

LIBRARIES IN THE AGE OF DIGITAL INFRASTRUCTURES (AND OTHER CATASTROPHES)

Toma Tasovac

IFLA Satellite Meeting 2017: Digital Humanities (15.08.2017)



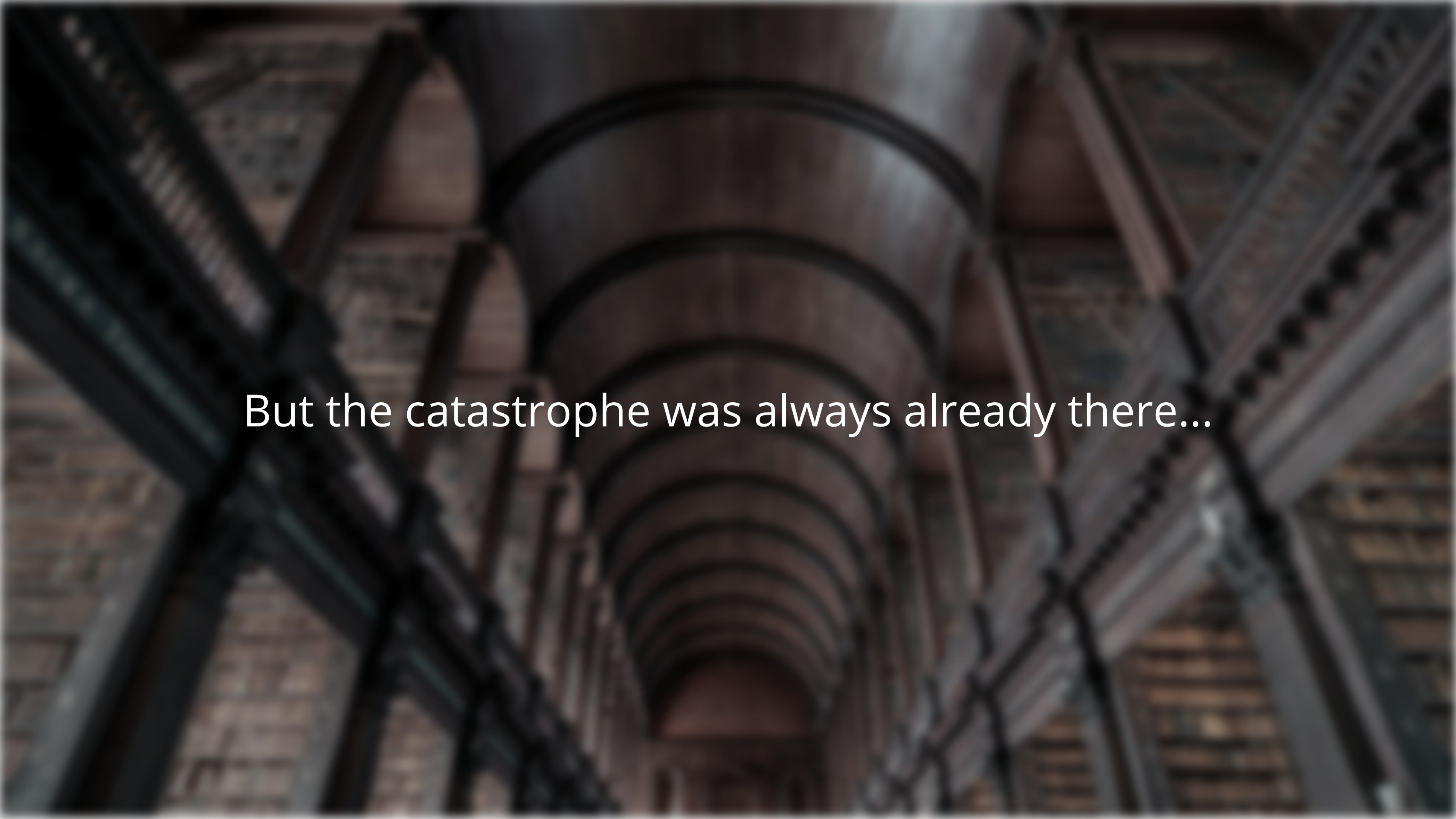
Καταστροφή!* (Disaster!)

καταστροφή = reversal, overturning, sudden disaster, death



MORBID FASCINATION

- death of the author
- death of the book
- death of the humanities



But the catastrophe was always already there...

This is what the printing presses do: they corrupt susceptible hearts. Yet the [...] silly asses do not see this, and brutes rejoice in the fraudulent title of teachers, exalting themselves with a song like this [...]: "O good citizen, rejoice: your city is well stuffed with books. For a small sum men turn themselves into doctors in three years. Let thanks be rendered to the printers!" Any uncultured person without Latin bawls these things.

— Filippo de Strata, *Polemic Against Printing* (c. 1473)





Technology is *not* and has never been the *other*
of the humanities

DIGITAL HUMANITIES

The ultimate revenge of global capitalism on traditional humanities?



TEXT AS DATA?

- there is nothing outside of text (Derrida, 1997: 158)
- text vs. context
- paratext: “a fringe of the printed text which in reality controls one’s whole reading of the text.” (Genette, 1997: 2)



DIGITAL HUMANITIES

A computational framework for
decontextualizing the humanities?



COMPUTERS
&

CAPITALISM
of america

**THEY'RE
GR-R-REAT!**

MATERIAL ARTIFACT VS. DIGITAL SURROGATE

What can you *do* with it?



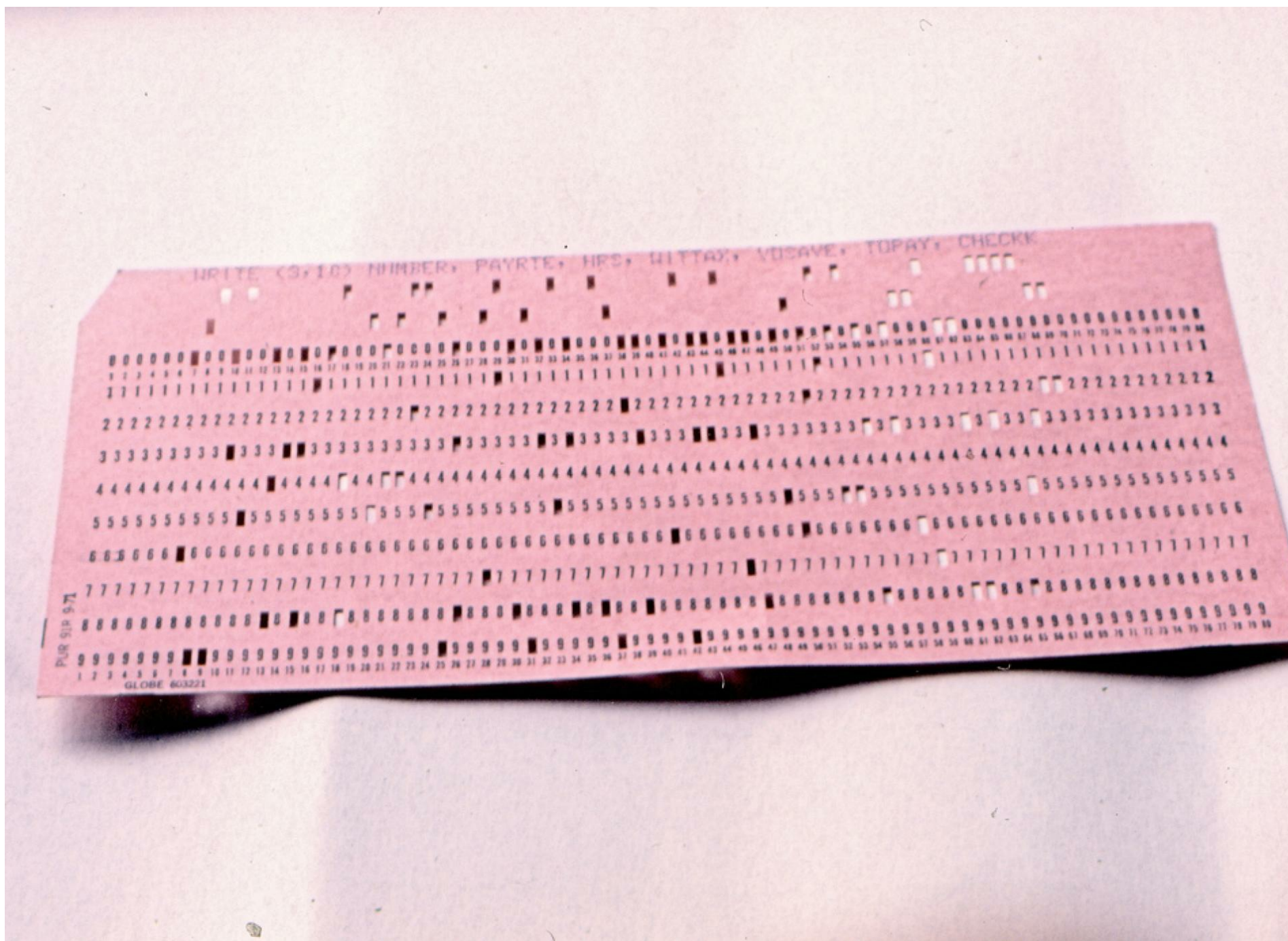
DH: RECONTEXTUALIZING THE HUMANITIES

- provide access (texts, tools, methods)
- discover patterns & generate new knowledge
- *supplement* close with distant reading



Roberto Busa





DH wasn't born yesterday: institutionalizing a community of practice

- Computers and the Humanities (1966)
- The Association for Literary and Linguistic Computing (London-based, 1973)
- Association for Computers and the Humanities (US-based, 1987)
- Literary and Linguistic Computing (1986)
- Alliance of Digital Humanities Organizations (ADHO, 2005)
- Master's and PhD programs



TEI: de facto standard for textual encoding

es textes...en référence à
normé.

exploitation des flux
produire des formes

structuré [formes]

n ajoutée (catalogage,
n, accès...) [interfaces]

stockage [pérennité]

ant des outils logiciels de
ou de mise en forme

é]

e dans de multiples formes,
multiples supports

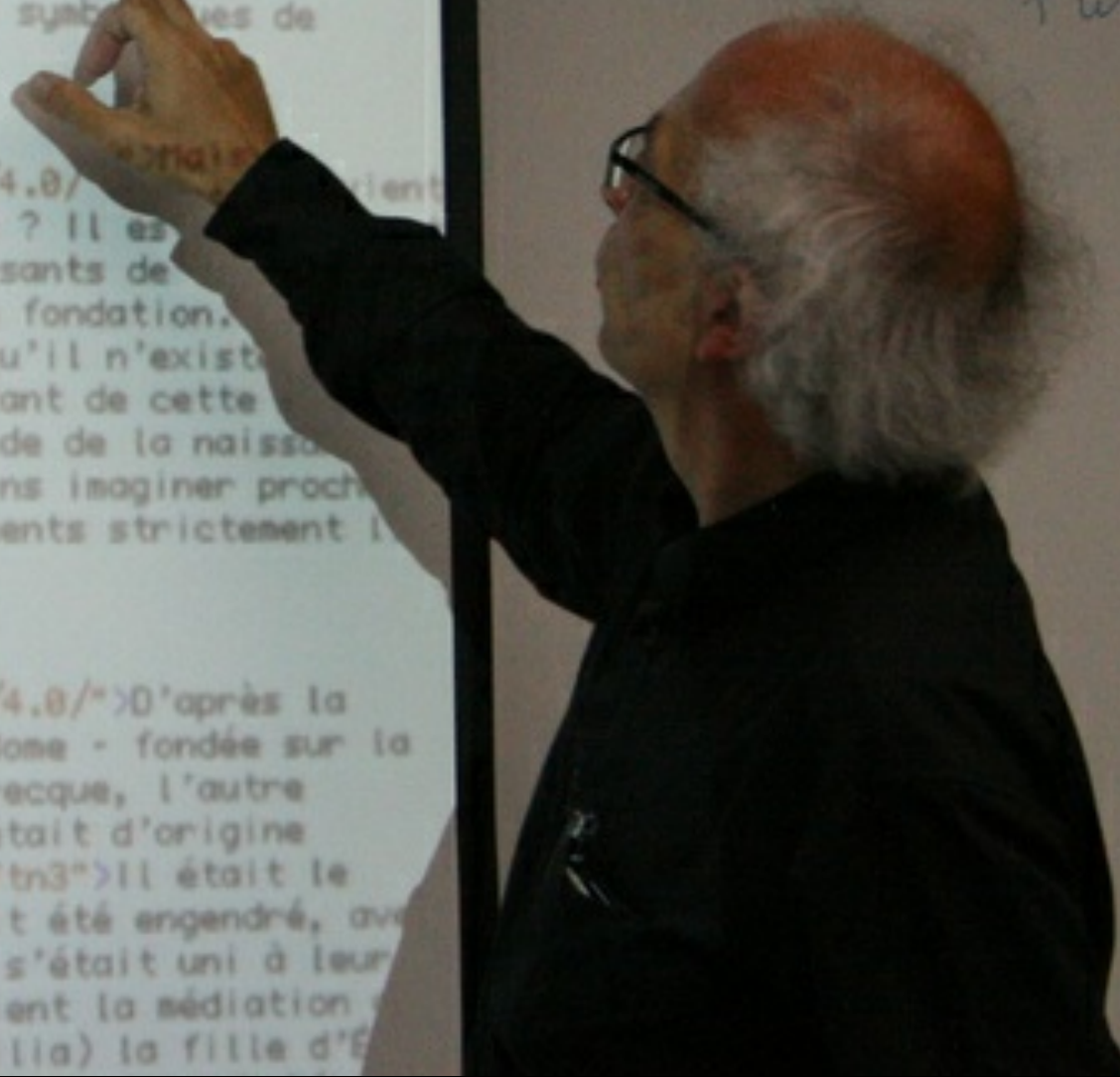
lité]

présentes dans le choix de la topographie architecturale, et elles peuvent être
celles des abbayes, comme pour le cloître de Saint-Pierre de Caudebec, ou
perforées, comme pour le cloître de Saint-Pierre de Caudebec, ou
xalis:ftn3">Voir V. Capobianco, « Le immagini simboliche e gli
stemmi di Roma », *Archivio della Società Romana di Storia
Patria*, n° 19, 1896, p. 349-356 ; M. Miglio, *Scrittura
Scrittori e Storia*, Manziana, Vecchiarelli, 1991, vol. 1,
p. 18-21 ; *Id.*, « Il leone e la lupa. Dal simbolo al
pasticcio alla francese », *Studi romani*, t. XXX, 1982,
p. 177-186 ; *Id.*, « Roma dopo Avignone. La rinascita
politica dell'antico », in *Memoria dell'antico nell'arte
italiana*, S. Settis (éd.), vol. 1, *L'uso dei
classici*, Turin, Einaudi (Biblioteca di Storia dell'Arte ;
1), 1984, p. 73-111 et G. Cavallo, « Il segno delle mura.
L'iconografia della città nel libro antico », in *Storia di
Roma*, IV, *Caratteri e morfologie*, Turin,
Einaudi, 1989, p. 267-300. Cette dernière image, récurrente
dans l'histoire culturelle de la ville, est celle d'une fesse à
laquelle sont associés des attributs réels et symboliques de
l'entité qu'elle représente.

`<p aid:pstyle="txt_Normal" xmlns:aid="http://ns.adobe.com/AdobeInDesign/4.0/">`
la tradition de personnifier la ville de Rome ? Il est
reconnu que l'un des aspects les plus intéressants de
mythographie d'une ville est la légende de sa fondation,
souvent d'une nécessité incontournable, puisqu'il n'existe
généralement aucun document historique attestant de cette
Comme pour d'autres villes antiques, la légende de la naissance
Rome est un mélange d'éléments que nous pouvons imaginer proches
la réalité, auxquels s'ajoutent d'autres éléments strictement liés
la sphère religieuse.

`<p aid:pstyle="txt_Normal" xmlns:aid="http://ns.adobe.com/AdobeInDesign/4.0/">`D'après la
version la plus répandue de la naissance de Rome - fondée sur la
fusion de deux légendes différentes, l'une grecque, l'autre
autochtone - Romulus, le fondateur éponyme, était d'origine
troyenne `<note aid:pstyle="txt_Note" xal:id="ftn3">`Il était le
descendant d'Énée par les rois d'Aibe et avait été engendré, avec
son frère jumeau Rémus, par le dieu Mars qui s'était uni à leur
mère, Rhéa Silvia. Certains auteurs supprimaient la médiation
de Mars et faisaient de Rhéa Silvia (ou Ilia) la fille d'Énée.

SSID :
Pud :



WHAT DO PEOPLE DO WHEN THEY DO DH?

- textual editing (digital editions, digital collections)
- data modeling and creation of structured data (XML)
- distant reading (as opposed to close reading)
 - statistical methods and processing
 - data mining
 - stylometry
 - visualization
 - network analysis
 - geospatial methods
- tools and services, incl. research infrastructures (CLARIN, DARIAH)



the traditional humanities focuses on a "minimal fraction of the literary field...[a] canon of two hundred novels, for instance, sounds very large for nineteenth-century Britain (and is much larger than the current one), but is still less than one per cent of the novels that were actually published: twenty thousand, thirty, more, no one really knows—and close reading won't help here, a novel a day every day of the year would take a century or so... And it's not even a matter of time, but of method: a field this large cannot be understood by stitching together separate bits of knowledge about individual cases, because it isn't a sum of individual cases: it's a collective system, that should be grasped as such, as a whole"



DH CHALLENGES

- interoperability
- communication
- integration
 - aggregation
 - federation



DH / LIBRARY CHALLENGES

- interoperability
- communication
- integration
 - aggregation
 - federation



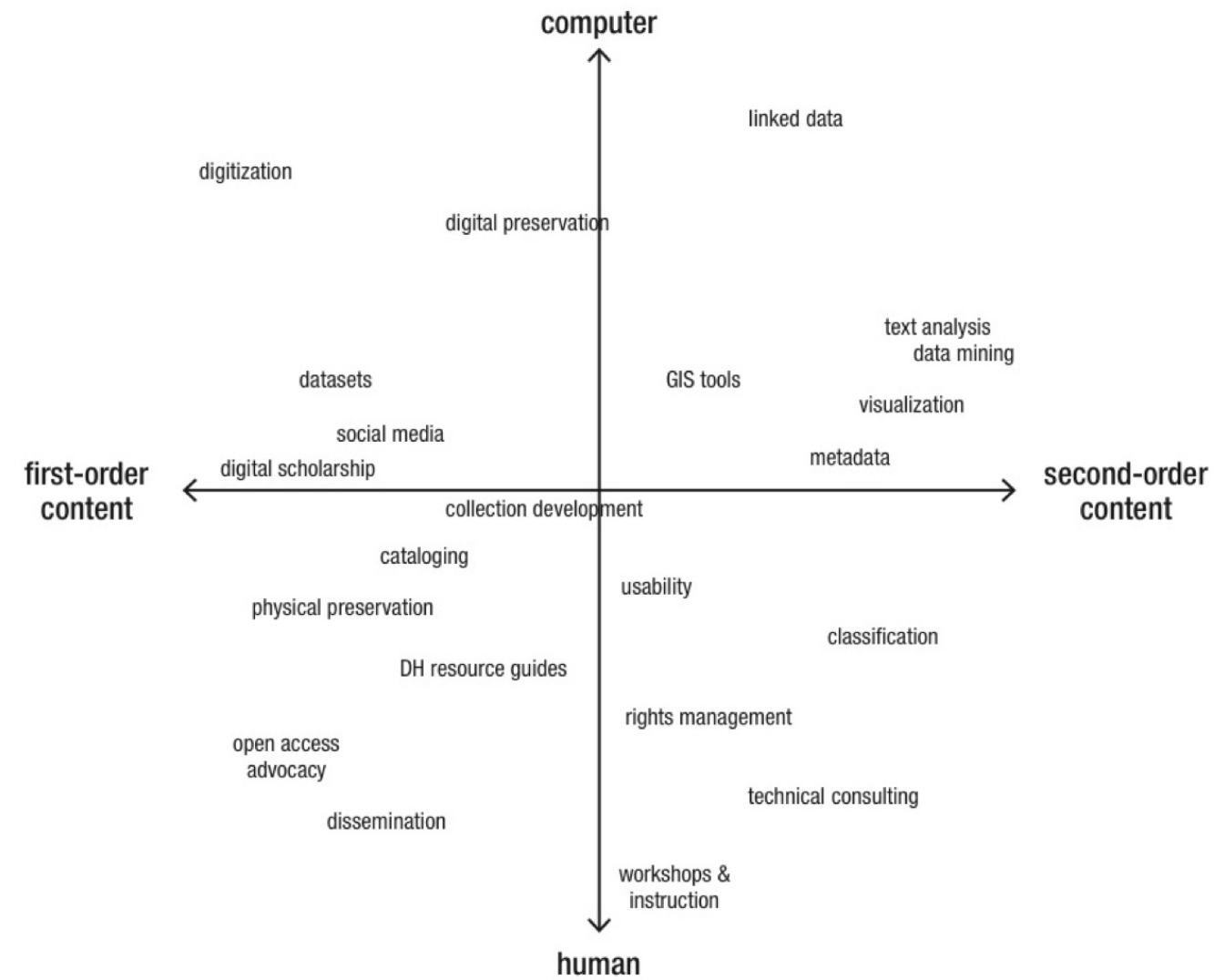
DH / LIBRARY / INFRASTRURAL CHALLENGES

- interoperability
- communication
- integration
 - aggregation
 - federation



Chris Alen Sula, "Digital Humanities and Libraries: A Conceptual Model", *Journal of Library Administration* 53(1) (2013): 10–26

Fig 3. A cultural informatics model for digital humanities and libraries



SCHOLARLY PRIMITIVES

- discovering
- annotating
- comparing
- referring
- sampling
- illustrating
- representing

Unsworth (2000)



TEXTS IN DIGITAL LIBRARIES?

- formats and delivery methods
- text/paratext boundaries
- computability, metrics and traces of use



TEXTS IN DIGITAL LIBRARIES

- must be machine-computable, not only machine-deliverable
- should be a kind of service as opposed to fixed object
- an interactive method as opposed to a static element



TEXT AS SERVICE

- metadata
- full text search
- retrieving identifiable fragments of texts
- persistent identifiers



TEXTS: OBJECTS → MODELS → SERVICES

- research infrastructures (cf. DARIAH / CLARIN)
- Open Annotation Collaboration (OAC) and the Web Annotation Working Group
- Canonical Text Services (CTS)



DH API LANDSCAPE

- HathiTrust
- DPLA
- Europeana
- IIF
- Open Siddur API
- Folger Digital Texts API
- correspSearch API
- Old Bailey API
- European Holocaust Research Infrastructure API
- Manuscripts Online API

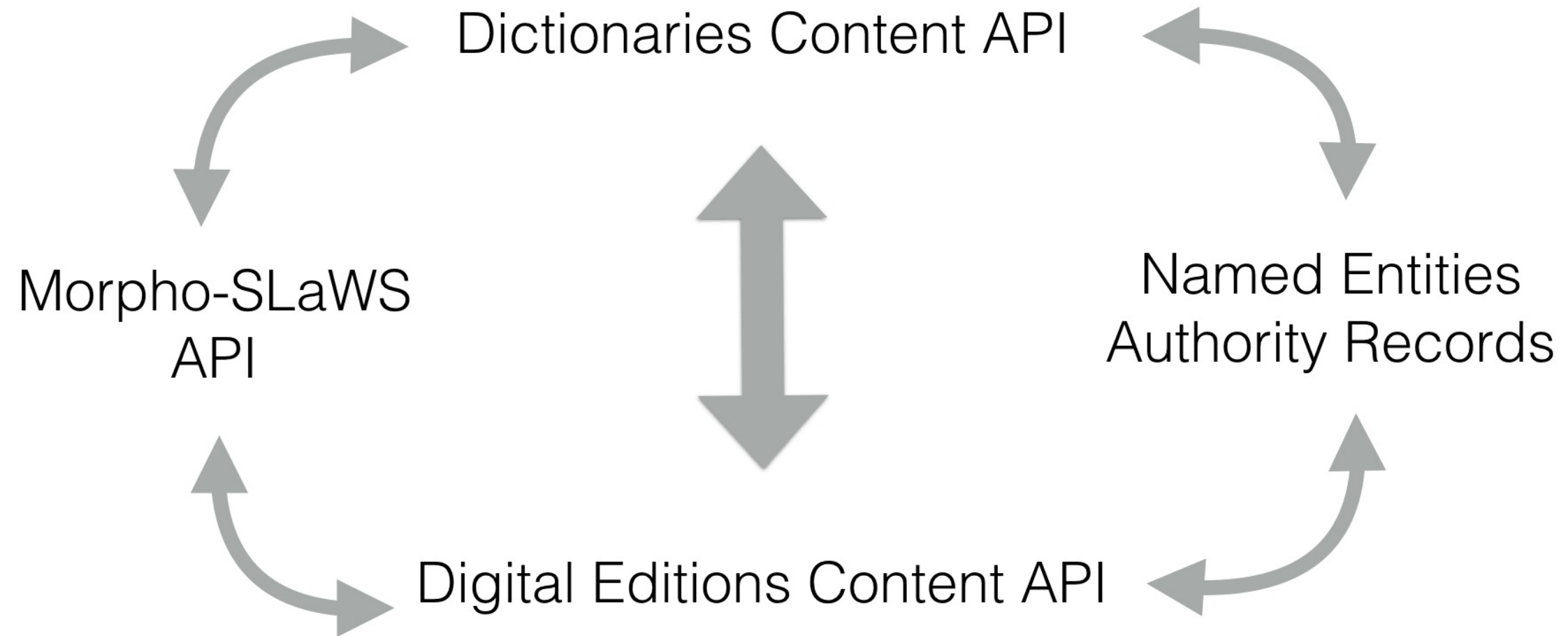


AT BCDH:

Raskovnik: Dictionary Platform
(<http://raskovnik.org>)

Izdanak: Digital Editions (in development)





Raskovnik API
Toma Tasovac · raskovnikapi

Documentation Inspector Editor Tests

Create a team

Fork on GitHub
BCDH/raskovnik-api

INTRODUCTION

Allowed HTTPs requests:
Description Of Usual Server Responses:
Pagination and Headers:

REFERENCE

Entries Collection
Slices Collection
Context Collection
Range Collection
Random Collection
Config Collection

Raskovnik API

INTRODUCTION

Raskovnik is an API that allows consumers to do the context-based search on TEI dictionaries, which can be either directly through searching their own paths or searching for specific lemmas. It also allows various other functionalities to view results in different forms as per the user.

Allowed HTTPs requests:

GET : Get a resource or list of resources

Description Of Usual Server Responses:

- 200 OK - the request was successful.
- 204 No Content - the request was successful but there is no representation to return (i.e. the response is empty).
- 400 Bad Request - the request could not be understood or was missing required parameters.
- 401 Unauthorized - authentication failed or user doesn't have permissions for requested operation.
- 403 Forbidden - access denied.
- 404 Not Found - resource was not found.

Parameters

Parameter	Description	Type
lemma	a lemma need to be searched Example: a6a	String
dict	retrieves resources for all or specific dictionaries based on their ids. Multiple ids should be comma-separated, for instance dict=VSK.SR,MZ.RGJS. Default: all	String
page	page of results expected by client. Default: 1	Number
limit	maximum number of records expected by client. Example: 1. Default: 10	Number
transform	name of the XSLT file (in the \$TRANSFORMATION directory) that should be applied to retrieved resources before returning them. Default: false	String
format	results expected by client should be in XML or JSON. Default: xml	String

```
GET http://raskovnik.org/api/v1/entries?lemma=%D0%B0%D0%B1%D0%B0&dict=all&page=1&limit=1&transform=false&format=xml&group=false&wrap=true&contains=false
```



поп| 🔍

пѡп	Караџић
поп	Златановић
пѡп	Елезовић
пѡп златѡр, пѡп зѡк	Караџић
пѡпа	Караџић
попа	Елезовић
пѡпадати	Караџић
попадѡка	Златановић
попадѡка	Златановић
попадѡдија	Караџић
попадѡја	Елезовић
попадѡдијин	Караџић
Попадѡћ	Елезовић
попѡз	Елезовић



Врањанске пословице

Одреднице које садрже пословице у Речнику говора јужне Србије.

Прикажи: списак речи целокупне одреднице

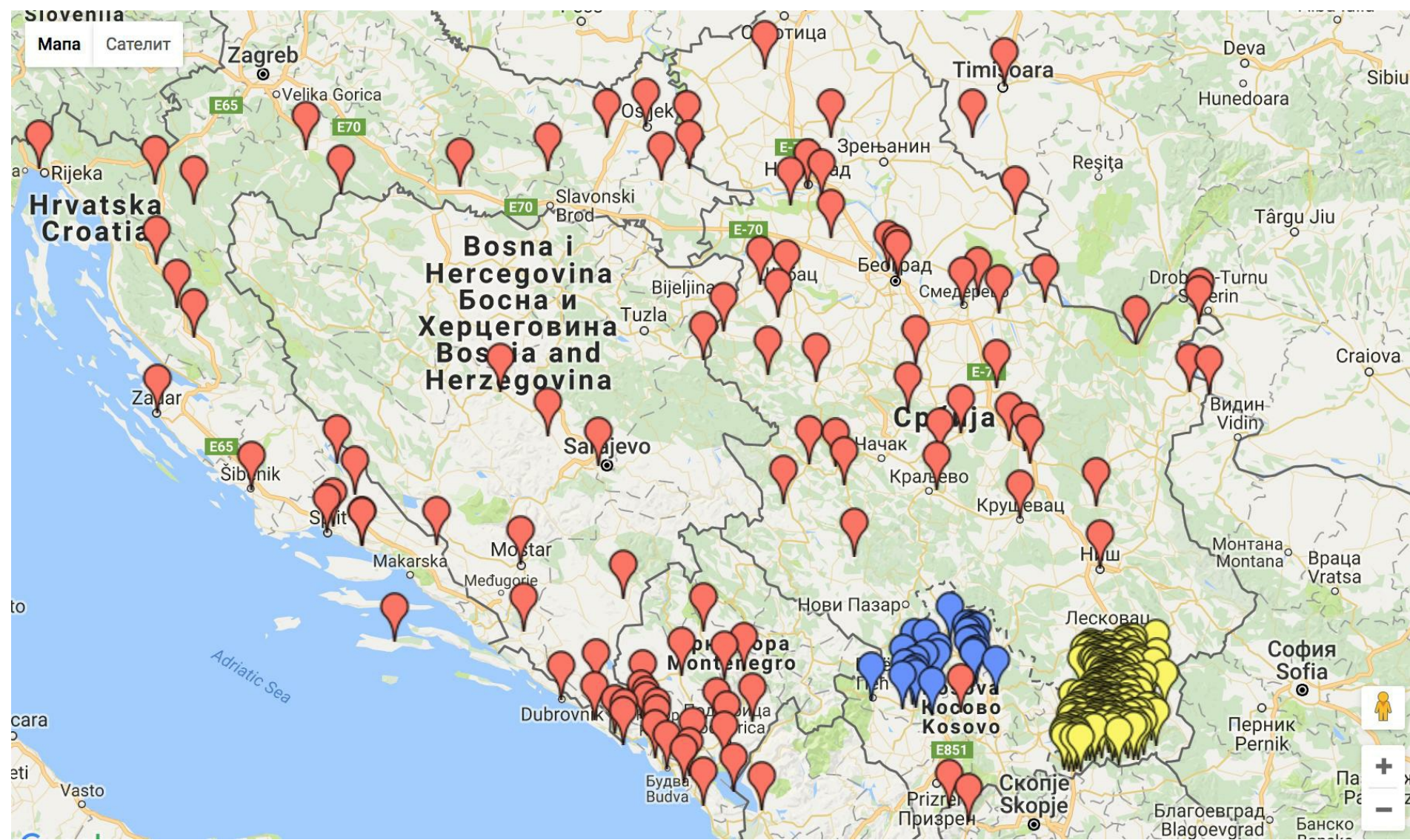
УКУПНО: 171

1	2	3	4	5
---	---	---	---	---

азър	беговица	врзан	еспап
ајлазүјем	Бежан	гревор	ешек
алим	благүјем	грчав	забављам (се)
апање	брéвам		
аспра	булүтин		
атъл	булча		
багламе	везирлѝк		
башка	вражѝлац		
башчаванѝја	вржем		

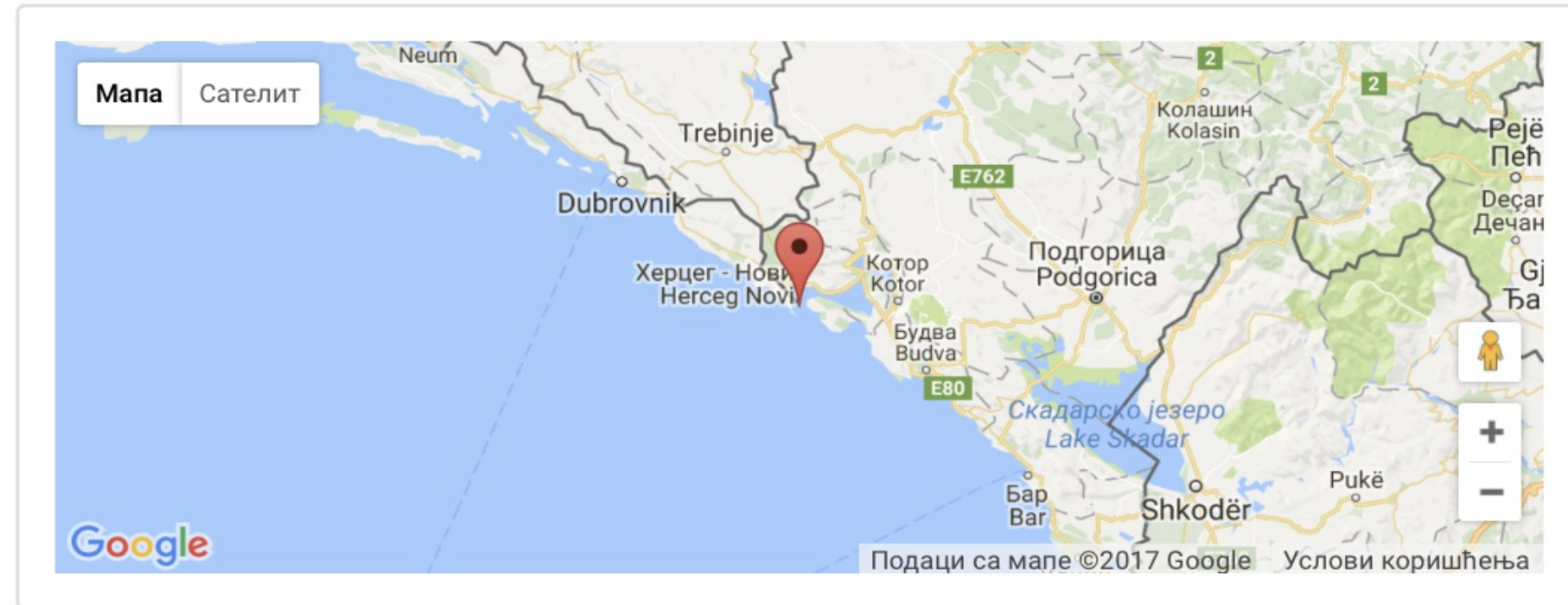
дрт, -а, -о стар и изнемогао. „Виђује се у авлију једн дрт старчина.“
„Не чувај јаре од дрту козу“ (посл.) (Врање) „Сас дрто магаре на пазар“
(Владичин Хан).





Бока Которска

географска област



Прикажи: списак речи целокупне одреднице

УКУПНО: **673**

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

âк

Àмат

àцâл

бáбовић

бàбушка

бàтyн

бàхорити

бàхорица

бèдрениˆпраг

бèрба

блèска

бљèснути

бобòтати

богàтáш

богòраднйк

брèме

бркљача

брњица

брхан

брцнути

Српски рјечник

баја|



бӑсма, *f.* оне ријечи скупа што бајалица или бајач г
Zauberspruch, carmen.

бӑсма* *f.* 1. gedruckte Leinwand, linteum pictum. 2.
рухидис ад пулверем нитратум [cf. [везма](#)].

← баскйјати

⇒ 10 ⌘ 1

Речник косовско-метохиског дијалекта

бӑсма, *e ж.* циц. У В. 2. У В. 2 **бӑсма ж.** 1) без оног др
његга наведено. У РЈА. 2. *bàsma ж.* 1).

← баскйн

⇒ 10 ⌘ 1

Бӑја	+
бӑја	+
бӑјалица	+
бӑјало	+
бӑјам	+
бӑјат, бӑјатан	+
бӑјати	+
Бӑјац	+
бӑјӑч	+
бӑјӑње	+
Бӑјко	+
Бӑјнӑ Лука	+



**TEXT → MODEL → API → SERVICE →
INFRASTRUCTURE**



ANCHORING DIGITAL SCHOLARLY METHODS IN DIGITAL LIBRARY WORKFLOWS

Developing tools and services for different types
of scholarly primitives in modular, dynamic,
research-driven digital collections



IN AN IDEAL WORLD, DH-FRIENDLY LIBRARIES WOULD

- provide access to machine-readable and machine computable texts
- develop well-documented APIs for both metadata and full-text access
- offer layers of automatic and semi-automatic structural, linguistic and semantic annotations with support for crowdsourced corrections
- integrate dictionary and other type of reference services to be queried directly from the text



DATA ENRICHMENT WORKFLOWS

Digital library → Scholars → Digital library



Von der Gottheit,
 und den
 Dingen, welche den Gottes-
 dienst angehen.

☩ Бжествѣ, и вѣщѣхъ, гаже Бжѣи
 слѣжбы касаютса.

Gott, Góttz.	Бѣз, Бѣгъ.
Die Gottheit.	Бжество, Божество.
ди Гóттхайтз.	Бѣз Оцъ,
Gott der Vater,	Бѣз Оцѣцъ.
Góttz derz Fátterz.	Бѣз Онъ,
Gott der Sohn.	Бѣз Оынъ.
Góttz derz Gónz.	Бѣз Дхъ стый.
Gott der heilige Geist,	Бѣз Дѣхъ сватый.
Góttz derz хайлиге	Отѣл Трѣца, сватѣл
Гайстз.	Трѣица.
Die heilige Dreifaltigkeit,	
ди хайлиге Дрѣифал-	
тигкайтз.	

... * « ...
 Г.
 un ь b«n
 Singen, ше
 1фе ben 1 bienft angelen*
 Л.
 4 *
 GO Кжсствнѣ, и кеци^х, гаже КжГ/л
 слѣжбы касаютса.
 | •
 &ott, rÓTTTU Brü, ®
 ie ©oft^eif. Бжктко, Божктво. *;
 _ / ^ „ -
 ди Гóтт^литор. . Нгч ©
 oft beifSoafec, Bors \
 f(iÄ Флтгря. КГА GN'Ä ,
 бес €>офп. БОГА вы НЗ.
 Тоттч А'f* GO'Мї. Вгх Д\х стый« ©
 oti бес) fällige ©eifī,
 ГОТТА Д|р А ХАЙЛИП
 ГАЙСТА.;, -, GT(ÍATfl)A,
 ie ſetitge SDrdfälligf ŷti, ТРОИЦА.
 I ъДИЛИГ4 * --y--'
 ТИГКАЙТХ.

DATA ENRICHMENT WORKFLOWS

Digital library → Scholars → Digital library



Καταστροφή!* (Disaster!)

καταστροφή = reversal, overturning, sudden disaster, death

Καταστροφή!

καταστροφή = 1. reversal, overturning, sudden disaster, death. 2. turning point (in tragedy), unravelling, denouement.

DIGITAL LIBRARY AS INFRASTRUCTURE



LIBRARIES AND DH

- recontextualizing the humanities
- providing small and big data (+ tools and methods)
- enabling close and distant reading
- developing data enrichment workflows



Thank you for bearing with me!

