Software Release and Progress Report

Start date: 2019-06-26 End date: 2019-10-14

Short Summary

A major new version of FLAT has been released this period, covering support for annotating arbitrary higher-order relations. This was an explicit part of the deliverables for T062/T063 and a prerequisite for later implementing cross-document support (think for instance parallel documents/corpora).

CLAM has had several bugfixes to make the Autosearch forward work in PICCL, as deployed by INT.

A new language identification tool has been written (`colibri-lang`) for Nederlab, with support for two variants of old dutch. This actually brings the total of FoLiA-aware language identification tools to three (`colibri-lang`, `FoLiA-textcat`, `folia-langid`), each builds on a different system than the others.

An experimental new FoLiA library for Rust was developed and released, the aim was to see if a high-performance library could be made with significantly less memory consumption and high speed, which may be useful in the future. The benchmarks prove succesful.

Talks have been started up regarding participating in federated authentication for our webservices (this was not granted as part of T018 though, as it was not deemed needed for CLST to be a CLARIN centre?). For proper inclusion in the CLARIN switchboard, it is relevant again though.

- Nelleke is finalizing her report on mapping metadata between SONAR and CGN
- On Aug. 30 those of us running instances of OpenSoNaR Plus decided to install a copy of INT's version, so that the same version will also run in Nijmegen and Tilburg
- Work on launching the CLARIN K-centre on Atypical Communication Expertise in November is ongoing
- Henk van den Heuvel gave a presentation at DH2019 about the Transcription Chain. In addition a workshop on the topic was organised in cooperation with the SSHOC project. A blog is in preparation.
- Henk van den Heuvel gave a presentation on 3 September about secure management of sensitive data at https://www.tapas-etn-eu.org/events/te3
- At the CLARIN AC in Leipzig (30 Sep. - 3 Oct.), Henk van den Heuvel presented the Transcription Chain and the DELAD initiative on corpora of disordered speech
- We have started preparations for the ASTA subproject on dialect recordings with Meertens
- Nelleke Oostdijk gave two lectures at the Xornada de estudios anotación e explotación de corpus orais, a network meeting of RED Temática en Estudios de Análisis del Discurso (FFI2017-90738-REDT) on 3 and 4 October 2019 in Santiago de Compostela. The titles of the lectures were resp. *The Spoken Dutch Corpus Revisited* and *Linguistically-motivated Annotation: Moving forward*.

CLAM

Project/task/deliverable references: CLARIAH-PLUS WP3 T142; CLARIAH-CORE WP3 T21 [D2.8 (software); D2.9 (doc)]

clam v2.4.13

Forward viewer fix (final?)

(Released by Maarten van Gompel on 2019-09-02) https://github.com/proycon/clam/releases/tag/v2.4.13

clam v2.4.12

Forward viewer fix (previous one was wrong)

(Released by Maarten van Gompel on 2019-09-02) https://github.com/proycon/clam/releases/tag/v2.4.12

clam v2.4.11

Fixed version compatibility check

(Released by Maarten van Gompel on 2019-09-02) https://github.com/proycon/clam/releases/tag/v2.4.11

clam v2.4.10

Another minor bugfix release for the forward viewer

(Released by Maarten van Gompel on 2019-09-01) https://github.com/proycon/clam/releases/tag/v2.4.10

clam v2.4.9

Remove unneeded indirection in ForwardViewer unless explicitly enabled (LanguageMachines/PICCL#51)

(Released by Maarten van Gompel on 2019-08-30) https://github.com/proycon/clam/releases/tag/v2.4.9

clam v2.4.8

Minor bugfix release to fix issue in forward viewer

(Released by Maarten van Gompel on 2019-08-27) https://github.com/proycon/clam/releases/tag/v2.4.8

clam v2.4.7

- attempted fix for problem in forward viewer (LanguageMachines/PICCL#51)
- deprecating Python 3.4 support (because lxml is doing so too)
- also count number of file when determining disk usage (not really used yet)

(Released by Maarten van Gompel on 2019-08-25) https://github.com/proycon/clam/releases/tag/v2.4.7

clam v2.4.6

Minor update release, adds the ability to disable a webservice from accepting new jobs by setting ENABLED = False in the service configuration. Useful for maintenance purposes.

(Released by Maarten van Gompel on 2019-07-05) https://github.com/proycon/clam/releases/tag/v2.4.6

FLAT

Project/task/deliverable references: CLARIAH-PLUS WP3 T062 & T063; CLARIAH-CORE WP3 T24 [D2.10 (software); D2.11 (docs)]

flat v0.9.0

This release introduces several big new features:

- Implemented support for relations (previously known as alignments) (#84)
  - This finishes the essential prerequisites for syntactic movement annotation (#138)
- Handle alternative annotations (#149)
- Added the ability to edit the structure's class (#152) Some other changes:
- Expanded the test suite according to the new features (#156) and updated documentation
- Major refactoring in the showeditor() function
- Output internal editor errors to javascript console
- Span annotations without set are now properly handled #150
- Documenting 'slices' configuration directive #153
- Fix toggles in menu
- Improved styling of activated menu elements on mouseover

(Released by Maarten van Gompel on 2019-08-30) https://github.com/proycon/flat/releases/tag/v0.9.0

foliadocserve v0.7.2

Updated tests, compatible with new FLAT v0.9.0, no further functional changes.

(Released by Maarten van Gompel on 2019-08-30) https://github.com/proycon/foliadocserve/releases/tag/v0.7.2

foliadocserve v0.7.1

- serialise alternative annotations for FLAT (proycon/flat#149) (for upcoming flat v0.8)
  - fix for alternative annotations
- handling of setless annotations (proycon/flat#150)
- fix regression problem upon submission of metadata (proycon/flat#155)

(Released by Maarten van Gompel on 2019-08-15) https://github.com/proycon/foliadocserve/releases/tag/v0.7.1

FoLiA

Project/task/deliverable references: CLARIAH-PLUS WP3 T108; CLARIAH-CORE WP3 T71 [D2.5 (libs); D2.6 (docs); D2.7 (tools)]

folia v2.2.1

Minor update release:

- added a syntactic movement example from proycon/flat#138
- added FQL documentation on adding relations (proycon/foliapy#11)
- allow AbstractInlineAnnotation in quote

(Released by Maarten van Gompel on 2019-08-20) https://github.com/proycon/folia/releases/tag/v2.2.1

folia v2.2.0

- The default text delimiter for `<part>` and `<ref>` is now space, but this can be overridden with the `space="no"` attribute. #61
- Revised "AS ALTERNATIVE" directive in FQL
- Various documentation, test and example updates

(Released by Maarten van Gompel on 2019-07-23) https://github.com/proycon/folia/releases/tag/v2.2.0

folia-rust v0.0.2

- Fixed text markup parsing (#3)

(Released by Maarten van Gompel on 2019-10-03) https://github.com/proycon/folia-rust/releases/tag/v0.0.2

folia-rust v0.0.1

Very first release of the FoLiA library for Rust (consider it an alpha version). Implements a lot of basic functionality (but no validation!). Note that this library may still be a bit rough around the edges and may contain major bugs.

(Released by Maarten van Gompel on 2019-10-02) https://github.com/proycon/folia-rust/releases/tag/v0.0.1

foliapy v2.2.2

Minor bugfix release:

- fixed bug with processor= keyword argument to "Document class

(Released by Maarten van Gompel on 2019-10-10) https://github.com/proycon/foliapy/releases/tag/v2.2.2

foliapy v2.2.1

- updated for FoLiA v2.2.1
- prevent adding a duplicate generator subprocessor
- relax parsing of the xref/@type attribute (it's status as either mandatory or optional was unclear, so let's keep it optional)
- [FQL] implemented support for relations #11
- alias was not properly serialised on declarations (#12)

(Released by Maarten van Gompel on 2019-08-20) https://github.com/proycon/foliapy/releases/tag/v2.2.1

foliapy v2.2.0

- Implemented AbstractWord.alternativelayers() method and made findspans() able to handle alternative spans
- Various FQL improvements, mainly regarding alternative handling and revised AS ALTERNATIVE directive
- Allow set-holding and set-less declarations to co-exist (proycon/folia#74)
- Expanded test suite

(Released by Maarten van Gompel on 2019-07-23) https://github.com/proycon/foliapy/releases/tag/v2.2.0

foliatools v2.2.4

- [alpino2folia] Fixed conversion of dependencies and syntax layer, was missing or a long time due to an unforeseen change in Alpino output (proycon/folia#49)
- [foliaspec] Added code generation for new Rust library
- [foliabench] Added a new benchmark tool for the foliapy library, implementing four benchmarks.

(Released by Maarten van Gompel on 2019-10-03) https://github.com/proycon/foliatools/releases/tag/v2.2.4

foliatools v2.2.3

Documentation update, added docstrings to all the tools for documentation (automatically harvestable for metadata)

(Released by Maarten van Gompel on 2019-09-09) https://github.com/proycon/foliatools/releases/tag/v2.2.3

foliatools v2.2.2

Bugfix release:

- [tei2folia] prevent generating empty folia comments in rare cases (La...
- [tei2folia] fix in t-ref, text ended up outside attribute
- [folialangid] language constraints were not propagated to underlying ...
- [folialangid] implemented confidence threshold and fallback language

(Released by Maarten van Gompel on 2019-08-18) https://github.com/proycon/foliatools/releases/tag/v2.2.2

foliatools v2.2.1

Minor bugfix release:

- [folialangid] exit code was wrong

(Released by Maarten van Gompel on 2019-07-30) https://github.com/proycon/foliatools/releases/tag/v2.2.1

foliatools v2.2.0

- [folialangid] adding support for using a different textclass as source
- [folialangid] fix in conflicting options for folialangid
- [tei2folia] Small adaptation for part annotation and spacing (proycon/folia#61)

(Released by Maarten van Gompel on 2019-07-23) https://github.com/proycon/foliatools/releases/tag/v2.2.0

foliautils v0.13

[Ko vd Sloot]

bug fixes:

- fix for https://github.com/LanguageMachines/foliautils/issues/35
- fix for https://github.com/LanguageMachines/foliautils/issues/36

[Maarten van Gompel]

new features:

- FoLiA-wordtranslate.cxx: frog should be able to deal with spaces now, no need for ugly non-breaking space hack that now causes other problems in frog (LanguageMachines/frog#77)
- FoLiA-wordtranslate.cxx: adding ability to constrain by language

(Released by Ko van der Sloot on 2019-08-30) https://github.com/LanguageMachines/foliautils/releases/tag/v0.13

libfolia v2.3

new features:

- autodeclare mode introduced (as in FoLiApy)
- folialint by default doesn't autodeclare. use -a or --autodeclare to use it
- Better detecting of declaration errors in general
- the select function now also enables the possibility to search recursively upto the first matching sibling.
- bumped library version to 13 other changes:
- some exceptions are changed.
- less exceptions are thrown. An empty result is returned instead.
- folialint now accept bote -d and --debug
- real fix for issue70
- small bug fixes and refactorings
- accept empty Comment and Description nodes

(Released by Ko van der Sloot on 2019-08-29) https://github.com/LanguageMachines/libfolia/releases/tag/v2.3

libfolia v2.2.1

Bug fix release:

- There were some problems handling NO setname vs. EMPTY setname, during incremental parsing in folia::Engine. This was sorted out now: https://github.com/proycon/folia/issues/74 This related to some ucto and frog issues too: https://github.com/LanguageMachines/ucto/issues/70 https://github.com/LanguageMachines/frog/issues/72

(Released by Ko van der Sloot on 2019-07-22) https://github.com/LanguageMachines/libfolia/releases/tag/v2.2.1

libfolia v2.2

Bug fix release.

- Folia::Engine choked on some complex FoLiA. Solved by refactoring and in fact simplifying some code. (Frog issue #77 revealed this)
- added flush() on document output to streams. (frog issue #72)
- improved output in debugging mode

(Released by Ko van der Sloot on 2019-07-15) https://github.com/LanguageMachines/libfolia/releases/tag/v2.2

Frog

Project/task/deliverable references: CLARIAH-PLUS WP3 T139; CLARIAH-CORE WP3 T22 [D2.1 (software); D2.2 (doc)]; T23 [D2.3 (froggen software), D2.4 (froggen docs)]

frog v0.18.6

- should really fix: https://github.com/proycon/tscan/issues/13

(Released by Ko van der Sloot on 2019-09-17) https://github.com/LanguageMachines/frog/releases/tag/v0.18.6

frog v0.18.5

Bug fix release.

- Fixes: https://github.com/proycon/tscan/issues/13

(Released by Ko van der Sloot on 2019-09-16) https://github.com/LanguageMachines/frog/releases/tag/v0.18.5

frog v0.18.3

Bug fix release:

- Fixes https://github.com/LanguageMachines/frog/issues/78

(Released by Ko van der Sloot on 2019-07-22) https://github.com/LanguageMachines/frog/releases/tag/v0.18.3

frog v0.18.2

Bug fix release:

- Fixes for:
    – https://github.com/LanguageMachines/frog/issues/75
    – https://github.com/LanguageMachines/frog/issues/77

(Released by Ko van der Sloot on 2019-07-15) https://github.com/LanguageMachines/frog/releases/tag/v0.18.2

python-frog v0.4.1

Minor update fixing a compatibility issue on some macs

(Released by Maarten van Gompel on 2019-08-07) https://github.com/proycon/python-frog/releases/tag/v0.4.1

LaMachine

Project/task/deliverable references: CLARIAH-PLUS WP3 T098; CLARIAH CORE WP3: LaMachine v2 plan in scope of CLARIAH WP3 VRE

codemetapy v0.2.2

Minor update to entrypoint extension: attempt to automatically read the docstring for each entrypoint and use it as a description for the entrypoint metadata

(Released by Maarten van Gompel on 2019-09-09) https://github.com/proycon/codemetapy/releases/tag/v0.2.2

LaMachine v2.10.0

- use latest available java version
- version bump because of other releases (FLAT, libfolia, ticcutils)

(Released by Maarten van Gompel on 2019-09-01) https://github.com/proycon/LaMachine/releases/tag/v2.10.0

LaMachine v2.9.1

- Added colibri-utils (except on mac)
- Adding user to vboxsf group for shared folder access (if it exists)
- ignore errors on creating webserver dirs (in case they already exist or are links)
- don't let nvm problems prevent lamachine activation

(Released by Maarten van Gompel on 2019-08-27) https://github.com/proycon/LaMachine/releases/tag/v2.9.1

LaMachine v2.9.0

- extract list of packages and ask to open installation manifest in editor or not (relates to #63), same thing for the configuration file
- Removing some packages from the default installation to conserve space:
    – Removing cltl-vu-python from the default packages (the folia-naf converter is not really in use nor maintained well nowadays and the git repo comes with some big files that take up too much space)
    – exclude xgboost from the default installation as it takes up a lot of space and is not a direct dependency

- add extra user input stage with updated instruction prior to ansible start. This ensures people know to expect a lot of output they can ignore and a long wait whilst things are installing, and so understand what to do when ansible asks for the BECOME password (addressed an issue raised by @pboot)
- fix for failure to start nginx #151
- adapting the strict python check and making the bootstrap also work when python 2 is absent in some cases
- ensured ~/bin is in $PATH immediately after bootstrap
- added more feedback to lamachine-start-webserver script
- implemented CSV export ability for lamachine-list
- do linux distribution detection and warn users if they run on an unsupported or outdated platform
- [documentation] updated FAQ

(Released by Maarten van Gompel on 2019-06-28) https://github.com/proycon/LaMachine/releases/tag/v2.9.0

Miscellaneous

Project/task/deliverable references: Dependencies/wrappers and or unforeseen tools (related to CLARIAH projects)

ticcutils v0.22

Major update:

- removed dependency on boost::regex. As a consequence, we DO NOT support gcc < 4.9 anymore.
- added the json.hpp include file from: https://github.com/nlohmann/json/ This opens the possibility to enable JSON support in server-like applications build on ticcutils.
- bumped the library version to 6
- some small changes:
    - clearer error message in ServerBase.cxx
    - added a 'protocol' option to ServerBase.cxx
    - added an extra function getAttributes() to XmlTools

(Released by Ko van der Sloot on 2019-08-29) https://github.com/LanguageMachines/ticcutils/releases/tag/v0.22

Nederlab

Project/task/deliverable references: Nederlab

colibri-utils v0.2.1

Minor update, only updated metadata

(Released by Maarten van Gompel on 2019-09-09) https://github.com/proycon/colibri-utils/releases/tag/v0.2.1

colibri-utils v0.2

Compatibility with latest libfolia

(Released by Maarten van Gompel on 2019-08-30) https://github.com/proycon/colibri-utils/releases/tag/v0.2

colibri-utils v0.1

First release of colibri-utils, containing colibri-lang, a language detection tool

(Released by Maarten van Gompel on 2019-08-20) https://github.com/proycon/colibri-utils/releases/tag/v0.1

PICCL & TICCL

Project/task/deliverable references: CLARIAH-PLUS WP3 ???; CLARIAH-CORE WP3 T26 [D1.3 (PICCL software); T1.4 (PICCL docs)]

PICCL v0.8.2

(Previous release was premature and didn't pass tests, fixed in this one)

(Released by Maarten van Gompel on 2019-08-25) https://github.com/LanguageMachines/PICCL/releases/tag/v0.8.2

PICCL v0.8.1

Intermediate bugfix release (#53 is still open):

- attempted fix for FoLiA-correct invocation (#50 #52)

(Released by Maarten van Gompel on 2019-08-25) https://github.com/LanguageMachines/PICCL/releases/tag/v0.8.1

Software Quality Guidelines

Project/task/deliverable references: CLARIAH-CORE WP2 Task 54.100

(no releases this period)

Timbl & Mbt

Project/task/deliverable references: Pre-CLARIN

(no releases this period)

Ucto

Project/task/deliverable references: CLARIAH-CORE WP3 T55 [D1.1 (software), D1.2 (docs)]

python-ucto v0.5.1

Fix release for compatibility with moved libxml2 on (some?) macs. (proycon/LaMachine#154)

(Released by Maarten van Gompel on 2019-08-07) https://github.com/proycon/python-ucto/releases/tag/v0.5.1

ucto v0.19

Bug fix release. solving:

- https://github.com/LanguageMachines/ucto/issues/72
- some problems with the newest libfolia.
- better provenance records

(Released by Ko van der Sloot on 2019-09-16) https://github.com/LanguageMachines/ucto/releases/tag/v0.19

ucto v0.18

Bug fix release. solving: https://github.com/LanguageMachines/ucto/issues/70

(Released by Ko van der Sloot on 2019-07-22) https://github.com/LanguageMachines/ucto/releases/tag/v0.18