilkis scho had of befor, quhar'for
  he askit a declaracion of the <Office>counsaile</Office> quhethir he was cause of the forsaid
  mater or noght, and than he removit and the <Office>counsaile</Office> habely decretit that he
  had made na sik cause na he had na wite tharof.</p>
```



# Search for Relations

{Name}({Token})\*5{Office}

Corpus: Entire datastore

Annotation set: All

Results:

Context size:



Search



Clear



Next page of 50 results

Context

of the contribucion to my Elizabeth' of Britane the hale counsaile has ordanit that it be

Token.string

of the contribucion to my Elizabeth' of Britane the hale counsaile has ordanit that it be

Name

Elizabeth'

Token

of the contribucion to my Elizabeth' of Britane the hale counsaile has ordanit that it be

Office

of Britane

Page 1 (4 results)



Export

Global

Left context	Match	Right context	Features
of the contribucion to my	Elizabeth' of Britane the hale counsaile	has ordanit that it be	Token.string=Elizabe
session of common council] that	Robert Blindzelle and Will' Woket, ball'	and ressawaris of the Martymes	Token.string=Robert
pay the xxx lib' to	John of Mar' alderman	to wirk the commownn werk	Token.string=John, c
under[wrytyn] sal be awd[itouris] to	John' of Fiffis cwn...ice of aldermanschip	of the yer. In primis	Token.string=John', .

Annotation Type	Count
Token	33120
Lookup	143
Name	64
P	28
Office	18
Location	16
RegisterEntry	15

# Challenges

- Reuse and develop Latin and Middle Scots NLP tools such as dictionaries, parsers, grammars, etc.
- Determine how to process irregularities.
- Structure the terminology into concepts.
- Identify and select external, linkable content.
- Demonstrate use of text analysis for legal historical analysis and for public access.

# Beyond and Contact

- Link to external web resources.
- Provide a web-based interactive interface.
- Create more projects. [Link to more projects.](#)
- Adam Wyner – [azwyner@abdn.ac.uk](mailto:azwyner@abdn.ac.uk)