Annotating and Querying Content within Machine-readable Legal Instruments

Dr. Adam Wyner
School of Law and Department of Computer Science
Swansea University, UK

Office of the Parliamentary Counsel, London May 21, 2019

Path

- Ask some questions
- Show our result by way of example
- Outline why and how we did it by way of example
- Inquire what you would like to do:
 - What questions would you like to answer with respect to your documents?
 - How would you like others to query your documents?
- Highlight some issues and problems

Handout

- Some questions before and after.
- Sample materials to annotate and query.
- Something to discuss together.
- An opportunity to share information and for me to learn from you.
- I would appreciate collecting the pages after you have made your notes.
- Add your name and email address, if you would be interested to establish some contact.
- Use additional paper if needed.

Part I

Questions and Result

Questions - before

- 1. What texts do you work on? Material.
- 2. What would you need/like to do with the materials? Aim.
- 3. What do you do with the materials to achieve the Aim? Action.
- 4. Who 'queries' the materials for what aims? Agents.

Pilot Project Scottish Smoking Legislation

- 1. Material small, cohesive corpus
- Aim find Scottish Ministers powers, compare clauses across document, check consistency of expression
- 3. Action annotate and query content
- 4. Agents drafters (how get help to revise, analyse, and relate), the public (what is in the law), businesses (how does the law bear on processes and compliance), other government agencies (how is the law relevant to my department)

On Annotation of the Textual Contents of Scottish Legal Instruments

Adam WYNER [a] 1 Fraser GOUGH [c], Francois LEVY [b], Matt LYNCH [c], and Adeline NAZARENKO [b]

^a University of Aberdeen, Aberdeen, United Kingdom

^b LIPN, Paris 13 University – Sorbonne Paris Cité & CNRS, Paris, France

^c Parliamentary Counsel Office, Scottish Government, Edinburgh, United Kingdom

Smoking Corpus

Scope so it is manageable and comprehensible, then scale up.

- 2010ActAsAmended.docx Tobacco and Primary Medical Services (Scotland) Act 2010 asp 3
- asp_20100010_en.docx Interpretation and Legislative Reform (Scotland) Act 2010
- asp_20160003_en.docx Smoking Prohibition (Children in Motor Vehicles) (Scotland) Act 2016
- asp 20160014 en.dox Health (Tobacco, Nicotine etc. and Care) (Scotland) Act 2016
- criminal procedure 1995 46 section 225.docx Criminal Procedure (Scotland) Act 1995
- ssi_20100406_en.docx The Sale of Tobacco (Prescribed Document) Regulations 2010ssi_20100407_en.docx The Sale of Tobacco (Register of Tobacco Retailers) Regulations 2010
- ssi_20110023_en.docx The Sale of Tobacco (Registration of Moveable Structures and Fixed Penalty Notices) (Scotland)
 Regulations 2011
- ssi 20110132 en.docx The Sale of Tobacco (Display of Warning Statements) (Scotland) Regulations 2011
- ssi_20130085_en.docx The Sale of Tobacco (Display of Tobacco Products and Prices etc.) (Scotland) Regulations 2013
- ssi_20170050_en.docx The Sale of Tobacco and Nicotine Vapour Products by Persons Under 18 (Scotland) Regulations 2017

Smoking Prohibition (Children in Motor Vehicles) (Scotland) Act 2016

2016 asp 3

The Bill for this Act of the Scottish Parliament was passed by the Parliament on 17th December 2015 and received Royal Assent on 21st January 2016

An Act of the Scottish Parliament to prohibit smoking in private motor vehicles in the presence of children, subject to limited exceptions; and for connected purposes.

- (1) It is an offence for an adult to smoke in a private motor vehicle when-
 - (a) there is a child in the vehicle, and
 - (b) the vehicle is in a public place.
- (2) Subsection (1) does not apply to a private motor vehicle that is designed or adapted for use as living accommodation and which, at the time the smoking occurs, is parked and is being used as living accommodation.
- (3) A person who commits an offence under subsection (1) is liable on summary conviction to a fine not exceeding level 3 on the standard scale.

2 Fixed penalties

The schedule provides for fixed penalties for offences under section 1.

3 Powers to enter and require information

- (1) An <u>authorised</u> officer of a local authority may enter and search a private motor vehicle which is parked within the area of the local authority in order to ascertain whether an offence under section 1 has been or is being committed there.
- (2) A person who—
 - (a) an authorised officer of a local authority reasonably believes-
 - (i) is committing or has committed an offence under section 1, or
 - (ii) has information relating to such an offence, and
 - (b) fails without reasonable excuse to supply the officer with the person's name and address on being so required by the officer,

Sample Queries

Scope so it is manageable and comprehensible, then scale up.

- What are the offences and associated penalties or defences?
- What prohibitions apply to tobacco products?
- What obligations have been placed on what entities, e.g. shop owners?
- What permissions are given to Scottish Ministers?
- Given a provision, what are related overriding or reparation provisions?

Smoking Prohibition (Children in Motor Vehicles) (Scotland) Act 2016

The Bill for this Act of the Scottish Parliament was passed by the Parliament on 17th December 2015 and received Royal Assent on 21st January 2016

To answer the questions:

How do we classify portions of the text?

How can we search for strings within textual portions?

How do we relate portions of the text one to the other?

An Act of the Scottish Parliament to prohibit smoking in private motor vehicles in the presence of children, subject to limited exceptions; and for connected purposes.

- (1) It is an offence for an adult to smoke in a private motor vehicle when-
 - (a) there is a child in the vehicle, and
 - (b) the vehicle is in a public place.
- (2) Subsection (1) does not apply to a private motor vehicle that is designed or adapted for use as living accommodation and which, at the time the smoking occurs, is parked and is being used as living accommodation.
- (3) A person who commits an offence under subsection (1) is liable on summary conviction to a fine not exceeding level 3 on the standard scale.

2 Fixed penalties

The schedule provides for fixed penalties for offences under section 1.

3 Powers to enter and require information

- (1) An authorised officer of a local authority may enter and search a private motor vehicle which is parked within the area of the local authority in order to ascertain whether an offence under section 1 has been or is being committed there.
- (2) A person who-
 - (a) an authorised officer of a local authority reasonably believes-
 - (i) is committing or has committed an offence under section 1, or
 - (ii) has information relating to such an offence, and
 - (b) fails without reasonable excuse to supply the officer with the person's name and address on being so required by the officer,

Annotations

Scope so it is manageable and comprehensible, then scale up.

- Clauses and relations in the smoking corpus has been annotated with:
 - Permission (prescriptive norm)
 - Obligation (prescriptive norm)
 - Prohibition (prescriptive norm)
 - Constitutive (definition)
 - Override (expresses 'exception')
 - Reparation (a relation between a prescription and a penalty; penalty is triggered when the prescription is violated)
 - Penalty
 - Number as index (optional)
- Definitions from the LegalRuleML Core Specification document.

Model

The set of annotations (classifications of text and relations between classified text) constitute a

model of the text

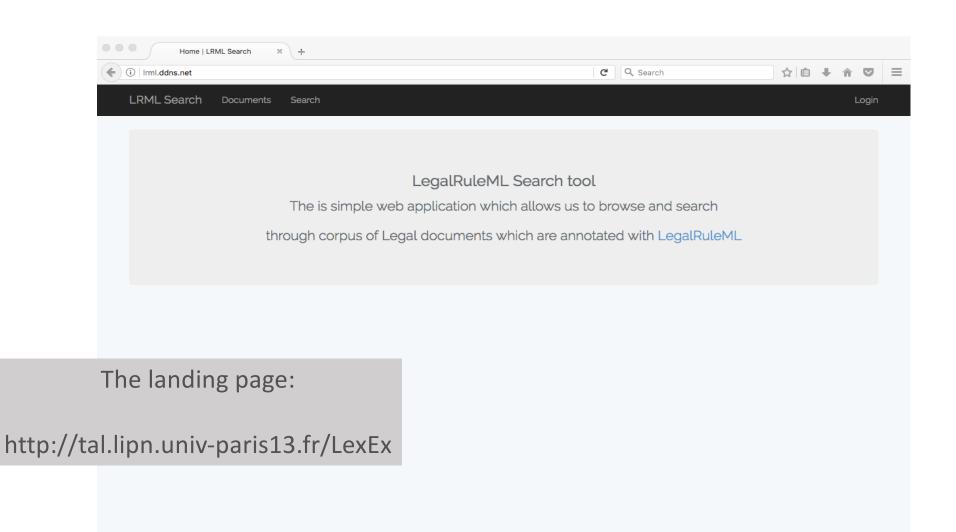
which represents an understanding of the meaning, relevance, and use of text.

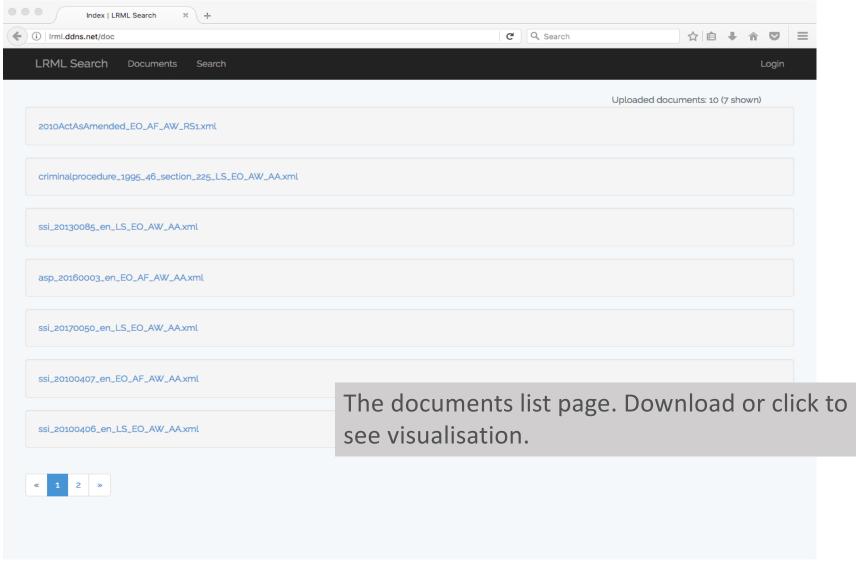
Annotating the text with respect to a model creates an instantiation of the model.

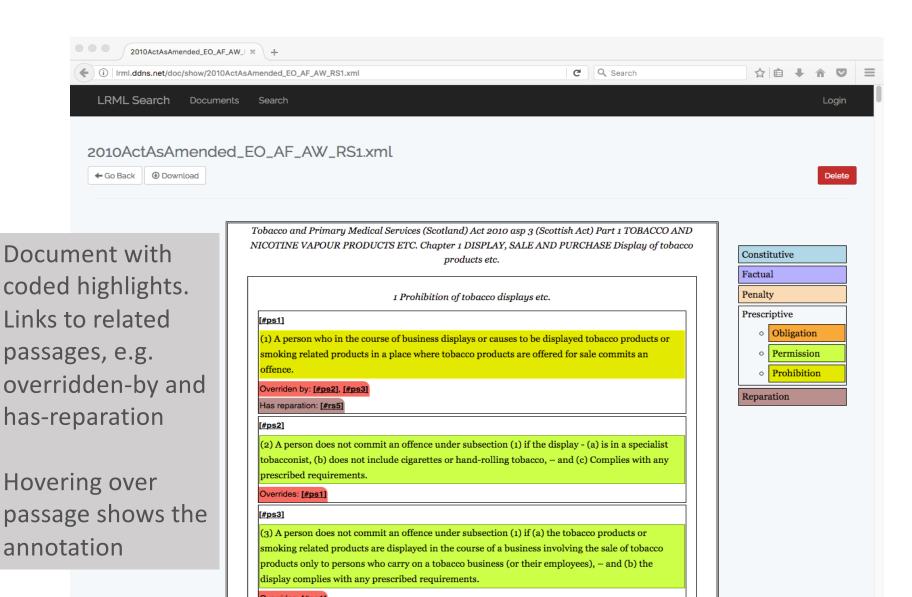
You can choose or create another model for other meanings, relevance, and uses.

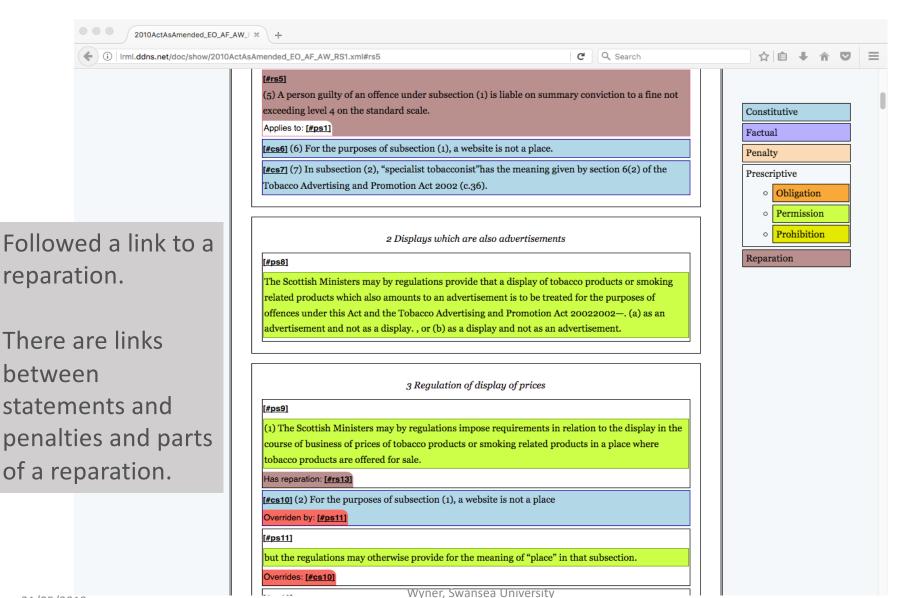
Visualisation and Search

- Searching for provisions in terms of semantic class and terminology.
- Finding connected provisions (or connected parts of a provision).
- Gathering information that is distributed across the corpus, which might otherwise be hard to find or understand.

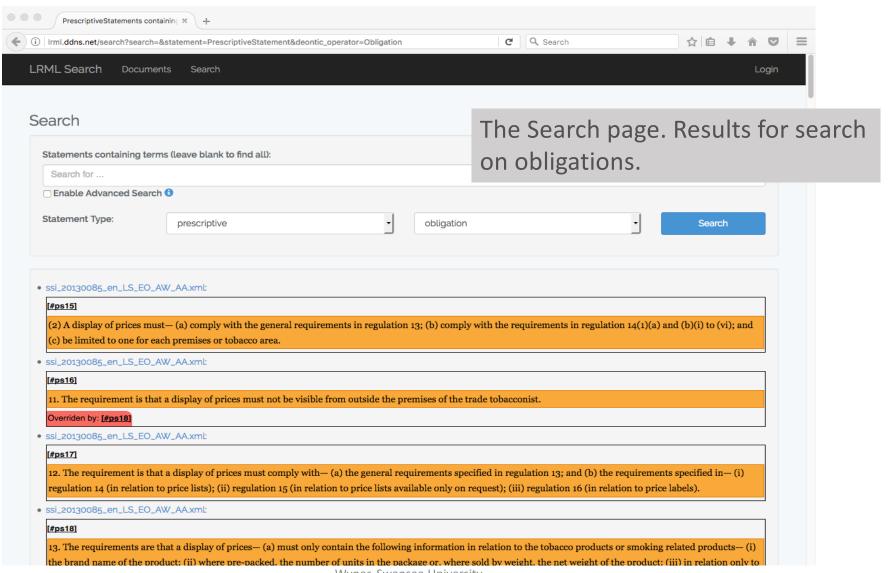


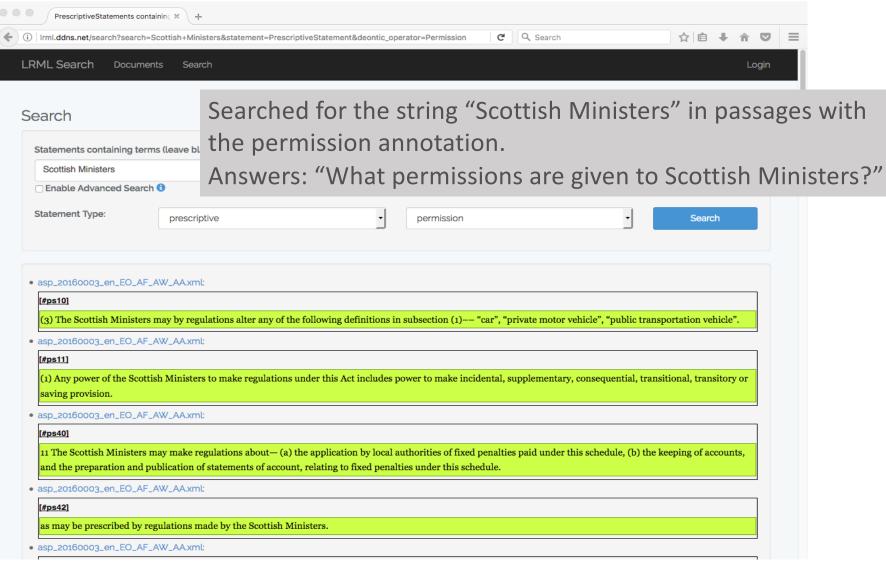


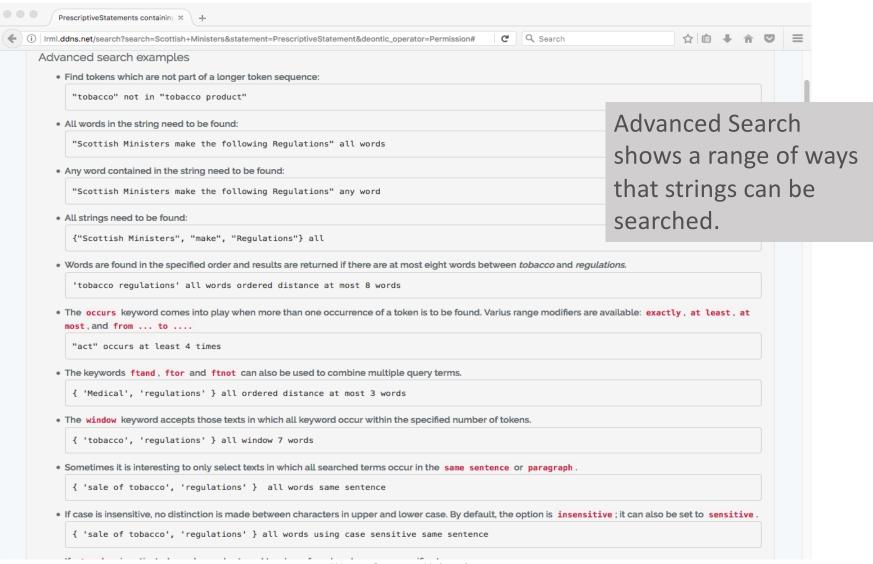


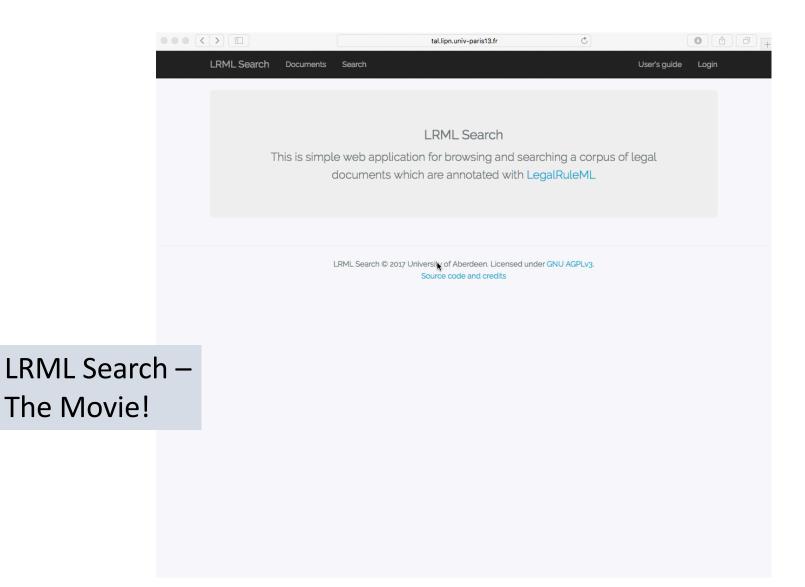


21/05/2019









Plus Point – Jurisdictional Distinctions between England and Wales

Issues about what laws are applicable within Wales and over which Wales has competence.

- The text in connection to jurisdictions is complex and dense.
- There are threads of dependencies, which may not always have an explicit reference.
- Jurisdictional applicability sometimes needs to be inferred.
- There are 'legal dead ends'.
- Want to be able to 'test out'

Can make a series of questions, annotations, and links to work with this. Encode drafters knowledge for drafters uses.

Thanks Terry Kowal, Welsh Govt. (errors or misrepresentations are mine)

Part I Summary

- Introduced some questions and how they can be addressed.
- Showed the corpus and text.
- Sampled the annotations.
- Illustrated some queries.

Part II

Why and How we did it

Rationale and Context

- The Scottish Government's Parliamentary Counsel Office are working to improve internal legislative drafting and information services as well as provide legislative information ``as a platform'' for a robust ecosystem of legal services.
- Problems:
 - Language and text structure
 - Volume and complexity of corpus
 - Distributed, yet interrelated, information
- A key part is to provide the law in electronically readable form.
- LegalDocML for document structure.
- LegalRuleML is tailored to semantic content.

Aim

- Translate legislation from source text to a formal, machine-readable language for rules, *LegalRuleML*.
- "Towards" automatic the translation.

Purposes/Uses

- Query, link, extract, and remix information from across a corpus.
- Reason with the rules automatically, e.g. from input to determinations (Oracle Policy Automation).

- Explore and understand the law.
- Help drafters and legislators to develop the law.
- Help business, government, and public to access the law.

What is LegalRuleML?

- A proposed OASIS open access/open-source standard XML markup language.
- Used to markup/annotate up a range of contentful information specific about rules in the legal domain.
- A bit of detail to come, but there is much more....

RuleML and LegalRuleML Elements

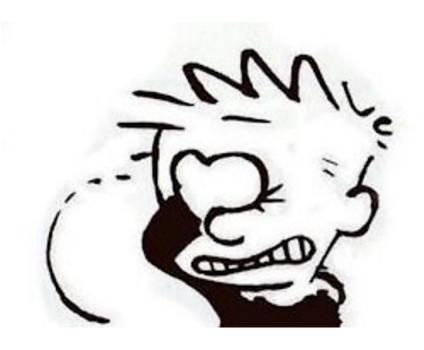
- Atom, if, then, Rule, Rel(ation), Ind(ividual), Formula, Var, slot, Time, Data, And,
 Or, Negation
- Agent, Authority, AuxiliaryParty, Bearer, ConstitutiveStatement, Context,
 DefeasibleStrength, FactualStatement, Jurisdiction, LegalSource, Obligation,
 Override, PenaltyStatement, Permission, Prohibition, Reparation, Right, Role,
 StrictStrength, TemporalCharacteristics, Violation

A person can engage in a credit activity if the person holds a licence.

Task: translate natural language text into the formal ML.

```
<lrml:PrescriptiveStatement key="ps2">
  <ruleml:Rule key=":rule2" closure="universal">
    <lrml:hasStrength>
      <lrml:DefeasibleStrength/>
    lrml:hasStrength>
    <ruleml:if>
      <rulem1:Atom>
        <rulem1:Rel iri=":hasLicence"/>
       <rulem1: Var>X</rulem1: Var>
      </ruleml:Atom>
    </ruleml:if>
    <ruleml:then>
      <lrml:SuborderList>
        <!rml:Permission>
          <ruleml:Atom>
            <ruleml:Rel iri=":engageCreditActivity"/>
            <rulem1:Var>X</rulem1:Var>
          </ruleml:Atom>
        </lrml:Permission>
      </lrml:SuborderList>
    </ruleml:then>
  </ruleml:Rule>
PrescriptiveStatement>
```

Problem is translation and TMI!!



Our Approach

- Start with 'lightweight' and 'high-level' semantic annotations.
- Fine-grained v coarse-grained. We go with coarse-grained.
- Goal-oriented to address user questions.
- Interesting point: still many (practical and interpretive) issues arise with coarsegrained annotations.

What We Did

- Manually annotate a corpus of Scottish legislation and regulation in a subpart of LegalRuleML.
- Transform to technical LegalRuleML.
- Once formally annotated, analysts can use a computer to query, link, reprocess, and transmit information.
- Facilitates analysis across the corpus.
- A small team of students, cooperation with legal drafters, Trello, Google Docs, annotation manual, limited annotations, development cycle, gold standard corpus....

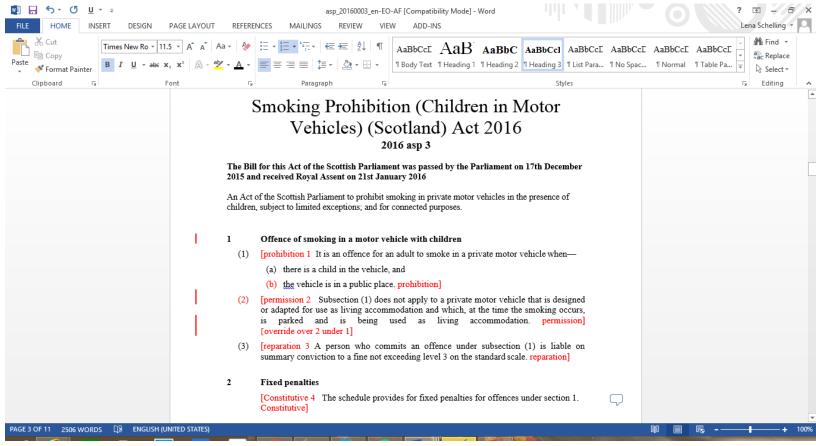
Manual - Initial Annotation

We annotated it as:

[prohibition 1 A person who in the course of business displays or causes to be displayed tobacco products or smoking related products in a place where tobacco products are offered for sale commits an offence. prohibition]

This is pretty close to what people do anyways in analysing text.

Manual – Example



To LegalRuleML – Why?

- Transform the Word document with annotations into LegalRuleML.
- On the way to search and extraction using computers.

LegalRuleML - Heading

```
<?xml version="1.0" encoding="UTF-8"?>
2 |<?xml-stylesheet type="text/css" href="stylesheet/css/style.css"?>
3 マ <lrml:LegalRuleML xmlns:lrml="http://docs.oasis-open.org/legalruleml/ns/v1.0/"
     xmlns:ruleml="http://ruleml.org/spec" xmlns:xm="http://www.w3.org/2001/XMLSchema-date
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://docs.oasis-open.org/legalruleml/ns/v1.0/ file:/home/rst/P
7
     <lrml:Comment> Tobacco and Primary Medical Services (Scotland) Act 2010 asp 3 (Scott)
      TOBACCO AND NICOTINE VAPOUR PRODUCTS ETC. Chapter 1 DISPLAY, SALE AND PURCHASE Disp
10
11
      products etc. </lrml:Comment>
12
13
     <!-- Legal Source -->
<lrml:LegalSource key="src1" sameAs="../Corpus/2010ActAsAmended_E0_0T.docx'"/>
15
     </lrml:LegalSources>
16
17
     <lrml:Statements key="stmts1">
18 🗢
      <lrml:Comment> 1 Prohibition of tobacco displays etc. </lrml:Comment>
19
20
```

LegalRuleML – Prohibition, Permission, and Override

```
21
      <!-- Prescriptive Statement: 1 -->
      <lrml:PrescriptiveStatement kev="ps1">
       <ruleml:Rule>
23 🗢
        <ruleml:then>
24 🗢
25 🗢
         <lrml:Prohibition>
          <lrml:Paraphrase> A person who in the course of business displays or causes to I
26 🗢
27
           tobacco products or smoking related products in a place where tobacco products
28
           for sale commits an offence. </lrml:Paraphrase>
29
         </lrml:Prohibition>
30
        </ruleml:then>
31
       </ruleml:Rule>
32
      </lrml:PrescriptiveStatement>
33
      <!-- Prescriptive Statement: 2 -->
34
      <lrml:PrescriptiveStatement key="ps2">
       <ruleml:Rule>
36 🗢
37 🗢
        <ruleml:then>
38 🗢
         <lrml:Permission>
          <lrml:Paraphrase> A person does not commit an offence under subsection (1) if tl
39 🗢
40
           (a) is in a specialist tobacconist, (b) does not include cigarettes or hand-ro
41
           - and (c) Complies with any prescribed requirements. 
         </lrml:Permission>
42
43
        </ruleml:then>
44
       </ruleml:Rule>
45
      </lrml:PrescriptiveStatement>
46
47 🗢
      <lrml:OverrideStatement>
       <lrml:Override over="#ps2" wwder-5"#ps2" University:Override>
48
      </lrml:OverrideStatementice of the Parliametary Counsel, London</pre>
49
```

21/05/2019

LegalRuleML

```
192
       <!-- Reparation Statement: 20 -->
       <lrml:ReparationStatement key="rs20">
193 🗢
        <lrml:Reparation>
194 🗢
         <lrml:Paraphrase> A person guilty of an offence under subsection (3) is liable on
195 🗢
          conviction to a fine not exceeding level 4 on the standard scale. </lrml:Paraphre
196
197
         <lrml:toPrescriptiveStatement keyref="#ps19"/>
198
        </lrml:Reparation>
       </lrml:ReparationStatement>
199
200
```

Various issues: only the prescription mentioned, not the related penalty.

LegalRuleML Virtues

- Extensible (add more parts of an analysis)
- Revisable (can change XML)
- Refinable (take a rough draft and refine it)
- Maintainable (a text file)
- Transmittable (available on the web)
- Linkable (connect one XML file to another anywhere on the web)

Part II Summary

- Indicated the rationale and aim.
- Pointed out some aspects of LegalRuleML and the complexity problem.
- Sketched our approach and annotation.
- Filled in a bit further the LegalRuleML representation of our text.
- Highlighted some advantages of the approach.

Part III

What questions would you like to answer with respect to your documents? How would you like others to query your documents?

Questions - after

- What texts do you work on? Material.
- 2. What would you need/like to do with the materials? Aim.
- 3. What do you do with the materials to achieve the Aim? Action.
 What would be your model of the text and corpus?
- 4. Who 'queries' the materials for what aims? Agents.

Other question that are not being asked here but ought to be asked?

- Evaluation
- Interpretation
- Context

Part IV

Issues and problems

Annotation - must

- On granting an application under section 11(3) the Scottish Ministers must issue to the applicant a certificate of registration in respect of each premises stated in the application.
- The list must be made available free of charge at all reasonable times.
- A person who carries on a tobacco business must display a notice in accordance with subsection (2) in any premises where that business is carried on.

Annotation - must

- In statements with 'must', when do we have an obligation and when a constitutive (definition) statement.
- Does every obligation statement have a correlated penalty for violation, or are some implicit penalties determined by judicial review? Epistemic v root interpretation of the modal.
- When 'must' and when 'commits an offence'?

Annotation - may

- (4) If it appears to the Secretary of State that there has been a change in the value of money since the relevant date, he may by order substitute for the sum or sums for the time being specified in the provisions mentioned in subsection (5) below such other sum or sums as appear to him justified by the change.
- (3) The Scottish Ministers may by regulations alter any of the following definitions in subsection (1)— "car", "private motor vehicle", "public transportation vehicle".

Annotation - may

• In statements with 'may', when do we have a permission and when a 'possibility'? A permission is something that one can do without interference. (epistemic v root interpretation of the modal)

Annotation – linguistic variants

• It is the duty of a council to enforce within its area the provisions of Chapters 1 and 2 and regulations made under them.

Why 'duty' and not 'obligation/must' or 'commits an offence'?

Annotation – linguistic variants

For the purpose of carrying out a council's functions under this Part, a council
officer is entitled to— (a) enter any premises (other than premises used only as a
private dwelling house)...

Why entitled to and not may?

Observation: the search tool could be helpful in drawing all similar terminological patterns together for comparison.

Annotation – prohibitions

 A right to enter any premises conferred by section 28 may be exercised only at a reasonable time.

A restricted right, which becomes a prohibition outside the scope of the right.

Annotation - reparations

• 6. The amount prescribed for the purposes of paragraph 4(1)(a) and (b) of schedule 1 to the Act for offences under sections 5 and 7 of the Act are respectively— (a) £50 (fixed penalty); and (b) £30 (discounted amount).

The treatment of reparations. These are relations between a prescriptive statement and a penalty, usually by references. References are addressed in LegalDocML. We have a partial representation.

Annotation - statement portions

• (2) [Constitutive 10 For the purposes of subsection (1), a website is not a place constitutive] [permission 11 but the regulations may otherwise provide for the meaning of "place" in that subsection. Permission]

Have tried only to annotate sentences. But there are instances where portions of sentences need to be annotated separately and in order to set interpretive scope.

Search

- 1. All offences by searching for the word "offence".
- 2. Defences are obtained by searching for "defence" or "offence".
- 3. Prohibitions which apply to tobacco products.
- 4. Obligations placed on shop owners.
- 5. Permissions given to "Scottish Ministers".

Others

- 1. Linking referential expressions.
 - "A person who fails to comply with a requirement made under subsection (1) or (2) commits an offence".
- 2. Annotation of whole sentences v parts (offence):
 - (1) A person who in the course of business displays or causes to be displayed tobacco products or smoking related products in a place where tobacco products are offered for sale commits an offence.

Next Steps

- Larger corpus or other corpora
- Adjudication of issues quality control
- Additional annotations:
 - Penalties
 - Conditional rule structures
 - Reparations
 - Agents
 - Additional LegalRuleML components
- Additional queries

Next Steps

- Expand network of collaborations and linking across jurisdictions.
- Other models of the text.
- More corpora and jurisdictions raises importance of common understandings and implementations.
 - Workshop on Legislation and Regulations on the Semantic Web 2019
 - https://sites.google.com/view/legregsw2019/home
- Project proposal.

Thanks for your attention!

- Contact:
 - Adam Wyner, a.z.wyner@swansea.ac.uk
- Comments, questions, suggestions....