

How to track Workflows  
for Painless Result Reproduction?



**Provenance!**

*Business Information Systems*

Cologne



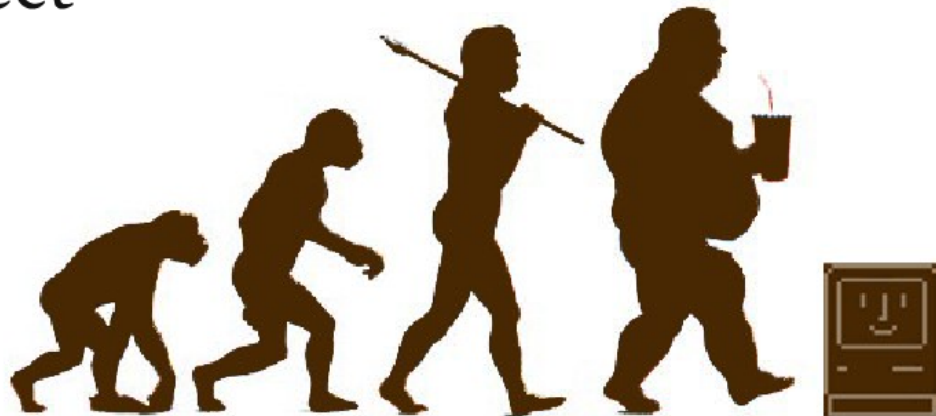
David Uhlig

*Summer Student*

Andreas Adelman

# What is Provenance?

- Provenance, from the French *provenir*, »to come from«, means the origin, or the source of something
- Provenance is accomplished by tracing the whole history of an object



- Provenance, a matter of documentation!

tracing of the whole history



Spiced Carrot Cake

Delicious!  
★★★★★



documentation



Why is Provenance important?



Reproduction needs recipe!

Automatically track and manage the provenance of processes



# VisTrails

- University of Utah
- OpenSource
- Licensed under GPL v3
- Python, Qt

Versioning

Standardization



4 Reasons for  
Provenance + VisTrails

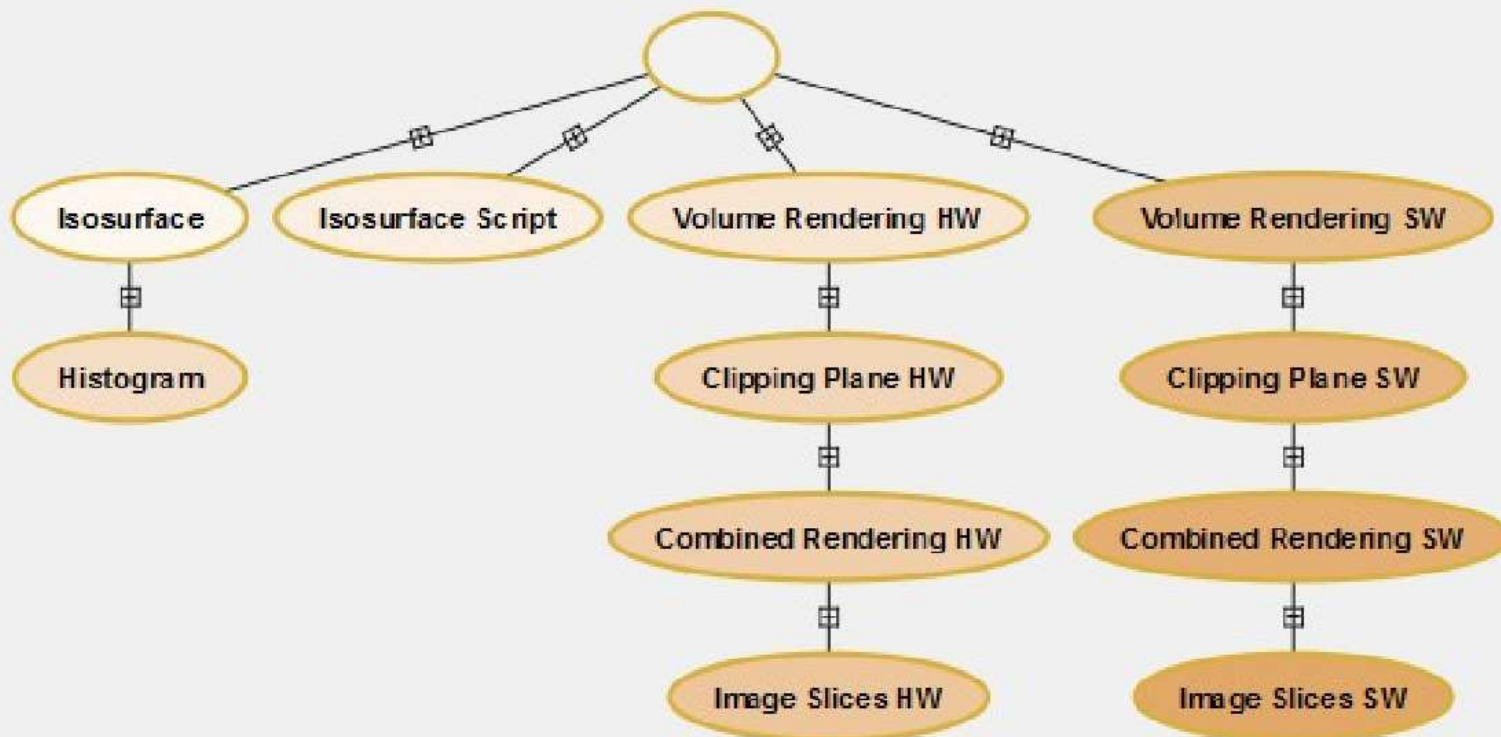


Automatization

Collaboration

# Advantages of VisTrails for SwissFEL + OPAL

i) Keeps up with Workflow changes



## Advantages of VisTrails for SwissFEL + OPAL

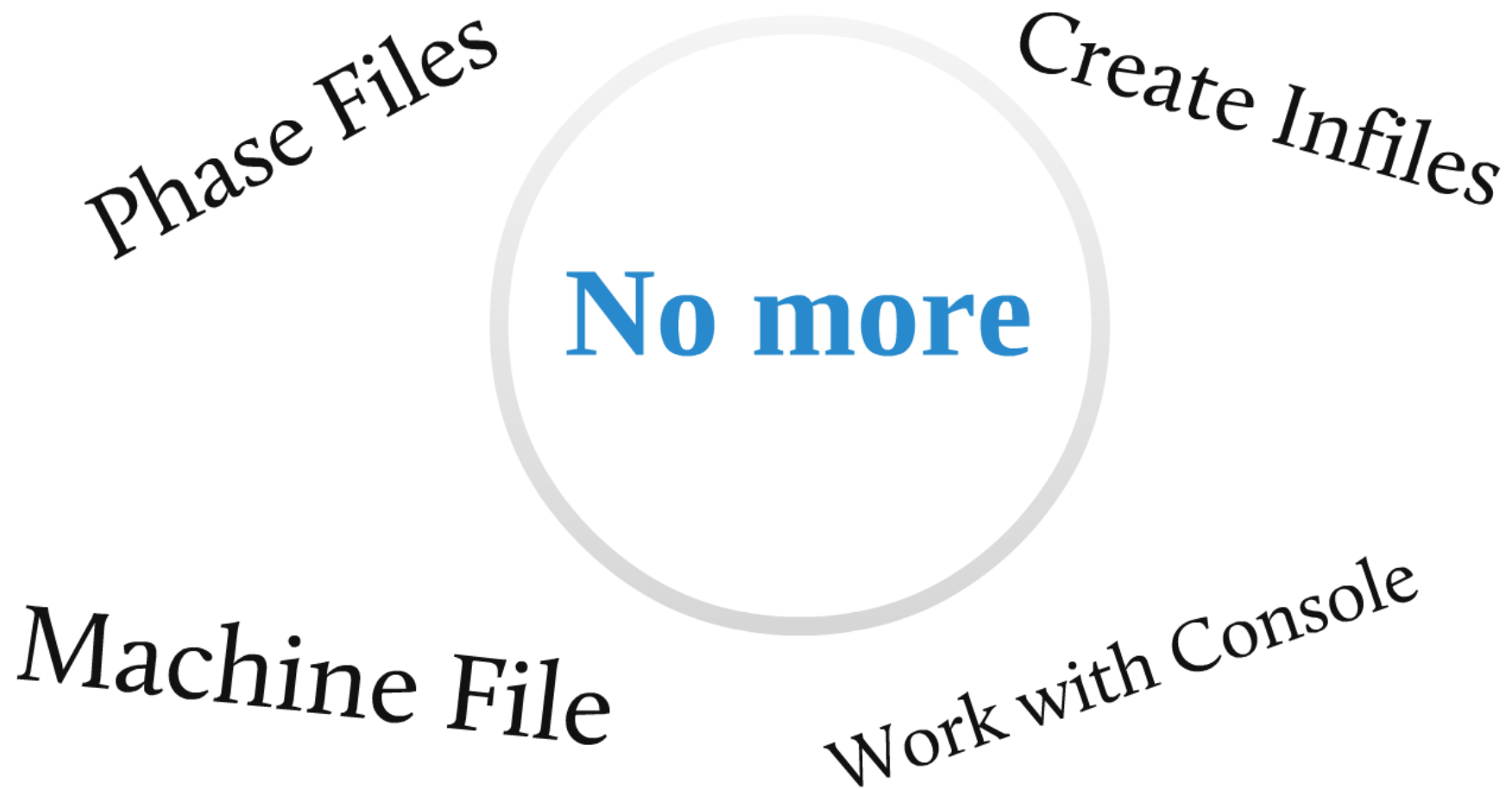
2) Collecting simulation results to one central place

- Accessible for everyone
- Reducing redundancy
- Enhancing efficiency (less time for searching)



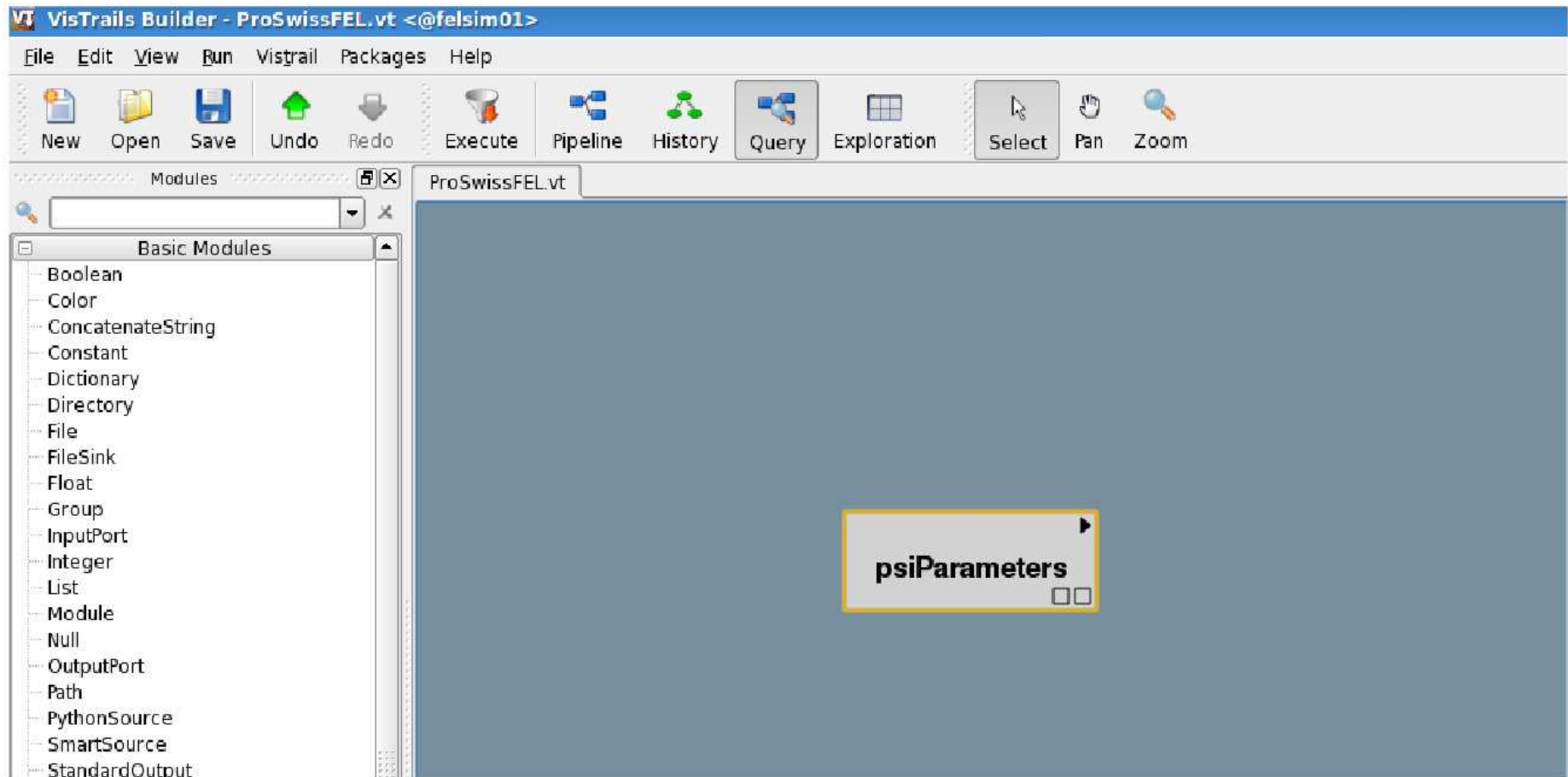
# Advantages of VisTrails for SwissFEL + OPAL

3) Simplifies process to get simulations started



# Advantages of VisTrails for SwissFEL + OPAL

## 4) Easy to find results



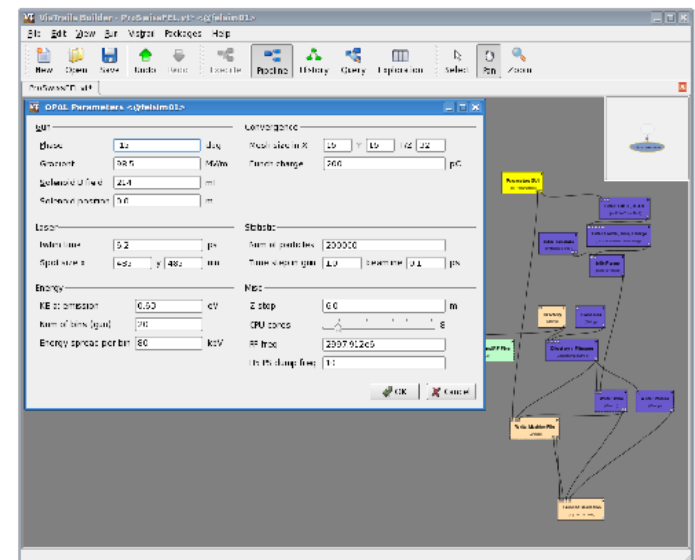
# Advantages of VisTrails for SwissFEL + OPAL

5) Takes care where data is stored

- Moving files of several GB → Expensive
- Better → Store data where it is generated

# State of Work on the Prototype

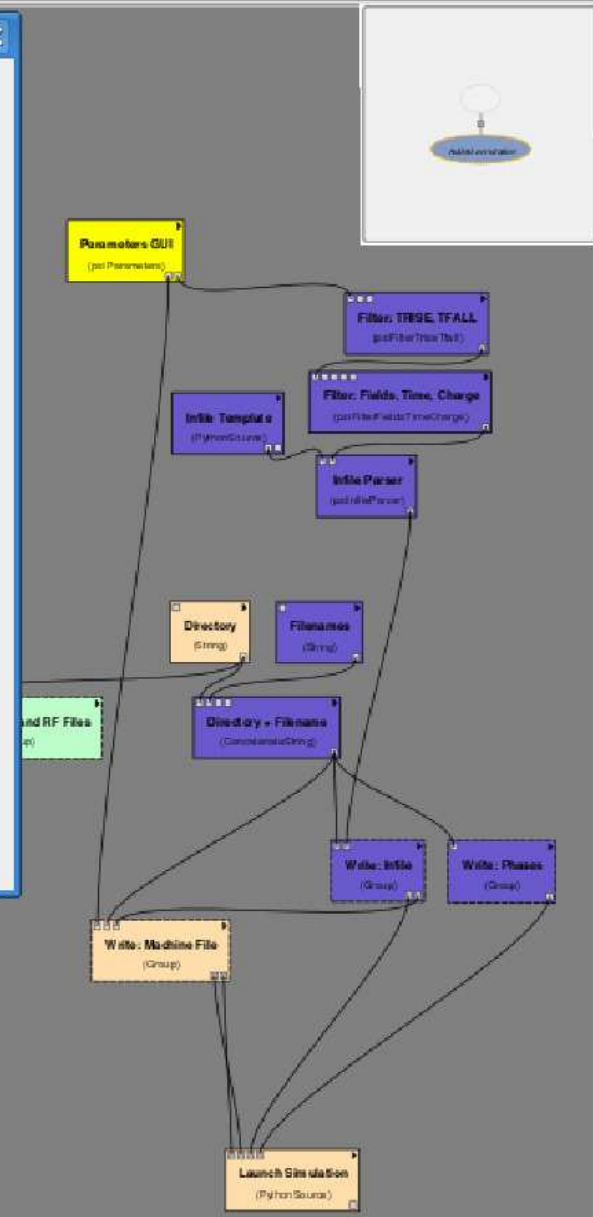
- VisTrails installed on FELSIM
- Workflow to start OPAL simulations
- GUI for easy parameter specification





VT OPAL Parameters <@felsim01>

<b>Gun</b>		<b>Convergence</b>	
Phase	<input type="text" value="-15"/> deg	Mesh size in X	<input type="text" value="16"/> Y <input type="text" value="16"/> T/Z <input type="text" value="32"/>
Gradient	<input type="text" value="98.5"/> MV/m	Bunch charge	<input type="text" value="200"/> pC
Solenoid B field	<input type="text" value="214"/> mT		
Solenoid position	<input type="text" value="0.0"/> m		
<b>Laser</b>		<b>Statistic</b>	
fwhm time	<input type="text" value="6.2"/> ps	Num of particles	<input type="text" value="200000"/>
Spot size x	<input type="text" value="485"/> y <input type="text" value="485"/> um	Time step in gun	<input type="text" value="1.0"/> beamline <input type="text" value="0.1"/> ps
<b>Energy</b>		<b>Misc</b>	
KE at emission	<input type="text" value="0.63"/> eV	Z stop	<input type="text" value="6.0"/> m
Num of bins (gun)	<input type="text" value="20"/>	CPU cores	<input type="text" value="8"/>
Energy spread per bin	<input type="text" value="80"/> keV	RF freq	<input type="text" value="2997.912e6"/>
		H5 P5 dump freq	<input type="text" value="10"/>



# Outlook



- Process and display simulation results
- Enable Provenance
- Enable collaboration -> Database