Software Release and Progress Report

Start date: 2018-06-22 End date: 2018-10-22

Short Summary

[Maarten van Gompel]

In no particular order:

- This period various improvements have been made to LaMachine.
    - added Oersetter for Fryske Akademy
    - adding Frisian-Dutch ASR (still in progress)
- Started work on the next major version (FoLiA v2), which features a completely revised and more formal documentation, as well as support for provenance data (https://github.com/proycon/folia/issues/46), which should be of benefit to the VRE. This can be considered a head-start for CLARIAH-PLUS.
- The Frog Named Entity Recogniser is being extended gazetteers with more fine-grained categories (https://github.com/LanguageMachines/frog/issues/59)
- Discussions with Antkse (VU) on the NAF-FoLiA converter and a shared pipeline, work is slowly starting again
    - Also trying to encourage the VU to participate in LaMachine ;)
- Discussions with Dieter van Uytvanck on Single-Sign-On for Language and Speech Tools hosted at Radboud (another CLARIAH-PLUS item)
- Attended the CLARIAH Provenance Workshop and gave a presentation on software metadata
    - Discussions with Jan Odijk about high-level centralised software metadata (CMDI) and lower-level metadata (codemeta), which both have good uses and can be complementary
- There is 3rd party interest in FLAT for syntactic movement annotation, leading to implementation requests/fixes that would also be covered in our CLARIAH-PLUS proposals (https://github.com/proycon/flat/issues/138)

[Ko van der Sloot]

- (TODO: add manually)

[Henk van den Heuvel; RDM & Data Curation]

- (TODO: add manually)

[Iris Hendrickx]

- Attended the CLARIN conference in Pisa and presented a poster on LaMachine (https://github.com/proycon/LaMachine/blob/r
- Working on a complete overhaul and long needed update of the ucto documentation

CLAM

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

clam v2.3.4

Minor update:

- Added codemeta metadata
- Extra information on output templates on `info/` page (accommodating a request by @henkvdheuvel)

(Released by Maarten van Gompel on 2018-08-28) https://github.com/proycon/clam/releases/tag/v2.3.4

FLAT

Project/task/deliverable references: CLARIAH WP3 T24 [D2.10 (software); D2.11 (docs)]

flat v0.7.14

Bugfix release:

- Fixes bug in adding certain span annotation elements (#138)

(Released by Maarten van Gompel on 2018-10-09) https://github.com/proycon/flat/releases/tag/v0.7.14

flat v0.7.13

Added support for URL prefix, i.e. hosting FLAT not at the root, needed by LaMachine #131

(Released by Maarten van Gompel on 2018-08-12) https://github.com/proycon/flat/releases/tag/v0.7.13

foliadocserve v0.6.8

Bugfix release:

- Fixed issue #14 (Metadata changes weren't saved properly if no other annotations were made)

(Released by Maarten van Gompel on 2018-09-19) https://github.com/proycon/foliadocserve/releases/tag/v0.6.8

FoLiA

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

NAFFoLiAPy v0.1.1

A release of the current status-quo with some latest fixes, still much to do

(Released by Maarten van Gompel on 2018-06-29) https://github.com/cltl/NAFFoLiAPy/releases/tag/v0.1.1

Frog

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

(no releases this period)

LaMachine

Project/task/deliverable references: CLARIAH WP3: LaMachine v2 plan in scope of CLARIAH WP3 VRE

codemetapy v0.2.1

- better failure and exit code if identifier was not found in registry

(Released by Maarten van Gompel on 2018-10-08) https://github.com/proycon/codemetapy/releases/tag/v0.2.1

codemetapy v0.2.0

- Added some simple support for converting debian package metadata from apt show to codemeta (#1)

(Released by Maarten van Gompel on 2018-09-17) https://github.com/proycon/codemetapy/releases/tag/v0.2.0

codemetapy v0.1.6

Minor bugfix release

(Released by Maarten van Gompel on 2018-08-31) https://github.com/proycon/codemetapy/releases/tag/v0.1.6

codemetapy v0.1.5

Minor update release:

- Added `--with-orcid` parameter to generate placeholders for ORCIDs in author details (#2)

(Released by Maarten van Gompel on 2018-08-30) https://github.com/proycon/codemetapy/releases/tag/v0.1.5

labirinto v0.2.4

- Added `REWRITE_HOST` configuration parameter that automatically rewrites the specified hosts (if occurring in entryPoints.urlTemplate) to the one determined at run-time, also introduced a `{host}` variable that does the same explicitly (proycon/LaMachine#97)
- Adding support for CLARIN specific attributes
- Added tooltips
- Show `dateCreated` property in interface

(Released by Maarten van Gompel on 2018-06-24) https://github.com/proycon/labirinto/releases/tag/v0.2.4

LaMachine v2.4.5

- Added Oersetter Frisian-Dutch MT (#121)
- Added Moses (#110)
- Metadata registration fixes

(Released by Maarten van Gompel on 2018-10-17) https://github.com/proycon/LaMachine/releases/tag/v2.4.5

LaMachine v2.4.4

Bugfix release:

- Added compatibility for BSD sed (on mac), not just GNU sed

(Released by Maarten van Gompel on 2018-10-12) https://github.com/proycon/LaMachine/releases/tag/v2.4.4

LaMachine v2.4.3

- Fixes in bootstrap for prebuilt VM

(Released by Maarten van Gompel on 2018-10-12) https://github.com/proycon/LaMachine/releases/tag/v2.4.3

LaMachine v2.4.2

- Added an `extra_disksize` variable for VM (`--disksize` parameter for bootstrap), configures an extra virtual disk to put LaMachine on in order to allow installation of large software that would not fit in the default disk image.
- Allow allow Tesseract German Fraktur installation to fail (not present on all distros, removed from ubuntu 18.04) should solve #117
- Do not require python-core for kaldi
- Various small fixes
- Adding fy_nl_ASR webservice #116 (not fully functional yet)
- Attempt to get kaldi working on Arch Linux still failed for now

- Added LaMachine poster presentation for CLARIN conference in Pisa

(Released by Maarten van Gompel on 2018-10-08) https://github.com/proycon/LaMachine/releases/tag/v2.4.2

LaMachine v2.4.1

Minor update release:

- Register metadata for debian packages (proycon/codemetapy#1)
- Register versions for nextflow and alpino

(Released by Maarten van Gompel on 2018-09-19) https://github.com/proycon/LaMachine/releases/tag/v2.4.1

LaMachine v2.4.0

- Integrated Jupyter Labs as a web-based notebook/IDE/terminal environment (#108)
- Included Python Course for the Humanities by @fbkarsdorp (and myself)
- Added `lamachine-add` script as a way to add software, i.e. automatically edit the installation manifest #111
- Added `lamachine-passwd` script to set passwords
- Improved documentation (README), included screenshots
- Fix in delegation to FLAT from CLAM webservices (viewer)
- Added webserver support for macOS virtualenv flavour, but still limited (uwsgi doesn't work yet, so webservices don't work) #114
  - **Upgrade notice**: If you come from an earlier installation you will need to set `webserver: true` in your LaMachine configuration for this to take effect
- Added Stanford CoreNLP (#60)
- Added wopr for macOS virtualenv flavour (i.e. added to homebrew now)
- Added extra utilities for data science (extra-utils), provides: jq, ack, graphviz, gnuplot, glances, csvkit, xmlstarlet and a lot more
- Added conll-u parse module to python-core

(Released by Maarten van Gompel on 2018-09-06) https://github.com/proycon/LaMachine/releases/tag/v2.4.0

LaMachine v2.3.0

- Always upgrade pip, as we want to rely on the newer `--upgrade-strategy only-when-needed` flag, preventing unnecessary updates of dependencies if not explicitly requested.
- An explicit `PYTHONPATH` is now set and used for global LaMachine installations (affects all VM and Docker installations). Users are recommended to start anew with a fresh LaMachine VM or Docker container if possible, as the regular `lamachine-update` will leave artefacts.
- Fixed CentOS 7 regression problem caused by v2.2.12
- FLAT integration completed #49
- Implemented (limited) custom version support (#51)
  - Custom versioning on macOS is not supported (and probably won't be supported in the future either unless there is a huge demand)
  - Custom versioning necessarily excludes packages supplied by the linux distribution's package manager. To ensure getting a LaMachine environment close to the original, ensure you build it on the same distribution.
  - Versioning of java/maven software is not supported yet (but will be in the future)
  - Some software is excluded as either upstream doesn't provide the necessary release history, or it needs more effort on my part to support that release history as it deviates from the default, examples are: Nextflow, Alpino

(Released by Maarten van Gompel on 2018-08-13) https://github.com/proycon/LaMachine/releases/tag/v2.3.0

LaMachine v2.2.12

- Added Stanford CoreNLP #60
    - Added support for pulling software from Maven Central
    - Set a default java classpath
- Fixed locale problems, set locale for docker and vagrant, always prevent C or POSIX locales #107
- Added PyTorch
- Always update pip (a proper recent pip that supports the upgrade policy parameter is needed)
- Detect if the virtualenv python diverges from the global one, and suggest reinstallation of LaMachine in that case #105
- Some Travis-CI fixes

(Released by Maarten van Gompel on 2018-08-09) https://github.com/proycon/LaMachine/releases/tag/v2.2.12

LaMachine v2.2.11

Minor fixes:

- Set $MANPATH in activation script (should solve #104)
- fixed hunspell-de package #102

(Released by Maarten van Gompel on 2018-07-15) https://github.com/proycon/LaMachine/releases/tag/v2.2.11

LaMachine v2.2.10

- added networkx to default python packages to install
- Added NAFFoLiAPy to LaMachine after first release
- Adding zip/unzip #100
- Disabling allow-unauthenticated parameter to apt module (solves #101)
- Bootstrap had no hashbang
- Adding pcre libraries to global dependencies, required in case of manual nginx compilation
- Updated kaldi metadata #32
- Added ubuntu 18.04 venv test
- Added hunspell tool and dictionaries if gecco is chosen #102

(Released by Maarten van Gompel on 2018-07-12) https://github.com/proycon/LaMachine/releases/tag/v2.2.10

LaMachine v2.2.8

Bugfix release:

- uwsgi template somehow failed to produce a newline, causing bug #95, this should fix it
- Set REWRITE_HOST for labirinto (v0.2.4) to allow rewriting hostname with the one detected at run time, allows accessing underlying services when not resolvable
- fixed bug in previous fix for #88 #94, hopefully now fixed for good
- documentation: add hostname to recommended docker command after bootstrap #97

(Released by Maarten van Gompel on 2018-06-24) https://github.com/proycon/LaMachine/releases/tag/v2.2.8

Miscellaneous

Project/task/deliverable references: Dependants of others

(no releases this period)

PICCL & TICCL

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

PICCL v0.6.3

- Better input validation #37
- Split list of enrichment steps in webinterface #33
- Minor fixes and increased debug in webservice script

(Released by Maarten van Gompel on 2018-07-12) https://github.com/LanguageMachines/PICCL/releases/tag/v0.6.3

Software Quality Guidelines

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

(no releases this period)