Provenance trails

Which person is responsible for what information & with what conf?

Human Source

I saw…
I think…
In my opinion…
My sister says that
my brother thinks that they say…

NLP pipeline

Which NLP module is responsible for which interpretation & with what conf?

Knowledge Store

Storyteller

Software Source

Which NLP module is responsible for which interpretation & with what conf?
Software provenance trails
NLP Provenance through NAF

<?xml version="1.0" encoding="UTF-8"?>
<NAF xml:lang="en" version="v3">
  <nafHeader>  
  </nafHeader>
  <raw><![CDATA[Qatar Holding sells 10% stake in Porsche to founding families]]></raw>
  <text>  
  </text>
  <terms>  
  </terms>
  <deps>  
  </deps>
  <entities>  
  </entities>
  <coreferences>  
  </coreferences>
  <timeExpressions>  
  </timeExpressions>
  <srl>  
  </srl>
  <constituency>  
  </constituency>
  <temporalRelations>  
  </temporalRelations>
</NAF>
Qatar Holding sells 10% stake in Porsche to founding families
NLP provenance trail

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<NAF version="v3" xml:>
  <nafHeader/>
  <linguisticProcessors layer="causalRelations">
    <lp beginTimestamp="2017-01-20T10:13:22+0100" endTimestamp="2017-01-20T10:17:24+0100"
      hostname="r13n3.lisa.surfsara.nl" name="CausalRelPro-FBK" version="2.0.0"/>
  </linguisticProcessors>
  <linguisticProcessors layer="topics">
    <lp beginTimestamp="2017-01-20T10:07:52+0100" endTimestamp="2017-01-20T10:07:58+0100"
      hostname="r13n3.lisa.surfsara.nl" name="ixa-pipe-topic-en" version="1.0.3-40be8debb88093b426ae3520d60df60161968e27"/>
  </linguisticProcessors>
  <linguisticProcessors layer="deps">
    <lp beginTimestamp="2017-01-20T10:12:08+0100" endTimestamp="2017-01-20T10:12:27+0100"
      hostname="r13n3.lisa.surfsara.nl" name="ixa-pipe-srl-en" version="1.0"/>
  </linguisticProcessors>
  <linguisticProcessors layer="timeExpressions">
    <lp beginTimestamp="2017-01-20T10:12:30+0100" endTimestamp="2017-01-20T10:12:43+0100"
      hostname="r13n3.lisa.surfsara.nl" name="TimePro-FBK" version="v1.5, based on Yamcha'"/>
  </linguisticProcessors>
  <linguisticProcessors layer="srl">
    <lp beginTimestamp="2017-01-20T10:12:08+0100" endTimestamp="2017-01-20T10:12:30+0100"
      hostname="r13n3.lisa.surfsara.nl" name="ixa-pipe-srl-en" version="1.0"/>
  </linguisticProcessors>
</NAF>
Confidence scores

<entity id="e12" type="PER">
  <references>
    <span>
      <!--Bell-->
    </span>
  </references>

  <externalReferences>
    <externalRef conf="0.9999999" ref="dbp:Bell_Canada" source="spotlight_v1"/>
    <externalRef conf="1.0654381E-7" ref="dbp:Bell_Mobility" source="spotlight_v1"/>
    <externalRef conf="3.168422E-8" ref="dbp:Bell_Helicopter" source="spotlight_v1"/>
    <externalRef conf="4.680139E-11" ref="dbp:Bell_County,_Texas" source="spotlight_v1"/>
    <externalRef conf="2.153721E-11" ref="dbp:Bell_Records" source="spotlight_v1"/>
    <externalRef conf="1.451566E-12" ref="dbp:Bell_System" source="spotlight_v1"/>
    <externalRef conf="5.5151933E-13" ref="dbp:Ian_Bell" source="spotlight_v1"/>
    <externalRef conf="1.6138333E-17" ref="dbp:Bell_Aircraft" source="spotlight_v1"/>
    <externalRef conf="2.7044043E-20" ref="dbp:Bell,_California" source="spotlight_v1"/>
    <externalRef conf="8.301236E-27" ref="dbp:Bell,_New_South_Wales" source="spotlight_v1"/>
    <externalRef conf="1.0" reference="dbp:Bell_Canada" source="POCUS"/>
  </externalReferences>
</entity>
Natural Language Annotation Format (NAF)

Advantages:
- NAF in, NAF out principle
- Extendible, modular, flexible
- Can build on previous layers
- Can handle ambiguity and alternatives
- Based on separate standards: SYNAF, MAF, SEMAF, TIMEML
- Interoperable across different languages and different software solutions

Disadvantages:
- Error propagation
- Overlapping and conflicting information
- Span mismatches
- Various ways to combine information across layers
- Unstable standards, lack of standards for higher-level interpretations, e.g. SRL, NERC
### NewsReader NLP benchmarking

**Table 7: Benchmarking of NLP modules using standard metrics and datasets.**

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Spanish</th>
<th>Dutch</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NERC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SoA Reference</td>
<td>90.90</td>
<td>81.39</td>
<td>Desmet [74]</td>
<td>Zanoli [75]</td>
</tr>
<tr>
<td>SoA F₁</td>
<td>91.36</td>
<td>84.16</td>
<td></td>
<td>84.91</td>
</tr>
<tr>
<td>NewsReader F₁</td>
<td>84.16</td>
<td>87.72</td>
<td></td>
<td>87.20</td>
</tr>
<tr>
<td><strong>NED</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dataset</td>
<td>TAC 2011</td>
<td>TAC 2012</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>SoA Reference</td>
<td>Barrena [76]</td>
<td>Monahan [77]</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>SoA F₁</td>
<td>81.55</td>
<td>62.22</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NewsReader F₁</td>
<td>62.90</td>
<td>50.00</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>SRL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dataset</td>
<td>CoNLL 2009</td>
<td>CoNLL 2009</td>
<td>SoNaR</td>
<td>N/A</td>
</tr>
<tr>
<td>SoA Reference</td>
<td>Nugues [71]</td>
<td>Zhao [78]</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>SoA F₁</td>
<td>85.63</td>
<td>80.46</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>NewsReader F₁</td>
<td>85.63</td>
<td>79.91</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td><strong>TIMEX detection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dataset</td>
<td>TempEval3</td>
<td>TempEval3</td>
<td>N/A</td>
<td>Evalita 2014</td>
</tr>
<tr>
<td>SoA Reference</td>
<td>Lee [? ]</td>
<td>Strötgen [51]</td>
<td></td>
<td>Mirza [53]</td>
</tr>
<tr>
<td>SoA F₁</td>
<td>83.10</td>
<td>85.33</td>
<td></td>
<td>82.70</td>
</tr>
<tr>
<td>NewsReader F₁</td>
<td>84.71</td>
<td>85.33</td>
<td></td>
<td>82.70</td>
</tr>
<tr>
<td><strong>TIMEX normalization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dataset</td>
<td>TempEval3</td>
<td>TempEval3</td>
<td>N/A</td>
<td>Evalita 2014</td>
</tr>
<tr>
<td>SoA Reference</td>
<td>Lee [? ]</td>
<td>Strötgen [51]</td>
<td></td>
<td>Manfred [79]</td>
</tr>
<tr>
<td>SoA F₁</td>
<td>82.40</td>
<td>85.33</td>
<td></td>
<td>70.90</td>
</tr>
<tr>
<td>NewsReader F₁</td>
<td>72.16</td>
<td>85.33</td>
<td></td>
<td>68.40</td>
</tr>
</tbody>
</table>
MEANTIME
120 Wikinews articles

Table 8: $F_1$ scores for out-of-domain benchmarking of NLP modules using MEANTIME.

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Spanish</th>
<th>Dutch</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>SoA Reference</td>
<td>Stanford NER</td>
<td>Stanford NER</td>
<td>Desmet [74]</td>
<td>Zanoli [75]</td>
</tr>
<tr>
<td>SoA $F_1$</td>
<td>66.96</td>
<td>47.48</td>
<td>48.44</td>
<td>46.85</td>
</tr>
<tr>
<td>NERC</td>
<td>70.90</td>
<td>62.14</td>
<td>63.93</td>
<td>46.85</td>
</tr>
<tr>
<td>NED</td>
<td>64.22</td>
<td>65.87</td>
<td>51.44</td>
<td>60.37</td>
</tr>
<tr>
<td>SRL</td>
<td>34.78</td>
<td>29.68</td>
<td>26.76</td>
<td>31.62</td>
</tr>
<tr>
<td>TIME detection</td>
<td>80.50</td>
<td>78.30</td>
<td>50.20</td>
<td>85.70</td>
</tr>
<tr>
<td>TIME normalization</td>
<td>68.50</td>
<td>62.20</td>
<td>41.90</td>
<td>64.60</td>
</tr>
</tbody>
</table>

- CoNLL (Reuter) and MEANTIME (Wikinews) same genre
- Annotation differences: spans, phrases or heads
- Different periods, styles
### Monosemous WSD errors due to PoS errors

<table>
<thead>
<tr>
<th>Competition</th>
<th>Monosemous</th>
<th>Wrong</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senseval2</td>
<td>499 (20.9%)</td>
<td>37.5%</td>
<td>gene.n (<em>suppressor_gene.n</em>), chance.a (<em>chance.n</em>) next.r (<em>next.a</em>)</td>
</tr>
<tr>
<td>Senseval3</td>
<td>334 (16.6%)</td>
<td>44.1%</td>
<td>Datum.n (<em>data.n</em>) making.n (<em>make.v</em>) out_of_sight (<em>sight</em>)</td>
</tr>
<tr>
<td>Semeval2007</td>
<td>25 (5.5%)</td>
<td>11.1%</td>
<td>get_stuck.v, lack.v, write_about.v</td>
</tr>
<tr>
<td>Semeval2010</td>
<td>31 (2.2%)</td>
<td>97.9%</td>
<td>Tidal_zone.n pine_marten.n roe_deer.n cordgrass.n</td>
</tr>
<tr>
<td>Semeval2013 (lemmas)</td>
<td>348 (21.1%)</td>
<td>1.9%</td>
<td>Private_enterprise, developing_country, narrow_margin</td>
</tr>
</tbody>
</table>
• The **US raises** import taxes
• Ford **exports** most cars from the **US**
• The **Ford crashed** into a **Toyota**

Table 2. Occurrences as Actor or Places per Class

<table>
<thead>
<tr>
<th>Class</th>
<th>Actor (total)</th>
<th>Place (total)</th>
<th>Actor (%)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>152,708</td>
<td>132,341</td>
<td>53.57%</td>
<td>40.93%</td>
</tr>
<tr>
<td>Person</td>
<td>106,318</td>
<td>9,447</td>
<td>91.84%</td>
<td>21.56%</td>
</tr>
<tr>
<td>Company</td>
<td>609,971</td>
<td>84,378</td>
<td>87.85%</td>
<td>29.63%</td>
</tr>
<tr>
<td>Motor Company</td>
<td>431,619</td>
<td>63,917</td>
<td>87.10%</td>
<td>28.46%</td>
</tr>
<tr>
<td>Automobile</td>
<td>116,893</td>
<td>5,804</td>
<td>95.27%</td>
<td>14.45%</td>
</tr>
</tbody>
</table>
Human provenance trails
Qatar Holding sells 10% stake in Porsche to founding families

Porsche family buys back 10pc stake from Qatar

fn:Commerce_money_transfer

type

fn:Buyer

fn:Sell

dbp:Porsche_family

fn:Goods

Entity23 10% stake

dbp:QatarHolding

sem:hasTime

2013-06-17
GRaSP
Model for Grounded Representation and Source Perspective
https://github.com/cltl/GRaSP
Porsche family buys back 10 pc stake from Qatar Holding

Qatar Holding sells 10% stake in Porsche to founding families

http://www.telegraph.co.uk

http://english.alarabiya.net
Pro-EU campaigners have hoped that big carmakers would also support the Remain campaign. In 2013 Carlos Ghosn, the chief executive of Nissan which makes cars in Sunderland, said of the Brexit debate: “If anything has to change, we would need to reconsider our strategy and our investments for the future.”
There is no such thing as a simple text

Eventualities

Carlos Ghosn, the chief executive of Nissan which makes cars in Sunderland, said of the Brexit debate: “If anything has to change, we would need to reconsider our strategy and our investments for the future.”
Pro-EU campaigners hoped carmakers would support. Carlos Ghosn, the chief executive of Nissan which makes cars in Sunderland, said of the Brexit debate: “If anything has to change, we would need to reconsider our strategy and our investments for the future.”
Pro-EU campaigners have hoped that big carmakers would also support the Remain campaign.
I won't go if Brexit happens says PM
British Prime Minister David Cameron said Friday he will step down after voters moved to withdraw from the European Union.
It all ends in tears: David Cameron stands down in wake of Brexit
# World stocks plunge on fears on US recession
# Last week, President Bush proposed a $145 billion stimulus plan to encourage more consumer spending.

```
{grasp

wikinews:World_stocks_plunge_on_fears_on_US_recession#char=13,19
  rdfs:label "plunge" ;
  gaf:hasAttribution w:wikinews:World_stocks_plunge_on_fears_on_US_recession/doc_attribute/Attr0.

wikinews:World_stocks_plunge_on_fears_on_US_recession/doc_attribute/Attr0
  rdf:value <CERTAIN_NON_FUTURE_POS> ;
  prov:wasAttributedTo w:wikinews:World_stocks_plunge_on_fears_on_US_recession.

wikinews:World_stocks_plunge_on_fears_on_US_recession
  prov:wasAttributedTo nw:author:Bawolff.

wikinews:World_stocks_plunge_on_fears_on_US_recession#char=1044,1053
  rdfs:comment "proposed" ;
  rdfs:label "encourage" ;
  gaf:generatedBy w:wikinews:World_stocks_plunge_on_fears_on_US_recession#char=1003,1011 ;
  gaf:hasAttribution w:wikinews:World_stocks_plunge_on_fears_on_US_recession/source_attribute/Attr2.

wikinews:World_stocks_plunge_on_fears_on_US_recession#char=1068,1076
  rdfs:comment "proposed" ;
  rdfs:label "spending" ;
  gaf:generatedBy w:wikinews:World_stocks_plunge_on_fears_on_US_recession#char=1003,1011 ;
  gaf:hasAttribution w:wikinews:World_stocks_plunge_on_fears_on_US_recession/source_attribute/Attr2.

wikinews:World_stocks_plunge_on_fears_on_US_recession/source_attribute/Attr2
  rdf:value <CERTAIN_u_POS>, <u_u_u>, <CERTAIN_NON_FUTURE_POS> ;
  gaf:wasAttributedTo dbp:Presidency_of_George_W._Bush.
```
<table>
<thead>
<tr>
<th>Event</th>
<th>Nr.</th>
<th>Perspective</th>
<th>Event</th>
<th>Nr.</th>
<th>Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>achieve</td>
<td>10</td>
<td>CERTAIN_u_POS</td>
<td>interest</td>
<td>8</td>
<td>CERTAIN_NON_FUTURE_POS</td>
</tr>
<tr>
<td>decision</td>
<td>6</td>
<td>CERTAIN_NON_FUTURE_NEG</td>
<td>interest</td>
<td>8</td>
<td>CERTAIN_u_POS</td>
</tr>
<tr>
<td>earnings</td>
<td>19</td>
<td>CERTAIN_FUTURE_POS</td>
<td>plans</td>
<td>9</td>
<td>CERTAIN_NON_FUTURE_POS</td>
</tr>
<tr>
<td>earnings</td>
<td>19</td>
<td>CERTAIN_NON_FUTURE_POS</td>
<td>plans</td>
<td>9</td>
<td>CERTAIN_u_POS</td>
</tr>
<tr>
<td>earnings</td>
<td>17</td>
<td>u_u_u</td>
<td>predict</td>
<td>17</td>
<td>CERTAIN_FUTURE_POS</td>
</tr>
<tr>
<td>focus</td>
<td>6</td>
<td>CERTAIN_FUTURE_POS</td>
<td>predict</td>
<td>17</td>
<td>CERTAIN_NON_FUTURE_POS</td>
</tr>
<tr>
<td>goal</td>
<td>7</td>
<td>CERTAIN_u_POS</td>
<td>predict</td>
<td>17</td>
<td>u_u_u</td>
</tr>
<tr>
<td>increase</td>
<td>18</td>
<td>u_u_u</td>
<td>stake</td>
<td>21</td>
<td>CERTAIN_u_POS</td>
</tr>
<tr>
<td>increase</td>
<td>15</td>
<td>CERTAIN_NON_FUTURE_POS</td>
<td>stake</td>
<td>19</td>
<td>CERTAIN_NON_FUTURE_POS</td>
</tr>
<tr>
<td>increase</td>
<td>11</td>
<td>CERTAIN_u_POS</td>
<td>stake</td>
<td>10</td>
<td>u_u_u</td>
</tr>
<tr>
<td>increased</td>
<td>9</td>
<td>CERTAIN_NON_FUTURE_POS</td>
<td>stake</td>
<td>19</td>
<td>CERTAIN_NON_FUTURE_POS</td>
</tr>
<tr>
<td>increased</td>
<td>8</td>
<td>CERTAIN_u_POS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Combing
Humand - Software provenance
Combining human & software provenance

- NLP: when, where
- WN3.0
- COREF v0.2
- CONLL 2011
- NED v0.5
- DBP 2011
- SRL v1.2
- NERC v1.2
- NED v0.5
- PROP BANK
- conf
- benchmark expectations
Digital humanities research ecosystem

Hypothesis
1
h-Scholars
2
Source & Data Selection

Source
Data
Hypothesis

Source
Data
Hypothesis

Source
Data
Hypothesis

World

World model

Silver Knowledge

Gold Knowledge

Crowd h-Scholars

Machine learning

Human annotation

cs-Scholars

Raw

RDF

Analytic tools

validation
evaluation

1
2
3a
3b
4a
4b
5a
5b
6
7
8
9

Digital humanities research ecosystem