

## Computational methods in Slavistic research and teaching

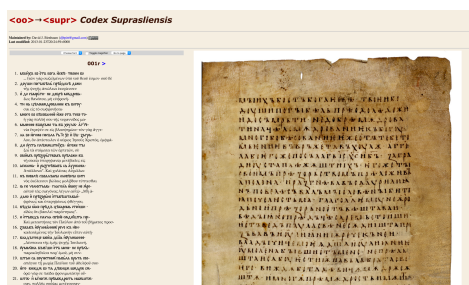
Kliment Oxridski University  
Sofia, 2016-06-15

David J. Birnbaum  
University of Pittsburgh  
djbpitt@gmail.com  
<http://www.obdurodon.org>

## Survey of projects

- Diplomatic editions of manuscripts: *Codex Suprasliensis*
- Critical editions of texts: *Rus' primary chronicle*
- Metadata and visualization: *The Repertorium of Old Bulgarian literature and letters*
- Textual transmission: *Daniel the Prisoner, Digenis Akritis*
- Distant reading: *Karolina Pavlova*
- Network analysis: *Menologies on line*
- Database: *Rusian genealogy*
- Language and prosody: *Meter, rhythm, rhyme*
- Language pedagogy: *Annotated Afanas'ev library*
- Learning DH: *Computational methods in the humanities*

## Codex Suprasliensis



## The Rus' primary chronicle

- Five principal witnesses
  - Text dates to early twelfth century
  - Oldest ms is *Lav* (1377)
- Full interlinear collation
  - Harvard UP, 2004; 3 v., 2368 p.
  - Aligned by “Karskij line”
  - Aligned by word

## Rus' primary chronicle

## “The story of Mary, Abraham’s niece”

- Part of the “Life of Abraham of Qidun”
  - Included in the *Paraenesis* of Efrem the Syrian
  - Syriac original, translated from Greek into Slavonic
- More than 40 known witnesses, most of the *Paraenesis*
- Slavic textology presupposes Greek textology
  - Cf. Mario Capaldo’s Greek parallels to the *Codex Suprasliensis*

### “The story of Mary, Abraham’s niece”

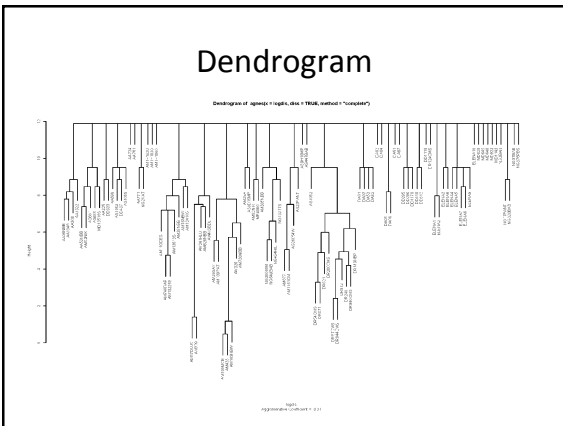
The screenshot shows a web interface for comparing manuscripts. At the top, it displays the title "The story of Mary, Abraham's niece" and a list of manuscripts under "Unit 3 (11-20)" and "Unit 4 (21-27)". The manuscripts are listed with their IDs, titles, and dates. A color-coded table below shows the relationships between these manuscripts, with colors indicating different levels of similarity or content overlap.

### The Repertorium

- *Repertorium of Old Bulgarian literature and letters*
- *Mixed-content miscellany* manuscripts are similar in the selection *and arrangement* of their contents
- Which manuscripts are similar to which other manuscripts in the corpus in their contents?
- How can we best visualize the relationships?
  - The old way: lists, tables
  - The new way: dendrograms, plectograms

### Tables

No. 29	1	2	3	4	5	6	7	8 <sup>II</sup>	9	10	11	12	13
No. 433	-	2	3	10	11	7	8	30 <sup>I</sup>	9	5	-	17	-
No. 149	3	4	5	12	13	9	10	32 <sup>I</sup>	11	7	-	19	-
No. 100	2	3	4	11	12	8	9	31 <sup>II</sup>	10	6	-	18	-
No. 794	31	21	3	4	5	6	7	8 <sup>II</sup>	22	12	-	-	26
No. 738	-	13	14	17	18	-	15	19 <sup>I</sup>	16	12	-	-	-
No. 326	-	4	5	10	6	-	8	7 <sup>I</sup>	9	11	3	18	-
Belovo	6	7	8	9	11	10	12	-	-	-	-	-	3



### Plectogram

- Trebniki (prayer books)
- Strong thematic unity
- Levels of coincidence
  - Individual articles
  - Sequences
- Microsequences

The plectogram shows two columns of manuscript IDs: DR671 XML (left) and DR672 XML (right). Lines connect corresponding IDs between the two columns, representing the relationships and coincidences between the two manuscript collections.

### Plectograms

- “[P]lectograms [...] are presented as working aids showing the degree of existing relationship, not as proposed genealogical models.” (125)
- “[T]he device lets us see at a glance the contours of the relationship. It allows us to give focus and priority to information that otherwise would be unwieldy; it allows us to set manageably small tasks for further work in strategic sampling of textual and other relationships.” (131)

### Mixed content miscellanies

- Very strong sequences
- Two pairs of texts appear twice each, in the same order and position
- Why twice?
  - Compilation from multiple sources
- Why twice in two manuscripts?
  - Shared textual transmission
  - Direct copying?
- Mind the gap
  - 0297 "Questions of St. John the Theologian to Abraham"
  - 0298 "Questions of St. John the Theologian to Abraham, fragment"

### Mixed content miscellanies

- Two closely related mixed content miscellanies and an outlier
- Similar content vs similar content *and* structure

### Mixed content miscellanies

### Mixed content miscellanies

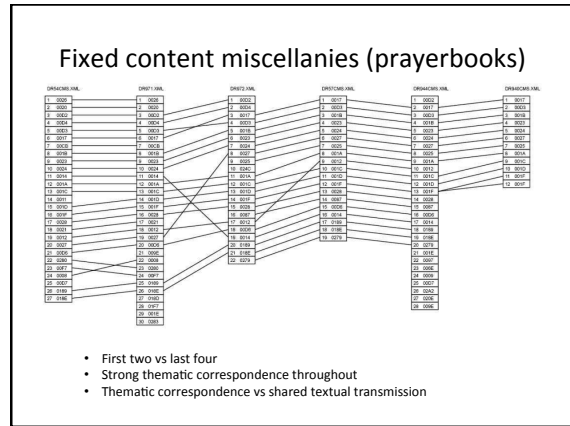
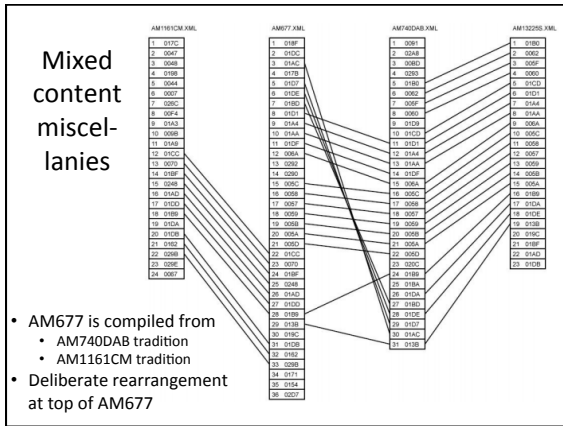
- "Twins"
- AM17DUJC is copied from AM309 (Miltanova)

### Mixed content miscellanies

- "Focusing" (Olmsted)
- Selection vs expansion

### Mixed content miscellanies

- Chiasmic correspondences
- Rearrangement and rebinding of NG649BAB



### Visualization

- Daniil Zatočnik (*Slovo* and *Molenie*)
- Digenis Akritis: Lay of the Emir
- Medieval saints as Facebook friends
- Karolina Pavlova, *A double life*

### Daniil Zatočnik

### Digenis Akritis: Lay of the Emir

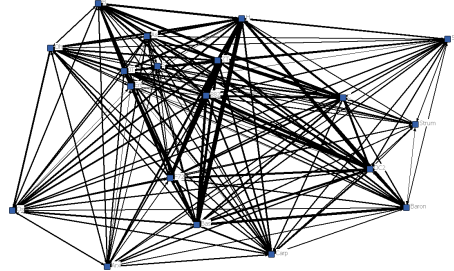
### Medieval saints as Facebook friends

- Gospel and Apostol manuscripts may include a calendar of saints (menology, synaxarion)
- Shared entries suggest shared textual transmission
- What is the “social network” described by shared entries?

## Facebook?!

- Manuscripts and saints are Facebook users
- Manuscripts “friend” saints by including them
- Manuscripts cannot friend manuscripts directly ...
  - ... but there’s an indirect network relationship based on the saintly friends they share
- Infer a weighted relationship between manuscripts

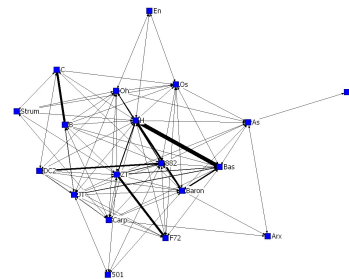
## The social network “hairball”



## “Loser” saints

- The social network “hairball”
  - Most saints occur in lots of manuscripts
  - Most manuscripts have roughly comparable connectivity
  - Too many lines, too poorly differentiated
- “Loser” saints have only two manuscript friends
  - Common shared values don’t differentiate within the corpus
  - Throw away data when it distracts from information

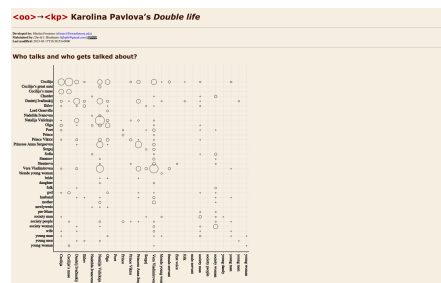
## The loser network



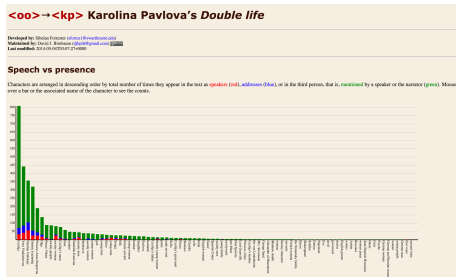
## Karolina Pavlova, *A double life*

- Cecilia, the protagonist, is the most present character in the novel.
- What is the relationship between presence and agency?
- Does she speak and act? Or is she present because others talk about her?
- How does her role change over the course of the novel?
- What, more generally, are the roles of men and women with respect to presence, speech, and agency? Old and young?
- <http://pavlova.obdurodon.org>

## Karolina Pavlova, *A double life*



# Karolina Pavlova, A double life



# Rusian genealogy

- Riurik genealogy
- Dynastic marriage
- Personal data
- Search by name, place, regnal year
- Under development: dynamic family trees

# Rusian genealogy

**<oo> -> <rg> Rusian genealogy**

Author: Svetlana Iaroslavich

**Sviatolav Iaroslavich**

This is a preliminary and preliminary about the marriage of Sviatolav Iaroslavich. This entry refers to the Russian medieval history. The following information would represent contemporary history of the end of the 10th century. In the 10th century, the Russian state was in the process of consolidation. The Russian state was in the process of consolidation. The Russian state was in the process of consolidation.

There are two main ways to represent it, **rg**, or in the very best of different ways, the list of which is below. The main difficulty for the representation is the lack of a single, unified system of representation. The Russian state was in the process of consolidation. The Russian state was in the process of consolidation. The Russian state was in the process of consolidation.

# Annotated Afanas'ev library

**<oo> -> <aal> Баба-яга и Василиса Прекрасная (104)**

**Linguistic commentary**

**Культурологический комментарий: Баба-яга**

В сказке «Баба-яга и Василиса Прекрасная» описаны события, связанные с жизнью Василисы Прекрасной. В сказке описаны события, связанные с жизнью Василисы Прекрасной. В сказке описаны события, связанные с жизнью Василисы Прекрасной.

В сказке описаны события, связанные с жизнью Василисы Прекрасной. В сказке описаны события, связанные с жизнью Василисы Прекрасной. В сказке описаны события, связанные с жизнью Василисы Прекрасной.

# Meter, rhyme, and rhythm

**<oo> -> <met> Meter, rhythm, and rhyme**

Author: Svetlana Iaroslavich

**Нелюбовь (Зинаида Николаевна Гиппиус)**

Line	Text	Meter	Rhyme	Stressed Syllables
1	Как ветер вихорь, ты былье в стволы,	0x10x(01)0x10x(01)	A	0 0 0 0
2	Как ветер черной, полых ты мой!	0x10x(01)0x10x(01)	B	0 0 0 0
3	Я вранью хвалю, и друг твой дамкой,	0x10x(01)0x10x(01)	A	0 0 0 0
4	Твой друг самый... отрой, отрой!	0x10x(01)0x10x(01)	B	0 0 0 0
5	Дарую и стужу, стужу ты моею,	0x10x(01)0x10x(01)	A	0 0 0 0
6	Дарюсь, но стужу стужу твою,	0x10x(01)0x10x(01)	B	0 0 0 0
7	Хвалю, хвалю, хвалю, хвалю,	0x10x(01)0x10x(01)	A	0 0 0 0
8	Мой друг востанет... хвалю, хвалю!	0x10x(01)0x10x(01)	B	0 0 0 0
9	Счастна хвалю, хвалю твой!	0x10x(01)0x10x(01)	A	0 0 0 0
10	Утворю и враню... враню, враню!	0x10x(01)0x10x(01)	B	0 0 0 0
11	Ты враню стужу, ты стужу мою,	0x10x(01)0x10x(01)	A	0 0 0 0
12	В стужу стужу... стужу, стужу!	0x10x(01)0x10x(01)	B	0 0 0 0
13	Охвалю, хвалю стужу,	0x10x(01)0x10x(01)	A	0 0 0 0
14	Люблю стужу стужу твою...	0x10x(01)0x10x(01)	B	0 0 0 0
15	Счастна хвалю, хвалю твою,	0x10x(01)0x10x(01)	A	0 0 0 0
16	Счастна хвалю... хвалю, хвалю!	0x10x(01)0x10x(01)	B	0 0 0 0

# Stress dictionary

**Stress report**

Line	Text	Stress	Stress	Stress
Line 1	Душею вихорь, ты былье в стволы,	0	1	0
Line 2	Как ветер черной, полых ты мой!	0	1	0
Line 3	Я вранью хвалю, и друг твой дамкой,	0	1	0
Line 4	Твой друг самый... отрой, отрой!	0	1	0
Line 5	Дарую и стужу, стужу ты моею,	0	1	0
Line 6	Дарюсь, но стужу стужу твою,	0	1	0
Line 7	Хвалю, хвалю, хвалю, хвалю,	0	1	0
Line 8	Мой друг востанет... хвалю, хвалю!	0	1	0
Line 9	Счастна хвалю, хвалю твой!	0	1	0
Line 10	Утворю и враню... враню, враню!	0	1	0
Line 11	Ты враню стужу, ты стужу мою,	0	1	0
Line 12	В стужу стужу... стужу, стужу!	0	1	0
Line 13	Охвалю, хвалю стужу,	0	1	0
Line 14	Люблю стужу стужу твою...	0	1	0
Line 15	Счастна хвалю, хвалю твою,	0	1	0
Line 16	Счастна хвалю... хвалю, хвалю!	0	1	0

## Computational methods course

**<oo>→<dh> Digital humanities**

Version 2.0 (2014-2015)

Lead: [djb@cs.pitt.edu](mailto:djb@cs.pitt.edu)

Course materials

**Administrative**

- Course website: [www.dh.pitt.edu](http://www.dh.pitt.edu)
- Course page and syllabus: [www.dh.pitt.edu/courses](http://www.dh.pitt.edu/courses)
- For more information on the course, please contact the lead.
- For more information on the course, please contact the lead.

**Course projects**

- *Project 1: Text analysis*
- *Project 2: Text analysis*
- *Project 3: Text analysis*

**General**

- The course is designed to be a hands-on learning experience.
- The course is designed to be a hands-on learning experience.
- The course is designed to be a hands-on learning experience.

**Required software**

- *Python*
- *Python*
- *Python*

## Summary: projects and courses

- Diplomatic editions of manuscripts: *Codex Suprasliensis*
- Critical editions of texts: *Rus' primary chronicle*
- Metadata and visualization: *The Repertorium of Old Bulgarian literature and letters*
- Textual transmission: *Daniel the Prisoner, Digenis Akritis*
- Distant reading: *Karolina Pavlova*
- Network analysis: *Menologies on line*
- Database: *Rusian genealogy*
- Language and prosody: *Meter, rhythm, rhyme*
- Language pedagogy: *Annotated Afanas'ev library*
- Learning DH: *Computational methods in the humanities*

## David J. Birnbaum

Thank you!

University of Pittsburgh  
 djb@pitt.edu  
<http://www.obdurodon.org>

