Domeniul Lingvisticii Computaționale

Masterat la Facultatea de Informatică a UAIC
CyRo – build a technology that interprets old Cyrillic Romanian

- Train OCR classifiers to decode printed, semi-uncial and cursive Cyrillic Romanian documents

- Ambitious goals of a mixt consortium
  - library curators
  - paleolinguists
  - image processing exp.
  - computational linguists
Google Translate

Example based translation

Masa netă a camionului era de 120 tone.
Masa compactă de oameni s-a adunat în Piața Parlamentului.
Și atunci ne-am așezat cu toții la masa din sufragerie.

The net weight of the truck was 120 tonnes.
Compact mass of people gathered in Parliament Square.
And then we all sat at the dining room table.
CoRoLa – achievement of a large corpus of texts

Corpus of Contemporary Romanian Language

Portal
Concordances

(KWIC – Key Word In Context)

...
QuoVadis – a corpus of book characters and their relations

- H. Sienkiewicz’s *Quo Vadis*
- A corpus and a technology
  - The corpus: developed with master students in CL:
    - entities of type person and god
    - relations: coreference, kinship, affective, social
  - The technology: recognising entities and relations
Altogether – the text: an extremely rich network of links

Petronius...

Vinicius was the son of his oldest sister, who years before had married his father, a man of consular dignity from the time of Tiberius.
love and worship relationships in Quo Vadis
Entity linking

- Challenges in entity linking:
  - name variations, ambiguities, absence
  - entities
  - link type

knowledge base
I need help to remember all kinship relations between characters.
### Characters in Forsyte Saga

<table>
<thead>
<tr>
<th>The old Forsytes</th>
<th>Thg young Forsytes</th>
<th>Their children</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ann</strong>, the eldest of the family</td>
<td><strong>Young Jolyon</strong>, Old Jolyon's artistic and free-thinking son, married three times</td>
<td><strong>June</strong>, Young Jolyon’s defiant daughter from his first marriage; engaged to an architect, <strong>Philip Bosinney</strong>, who becomes Irene’s lover</td>
<td><strong>Val</strong>, Winifred and Montague’s son; fights in the Boer Wars; marries his cousin Holly</td>
</tr>
<tr>
<td><strong>Old Jolyon</strong>, the patriarch of the family, having made a fortune in tea</td>
<td><strong>Soames</strong>, James and Emily's son, an intense, unimaginative and possessive solicitor, married to the unhappy <strong>Irene</strong>, who later marries Young Jolyon</td>
<td><strong>Jolly</strong>, Young Jolyon's son from his second marriage; dies of enteric fever during the Boer Wars</td>
<td><strong>Imogen</strong>, Winifred and Montague's daughter</td>
</tr>
<tr>
<td><strong>James</strong>, a solicitor, married to <strong>Emily</strong>, a most tranquil woman</td>
<td><strong>Winifred</strong>, Soames’s sister, one of the three daughters of James and Emily, married to the foppish and lethargic <strong>Montague Dartie</strong></td>
<td><strong>Holly</strong>, Young Jolyon's daughter from his second marriage, to June's governess</td>
<td><strong>Others</strong></td>
</tr>
<tr>
<td><strong>Swithin</strong>, James's twin brother with aristocratic pretensions; a bachelor</td>
<td><strong>George</strong>, Roger's son, a dyed-in-the-wool mocker</td>
<td><strong>Jon</strong>, Young Jolyon's son from his third marriage, to Irene, Soames’s first wife</td>
<td><strong>Parfitt</strong>, Old Jolyon's butler</td>
</tr>
<tr>
<td><strong>Roger</strong>, &quot;the original Forsyte&quot;</td>
<td><strong>Francie</strong>, George's sister and Roger's daughter, emancipated from God</td>
<td><strong>Fleur</strong>, Soames's daughter from his second marriage, to a French Soho shopgirl <strong>Annette</strong>, Jon's lover; later marries a baronet, <strong>Michael Mont</strong></td>
<td><strong>Smither</strong>, Aunts Ann, Juley and Hester's housekeeper</td>
</tr>
<tr>
<td><strong>Julia (Juley)</strong>, a fluttery dowager; Mrs. Septimus Small</td>
<td></td>
<td></td>
<td><strong>Warmson</strong>, James and Emily’s butler</td>
</tr>
<tr>
<td><strong>Hester</strong>, an old maid</td>
<td></td>
<td></td>
<td><strong>Bilson</strong>, Soames's housemaid</td>
</tr>
<tr>
<td><strong>Nicholas</strong>, the wealthiest in the family</td>
<td></td>
<td></td>
<td><strong>Prosper Profond</strong>, Winifred's admirer and Annette's lover</td>
</tr>
<tr>
<td><strong>Timothy</strong>, the most cautious man in England</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Susan</strong>, the married sister</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The adult cerebral blood flow is about 750-1000 ml/min, representing 15-20% of the heart blood flow.
Tash Aw: Map of the invisible world

In books time does not flow linearly on...
Very often, the author brings us backward in time

- **[In the kitchen, waiting for the water to boil, Margaret wondered how was it that]**

- **[she jumped so naturally to embrace Adam.]**

- **[She could not realise what had made her]**

- **[feel so happy to see him again.]**
Our perception of time
From time segments to time yards
iCub – the robot trained in the Poeticon++ project
From semantics to pragmatics

A salad can be tossed
A spoon is a tool
A carrot can be cut
A spoon cannot cut things
Cut with spoon the carrot!

valid affordances
refused affordances
iCub decides by itself how to solve a command
Inițiată în 1993, EUROLAN, în colaborare cu Academia Română, a devenit un eveniment important la nivel internațional pentru cercetarea și dezvoltarea limbajului natural.

EUROLAN 2001, ediția a opta a acestui eveniment, a avut loc la București, România, cu participarea a numeroși școlitori și ștudiendi din diferite țări, precum și a numeroși invitați internaționali.

Evenimentul a avut ca obiective promovarea cercetării și dezvoltării limbajului natural, precum și promovarea cooperării internaționale în acest domeniu.

EUROLAN 2001 a oferit participanților o oportunitate unică de a încuraja colaborarea între științe și cercetare, precum și de a dezvolta și de a îmbunătăți înțelesul și utilizarea limbajului natural în diferite domenii, inclusiv cercetare, învățământ și comunicații.
Seria anuală de conferințe “Resurse lingvistice și instrumente pentru prelucrarea limbii române” din 2001

În colaborare cu Academia Română