Developing a service orchestrator for the annotation platform AdNotare / PACTE

Lise Rebout
Computer Research Institute of Montreal

Canadian Research Software Conference 2018
Background

AdNotare / PACTE: Collaborative Electronic Text Annotation Platform

NLP pipelines

simple example of temporal expression tagging

Orchestrator requirements

- Call the different annotators in parallel or in sequence
- Integrated as an AdNotare / PACTE service (Celery Worker)
- Flexible: designing new workflows
- Lightweight: workflow parameters as JSON files
Our solution

**Workflow Language**: Common Workflow Language ([CWL](https://cwltool.io/))

- Open standard: follows the Open-Stand.org principles
- Explicitness and isolation of the tasks
- Parameters’ definition in JSON or YAML

**Workflow Engine**: [Toil](https://github.com/uctcscl/toil) (UCSC Computational Genomics Lab)

- Open-source pure-Python workflow engine

**Deployment / Containerization**:

- Docker Container as a Celery Worker

**Next steps**:

- Cancellation of tasks
Our solution

**Workflow Language**: Common Workflow Language ([CWL](https://cwl.readthedocs.io/en/latest/))
- Open standard: follows the Open-Stand.org principles
- Explicitness and isolation of the tasks
- Parameters’ definition in JSON or YAML

**Workflow Engine**: [Toil](https://toil.readthedocs.io/en/latest/) (UCSC Computational Genomics Lab)
- Open-source pure-Python workflow engine

**Deployment / Containerization**:
- Docker Container as a Celery Worker

**Next steps**:
- Cancellation of tasks
**Our solution**

**Workflow Language**: Common Workflow Language (CWL)
- Open standard: follows the Open-Stand.org principles
- Explicitness and isolation of the tasks
- Parameters’ definition in JSON or YAML

**Workflow Engine**: Toil (UCSC Computational Genomics Lab)
- Open-source pure-Python workflow engine

**Deployment / Containerization**:
- Docker Container as a Celery Worker

**Next steps**:
- Cancellation of tasks