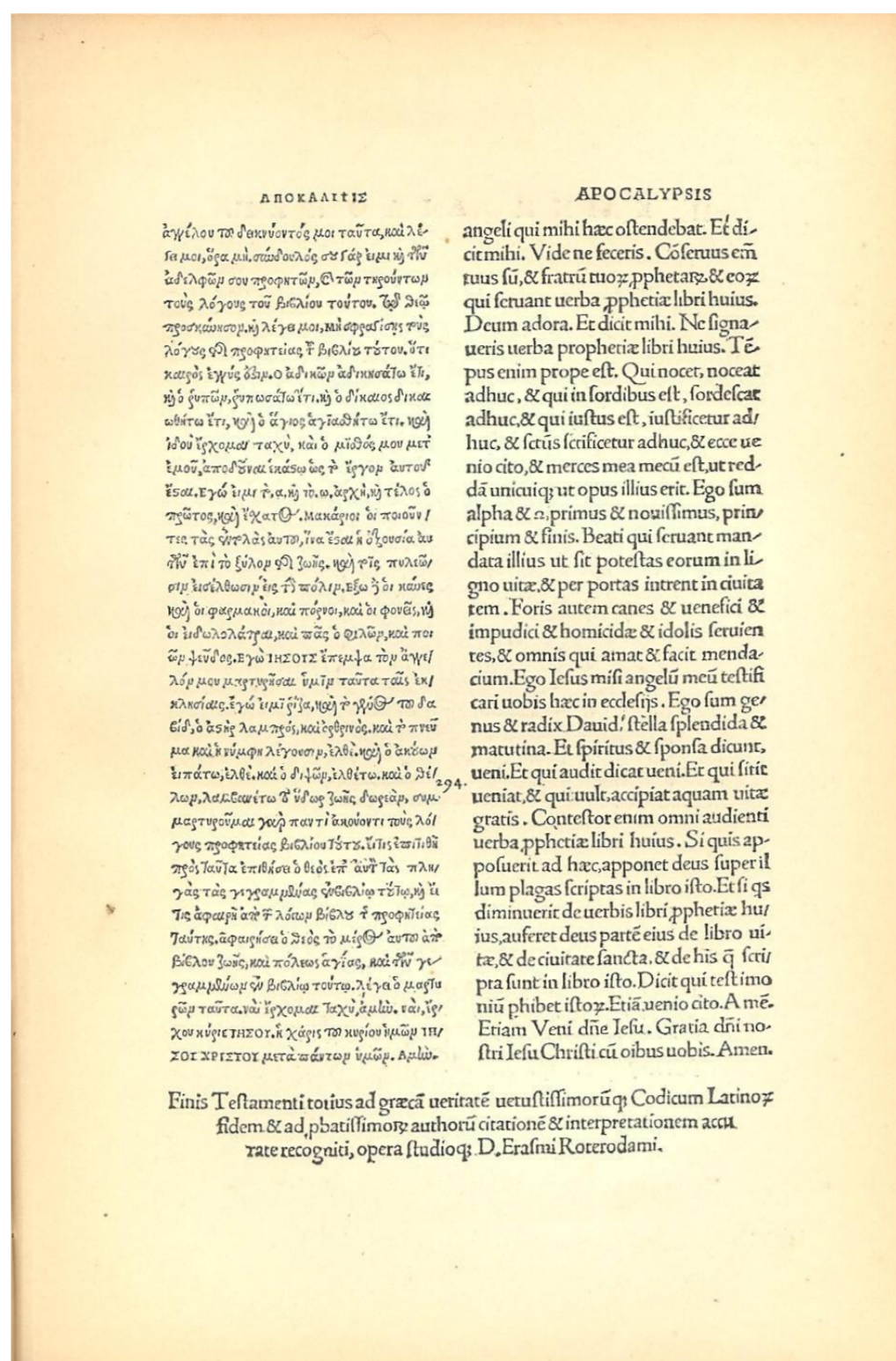


Automatic creation of tool workflows in CLARIN-DK



For hundreds of years, scholars have analysed, annotated and interpreted sources that were available to them, applying the insights and technology of their time.



In our time, researchers in humanities have access to enormous amounts of data.

There also exist a great variety of tools that can analyse, annotate and transform all those data, in ways that were unimaginable in the past.

But often a single tool cannot do what the researcher wants to attain. Instead she needs to take data through a number of steps, a *workflow*.

Planning a workflow can be very complicated, requiring a lot of training and experience.



As systems become more complex, the risk of human oversights during planning grows and may result in failure to reach the set goal.

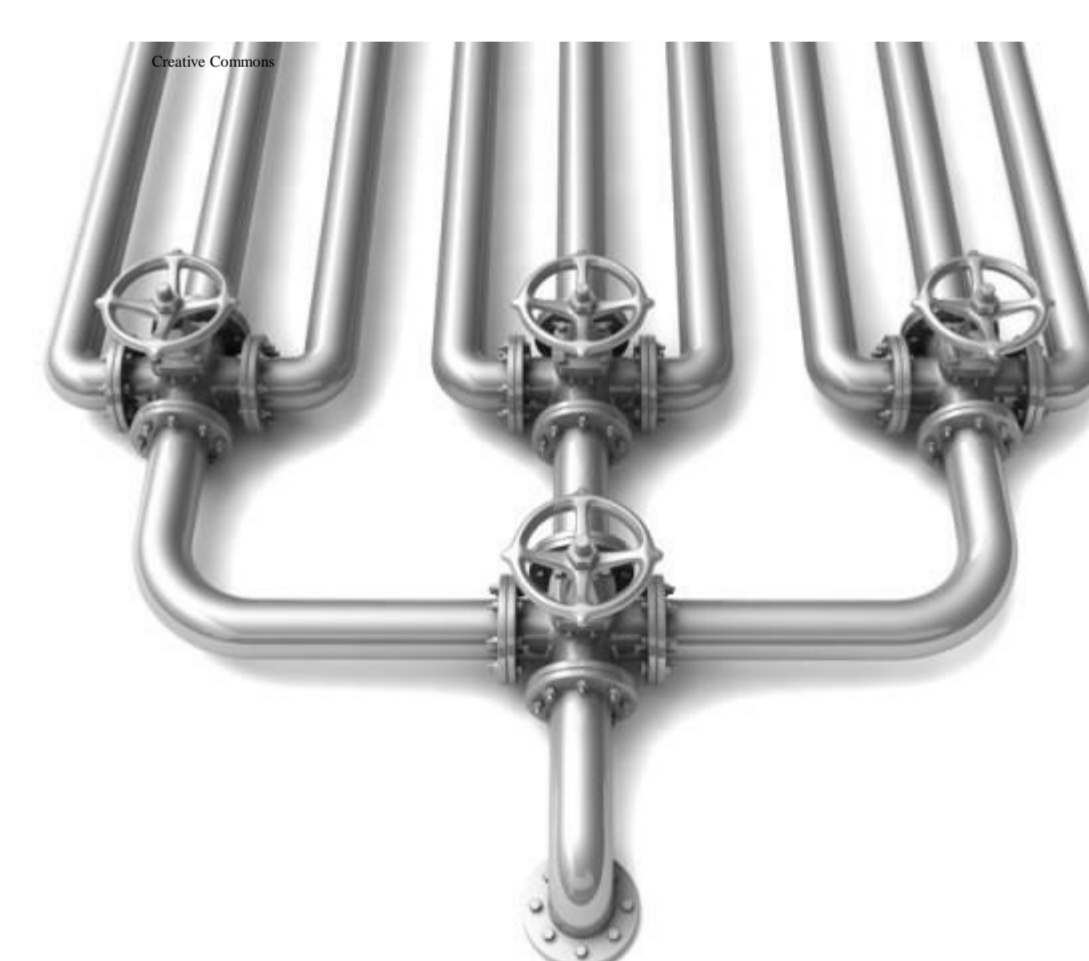


CLARIN-DK allows researchers in humanities to concentrate on what they are good at, delegating workflow planning and execution almost completely to the machine. The researcher only has to ...

... provide input ...



... specify a goal ...

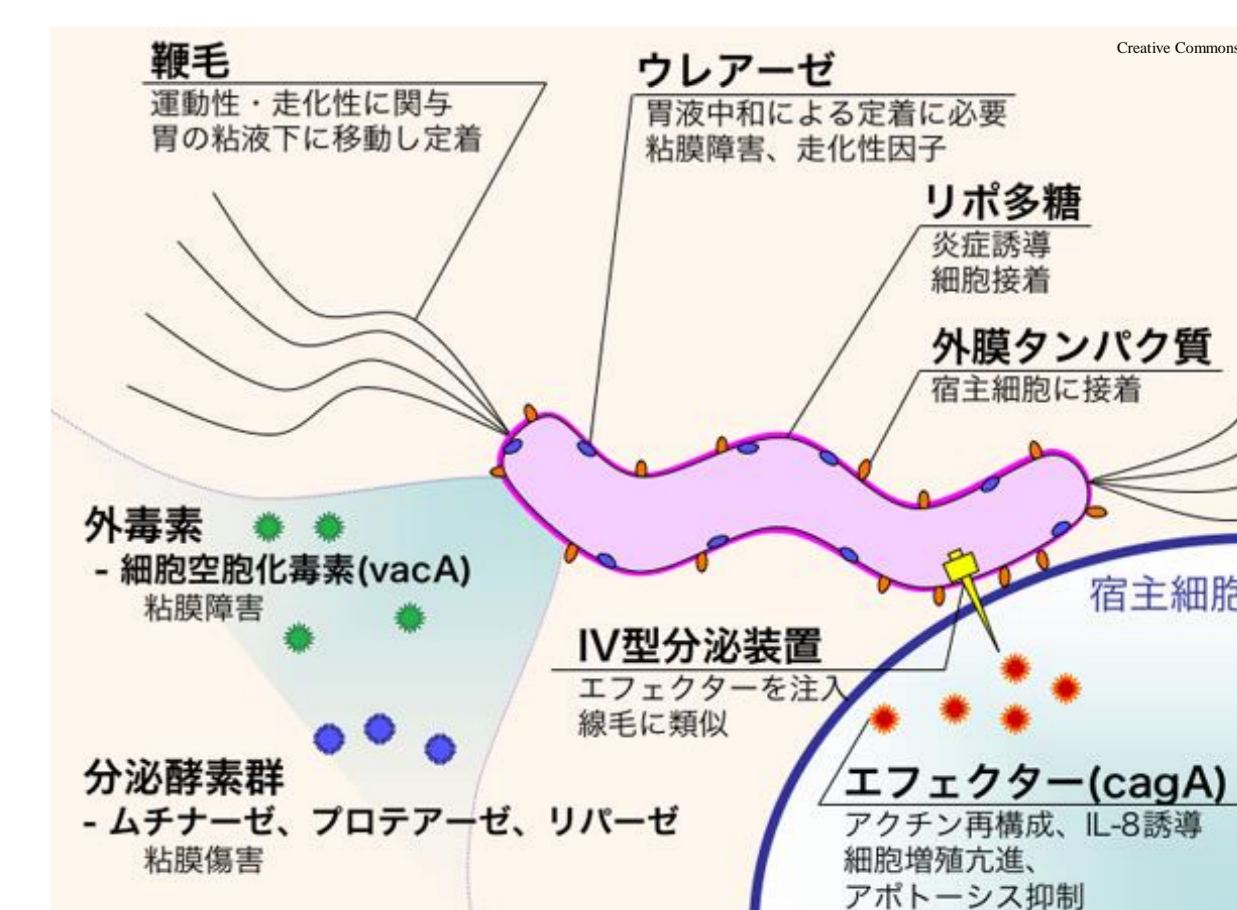


... select one of the proposed workflows that can realize the goal ...



... press this ...

... and wait for the output to arrive.



Check whether the input matches the goal.



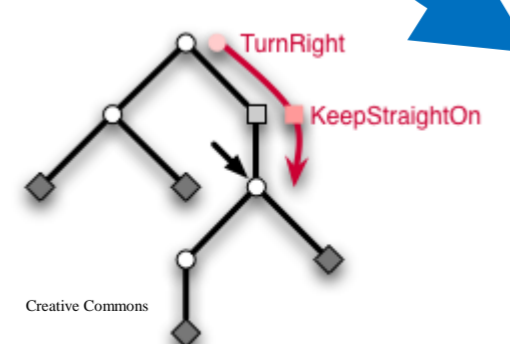
Find a tool in the toolbox ...



... with an output profile that matches the current goal.



Then move the goal to the input profile of that tool.



Backtrack and find alternative workflows that also satisfy the initial goal, threading through tools in all combinations that lead to success.

Let the researcher choose one of the workflows, execute that and present the result.



Registered tool metadata

- Boilerplate:
- Tool Name
 - URL of webservice
 - Description of tool
 - ...
- I/O Profile:
- type of content
 - style
 - format
 - style
 - language
 - style
 - [... more features ...]

For each feature, the profile describes input and output.

At least one feature must be profiled.

Style information is optional. (e.g. "Penn Treebank")