

CapiTainS

Challenges for the generalization and adoption of open source software

Bridget Almas



Head Engineer
Alpheios
balmas@gmail.com
Twitter : @BridgetAlmas
Github : @balmas
alpheios.net

Thibault Clérice

Head of TNAH MA
École Nationale des Chartes
thibault.clerice@chartes.psl.eu
Twitter : @ponteineptique
Github : @ponteineptique

Matt Munson



Project Head Engineer
Universität Hamburg - Formulae
matthew.munson@uni-hamburg.de
Twitter : @sonofmun
Github : @sonofmun
www.formulae.uni-hamburg.de

Plan

1. Introduction
2. What's CapiTainS ?
3. Expanding the user base from a software point of view
4. What's documentation ?
5. Organizing Workshops : for whom and for what ?
6. Conclusion

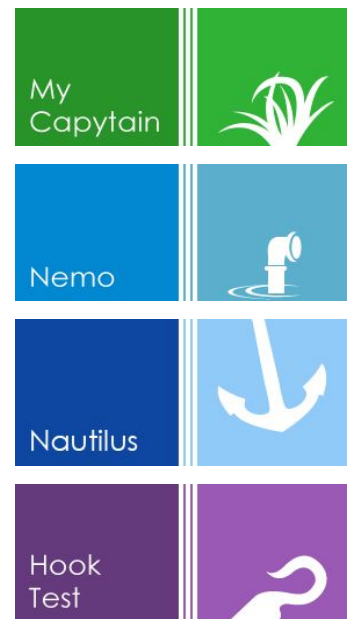
What's Capitains ?

- = Software Suite + Guidelines
- Not specific to any institution
 - But built thanks to them !
- Partial and separate answers to
 - Editing
 - Serving
 - Curating
 - Reading
- Bootstrapping development of project
- Born in the context of
 - Perseus
 - Perseids
 - Humboldt Chair of Digital Humanities



What's Capitains

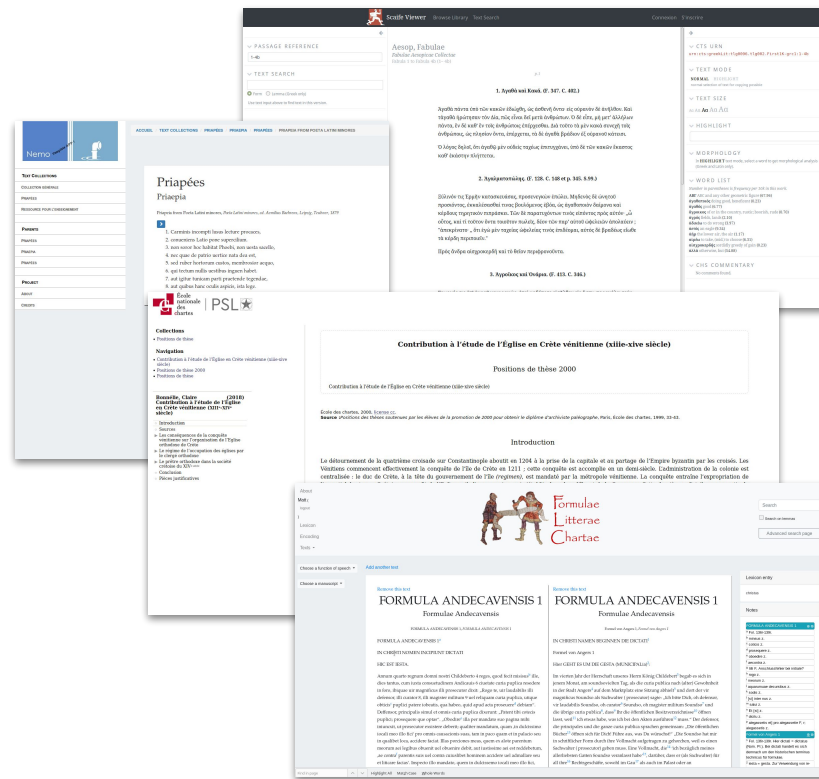
- Guidelines
 - Metadata + TEI or Epidoc
- Sparrow
 - Javascript
 - Old and needs to be rebuilt to have better quality software
- MyCapytain
 - = “Standard” library for CapiTainS
 - = Kind of lego set for Capitains
- HookTest
 - Testing corpora (locally or decentralized)
- Nemo
 - Reading Interface Skeleton
- Nautilus
 - API Generator



Capitains Python Realms

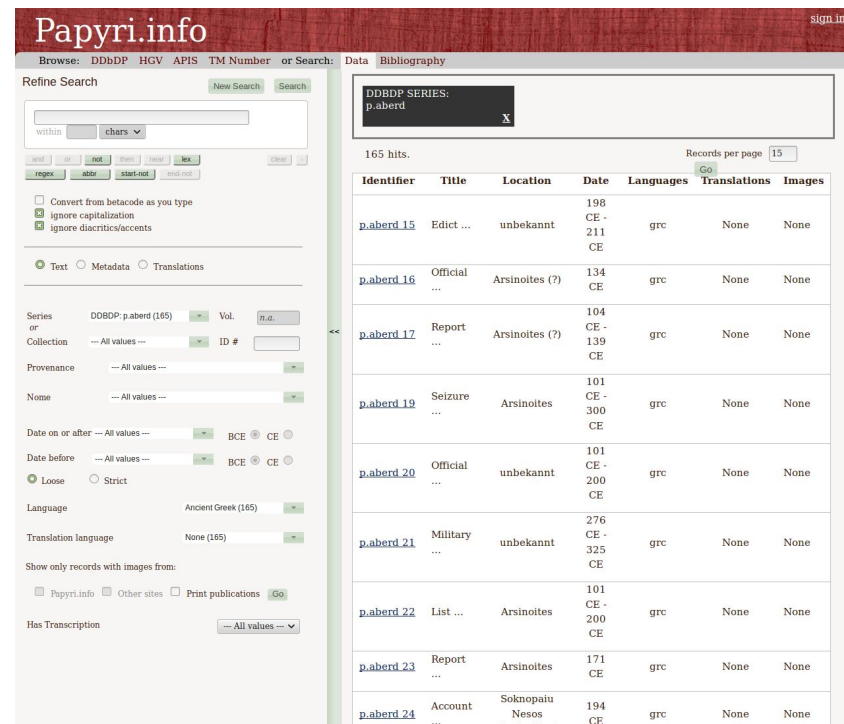
Expanding the user base from a software point of view

- What can we do to support a wider variety of textual data ?
- What are our limitations ?



CTS and its limitation as an API and a data model

- Identifiers and metadata scheme
- Amount of texts
- Reliability of responses across implementations



The screenshot shows the Papyri.info search interface. The search results are filtered to the series 'p.aberd', showing 165 hits. The results are displayed in a table with columns for Identifier, Title, Location, Date, Languages, Translations, and Images.


Identifier	Title	Location	Date	Languages	Translations	Images
p.aberd_15	Edict ...	unbekannt	198 CE - 211 CE	grc	None	None
p.aberd_16	Official ...	Arsinoites (?)	134 CE	grc	None	None
p.aberd_17	Report ...	Arsinoites (?)	104 CE - 139 CE	grc	None	None
p.aberd_19	Seizure ...	Arsinoites	101 CE - 300 CE	grc	None	None
p.aberd_20	Official ...	unbekannt	101 CE - 200 CE	grc	None	None
p.aberd_21	Military ...	unbekannt	276 CE - 325 CE	grc	None	None
p.aberd_22	List ...	Arsinoites	101 CE - 200 CE	grc	None	None
p.aberd_23	Report ...	Arsinoites	171 CE	grc	None	None
p.aberd_24	Account ...	Soknopaiu Nesos	194 CE	grc	None	None

How we actually repeated our errors

- MyCapytain 1.0.0 > 2.0.0
 - Data model = Distributed Text Services
 - Underlying hardcoded models
 - Remainings of CTS dependency
- MyCapytain 2.0.0 > 3.0.0
 - Data model ~ DTS
 - Open Public Discussions ([Issue 165](#))

Citation refactoring #165

[Open](#) Pontetneptique opened this issue 28 days ago · 5 comments

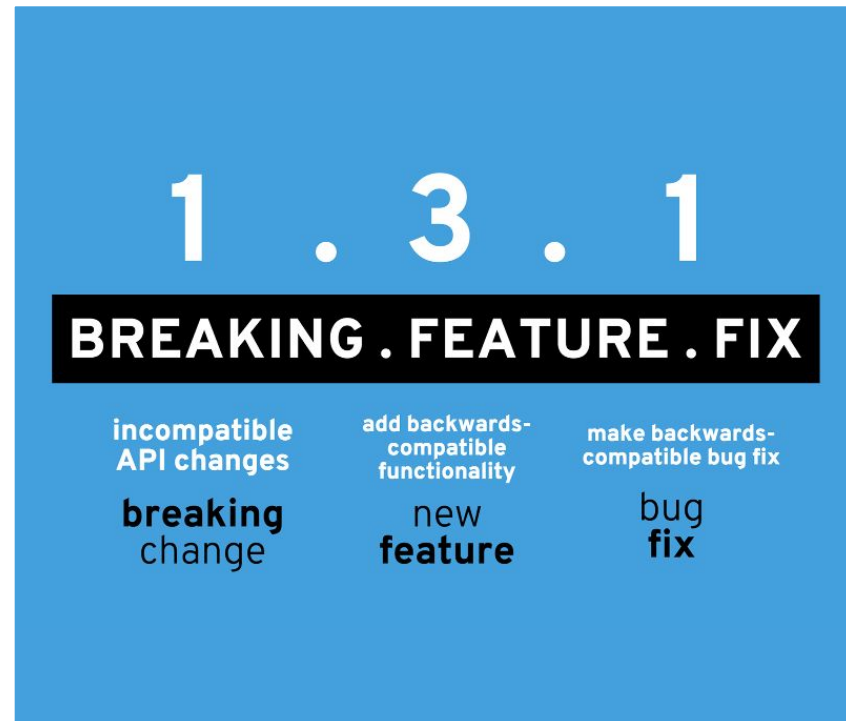
 Pontetneptique commented 28 days ago · edited · Member

Currently on <https://github.com/Captains-MyCapytain/tree/dts-retriever-MyCapytain/common> reference, I have started refactoring CTS Focused object. In this context, the original objects will retain their CTS specificity.

I have not changed the basic sense of most citations, keep in mind one of the objective is to prepare the arrival of the new guidelines : <https://github.com/Captains/guidelines#tests>

I would really like to gather feedback on the current BaseCitation and if you have any wishes regarding these objects.

- New Generic : Prototype CitationBaseObject
 - New ..extract(passages[]): function
- New Generic : Prototype Citation Object
 - New ..setLaten: Property
 - New ..length: property (Not ported to CTS Citation object yet)
 - New meaning of ..[...]: magic method
- New ..root and ..k_root: properties to link back to the citation set (for match implementations)
- Guidelines / CTS Citation Object
 - Implementation of the ..extract(passages[]): method



1 . 3 . 1

BREAKING . FEATURE . FIX

incompatible API changes	add backwards- compatible functionality	make backwards- compatible bug fix
breaking change	new feature	bug fix

<http://digitalcommunications.wp.st-andrews.ac.uk/2017/03/17/what-our-version-numbers-mean/>

What's documentation ? What we did wrong

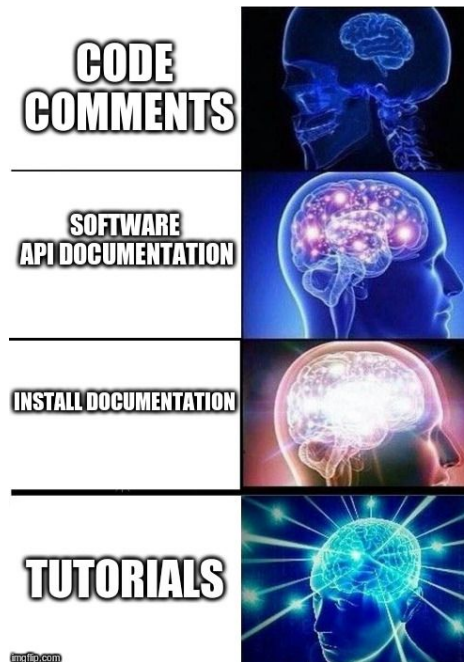
- Focused on developer documentation
- Some technical explanations about the why
- Still, difficulties to catch up :
 - a. Code base starting to be large
 - Naming convention can be as efficient as difficult
 - b. Focusing on tutorials and python notebooks would have been better ?



Legend : Which Capitains software should I use again ?

What's documentation ? What we should have done

- Priorities : making the most useful packages easier to pick up
- Tutorials
 - <https://github.com/Capitains/tutorial-nemo>
- Nemo Template App
 - <https://github.com/capitains/nemo-template-app>



imgur.com/a/GB2zAOo

Organizing workshops : Whom for ? What for ?

- Lot of workshops held in Europe + USA
- User base = Original Institutions + New Institutions of former members of the original institutions
- So what did we do wrong ?



Capitains trainer facing the results of years of workshops

Organizing workshops : Past Results

- Focused mainly on data production
- But what's the point if you do not have a website ?



Munich, Antikensammlungen J 1493, d'après LIMC, s. v. Sisyphos n° 6.
<http://books.openedition.org/pupo/1834>

Organizing workshops : Target software ?

- Hooktest ?
- MyCapytain ?
- Nemo / Nautilus ?



Organizing workshops : Nemo

- Make using Nemo simpler
- Reduce or take into account skills required before your own software
 - Reevaluate after each training !
- Think about the technical needs
 - We all hate windows but some users will have it. Try to prevent being surprised.
- Train to hack, not to know.



*Title : Our Wake Up Call
(Capitains does not promote violence against fellow
open source programmers)
<https://imgflip.com/i/2bt3or>*

Organizing Workshops : The future

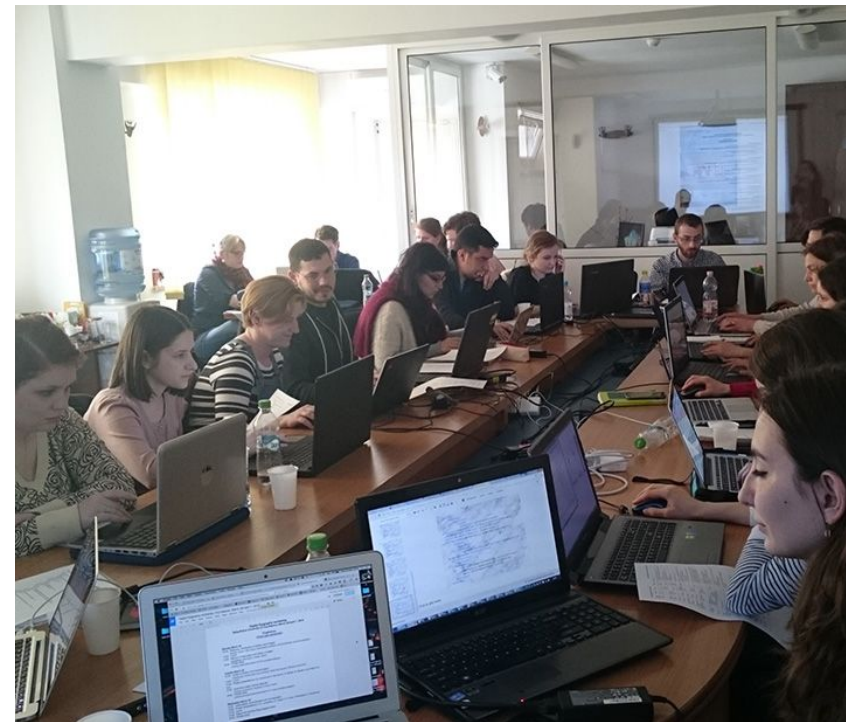
- People should leave with real knowledge
- Workshops
 - Intensive and in person
 - 6/8 trainees per trainer
- Try to have a simpler version of Nemo, the simplest possible



*It should be as simple and as complex as this.
It's not a perfect plane, but it still is one*

Conclusion

- Regular call for feedback
 - For code
 - For improving accessibility
 - <https://github.com/Capitains/MyCapytain/issues/149>
- Limit the concentration of power and knowledge
 - Build a real community of discoverers, users and developers knowing that there is a large spectrum in between these categories
 - Make things clearer !



Cluj-Napoca, Romania, March 28–April 1, 2016 (photo S. Stoyanova)



Thanks!

Capitains.org

GitHub.com/Capitains