

# A Workflow for Polarized Light Imaging Using UNICORE Workflow Services

June 24, 2014 | [B. Hagemeier](#) O. Bücker A. Giesler R. Saini B. Schuller

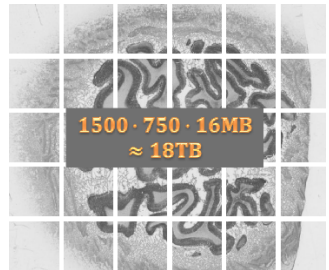
## Overview

- Polarized Light Imaging
- Manual processing
- UNICORE workflow
- Technical challenges

# Polarized Light Imaging (PLI)

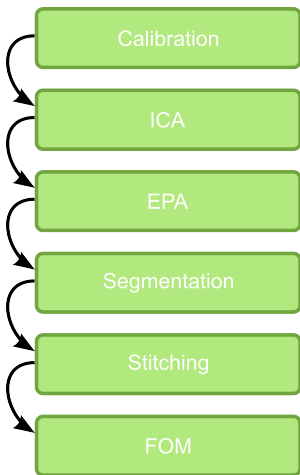
## The Method

- Developed at Institute for Neuroscience and Medicine, Forschungszentrum Jülich
- Aim: extract the course of single nerve fibers
- Sections of postmortem human brain, each  $70\mu m$
- Imaging the sections under linearly polarized light
  - $\sim 1500$  sections/brain
  - 30x25 tiles/section
  - 2048x2048 pixels/tile



# Polarized Light Imaging

## General Process

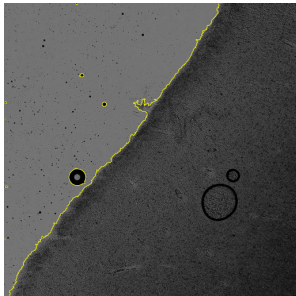


- Much of the data needed in many steps of the process

# Identifying Brain Matter

## Segmentation by Seeded Region Growing

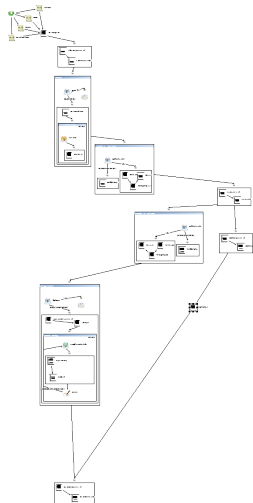
- Used to separate brain matter from background
- Requires manual intervention to determine threshold



# UNICORE Workflow

## General Structure

- Workflow variables to parameterize the workflow
- Housekeeping jobs to prepare actual jobs



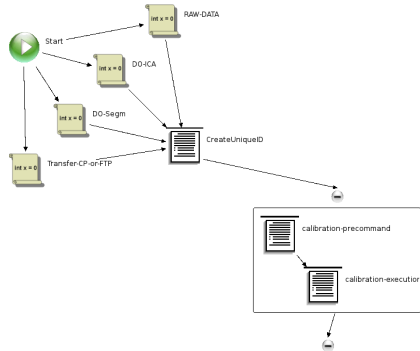
# UNICORE Workflow

## Data Handling

- Moving 1TB of data between workflow steps consumes time
- Duplicating data is a waste of space
  - Re-using data from the same location between workflow steps works better
  - Transport only references to local working directories between jobs
  - Use workflow parameters to select storage and storage mount point for workflow execution

# UNICORE Workflow

## Housekeeping Jobs



- Setup shared working directories for each workflow instance
- UC\_PREFER\_INTERACTIVE\_EXECUTION
  - Undocumented feature in TSI
  - Set variable in job to true to run it on the login node



## Individual workflow directory

### CreateUniqueID

#### Idea

- Create working directory accessible by all/most jobs in a workflow
- Reduce or avoid file transfers or even copies

```
1 tmp=$USER.XXXXXXXXXX
2 WFDIR=$(mktemp -u ${tmp}) || { echo "Failed ..." }
3 echo ${WFDIR}>PLI_unique_workflow_dir.txt
```

- Make information available in an exported file
- Output variables for workflow jobs would be an alternative

# Individual workflow directory

## Making use of it

```
1 ## Read working directory name ##
2 WFDIR='sed -n 1p
      CreateUniqueID_PLI_unique_workflow_dir.txt'
3 USPACE='pwd'
4
5 ...
```

- File name should actually be taken from variable

# UNICORE Workflow

## File List Generation

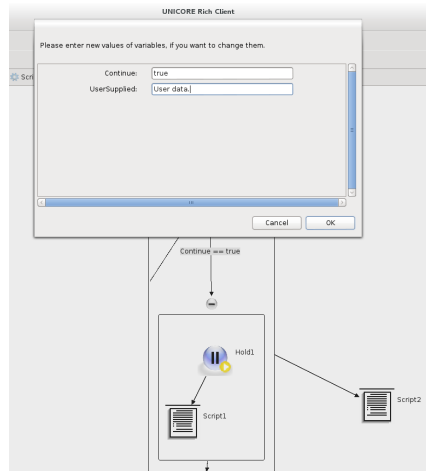
- For-Loops can process file lists read from files
- File list within file must contain storage locations
- Generate file list including storage location
  - supplied via workflow variable or statically coded

```
1  ...
2  ## Generate file list ##
3  cd ${WORK}
4  FILELIST='find PLI/${WFDIR}/caliOutputDir -type f'
5
6  for file in ${FILELIST}; do
7      echo 'BFT:https://fzj-unic.fz-juelich.de:9112/ \
8          FZJ_JUDGE/services/StorageManagement? \
9          res='${USER}'-WORK#'$${file} >>
          ${USPACE}/generated.txt
10 done
```

# URC Improvements

## Hold and Resume

- Generally possible in 6.x.x
- No means in URC to provide values for WF variables
- Workaround: provide values through a file modified during hold
- 7.x.x: WF engine and URC allow for supplying values on resume



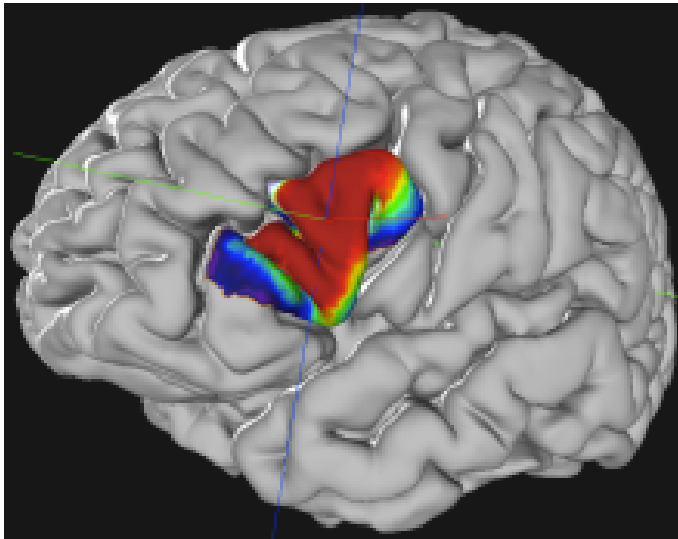
## Summary

### PLI method

- Extract brain fiber directional information by imaging with polarized light
- A number of tools involved
- Can be processed semi automatically

### UNICORE aspects

- Making use of workflow variables to parameterize workflows
- Handling large data sets between workflow jobs
- Interaction with workflow (hold and resume)
- Preparing file lists for subsequent jobs
  - Use of UNICORE service information within jobs desirable
- Make use of “housekeeping” jobs



Institute for Neuroscience and Medicine, Forschungszentrum Jülich