SCIENCE3D

pandata ODI

DESY

Helmholtz-Zentrum Geesthacht
Zentrum für Material- und Küstenforschung

Science3D - an open access project for 3D x-ray tomography

Frank Schlünzen
Deutsches Elektronen-Synchrotron (DESY)

HELMHOLTZ GEMEINSCHAFT
Open Science

(cc) BY

Outline

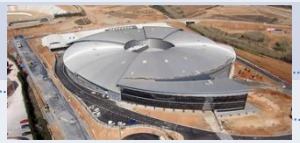
- Origin of the Science3D project
- Synchrotron radiation tomography
- Applications
- Services
- Hurdles

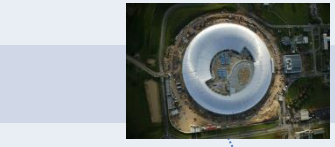
(cc) BY

PaNdata Facts

Consortium of 13 Photon and Neutron RIs

- **Focus** entirely on topics around data management
- **Number of instruments** > 200
- **Number of users >>** 30.000 users/yr
- **Investment** > €4.000.000.000*
- **Running costs** > €500.000.000/yr*
- **Publications** > 10.000/yr*
- **RCosts/Publication** ~ €50.000*%
- **Data volume** >> 10PB/yr*





% ESRF: operation budget 80-100Mio
# publications: ~1850

* wild guess

Data access

- Huge investment (tax money)
  - No access to data
  - Limited access to publication
  - Limited access to supplementary data
  - Useful meta-data only in the publication
  - Few tools for remote visualization, validation
- Quite limited return value (tax payer)

(cc) BY

PaNdata Virtual Labs

- PaNdata project tasks:
  - Demonstrator for a data catalogue integrating several services
- Drawback:
  - Projects don’t last (this one ended 09/2014)
  - Demonstrators hardly ever survive beyond the end of a project
    - Unless there is a real use case
  - Frequently wasted time & effort
- Alternative:
  - Turn the demonstrator into something useful
  - Create the use case: Open Access Tomography Database

(cc) BY

Aims

- Collect & make openly accessible
  - Raw data & meta-data
  - Derived data and any kind of supplement
  - Publications
  - Printable 3-D models
  - Tutorials & background information
- Provide services:
  - Visualization
  - Tools
  - Long term storage
- Why tomography
  - Applied science, comparably easy to capture the outcomes
  - Useful in a wide spectrum of applications; widely known technique

MARCO REUS

Tränen nach dem WM-Aus

Nach dem WM-Aus rief er gleich seine Eltern an

BILD MENÜ

Mainz.

Mit dabei: Patrick Ingelfinger, Oberarzt an der Mainzer Uni-Klinik und Mannschaftsarzt von Mainz 05. Er hat die Aufnahmen von MRT dabei. Gemeinsam mit DFB-Arzt Dr. Hans-Wilhelm Müller-Wohlfahrt werden die Aufnahmen begutachtet.

Bis um 2 Uhr morgens feststeht: Reus ist raus! Er wird bei der WM in Brasilien fehlen. Die Verletzung ist zu schwerwiegend.

Marco Reus verlässt auf Krücken das Knappschaftskrankenhaus in Dortmund

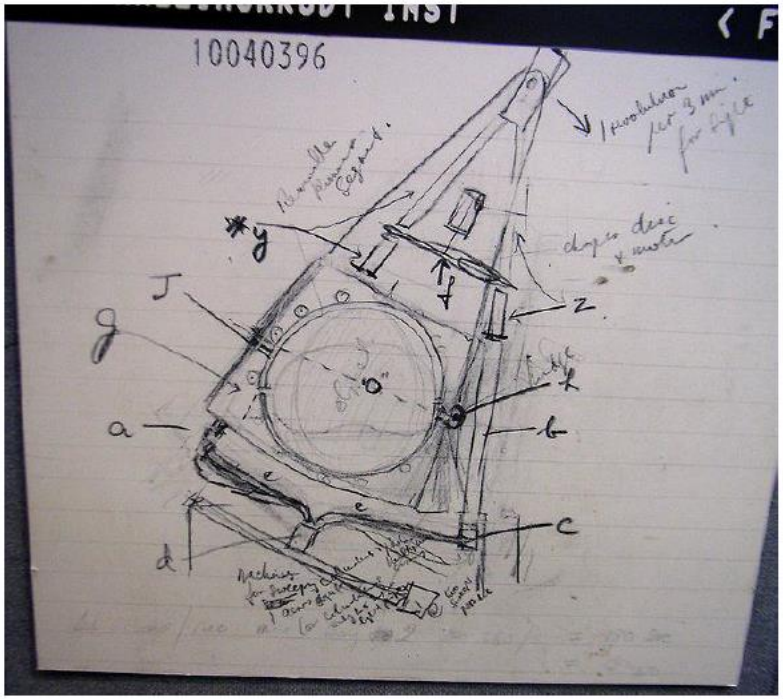
(cc) BY

Tomography

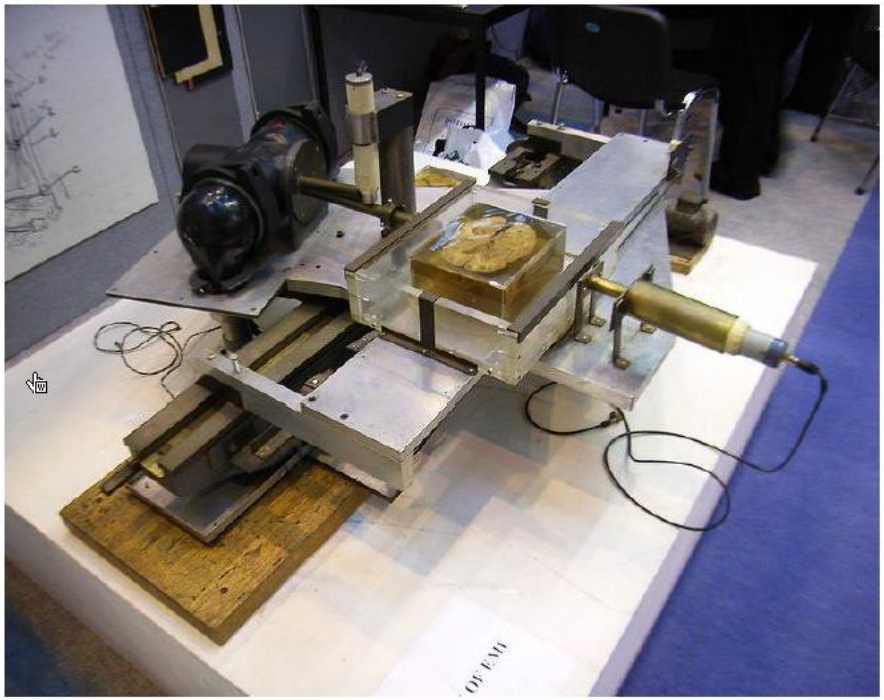
History: 2D ⇒ 3D

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

- Medicine CT was developed in the late 1970s based on the idea of the British engineer Godfrey Hounsfield and the American physicist Allan McLeod Cormack
- Hounsfield and Cormack shared the 1979 Nobel Prize in Medicine



Original sketch from Housfields notebook



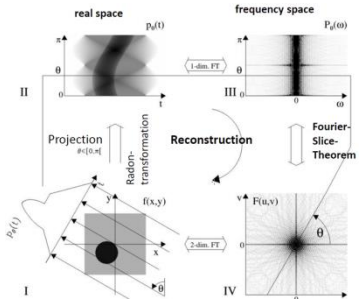
The very first CT-scanner prototype invented by Houndsfield at EMI

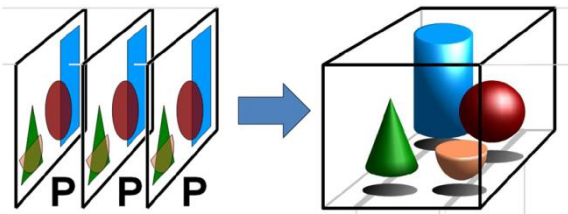
Stolen from: Astrid Haibel talk at science3d workshop

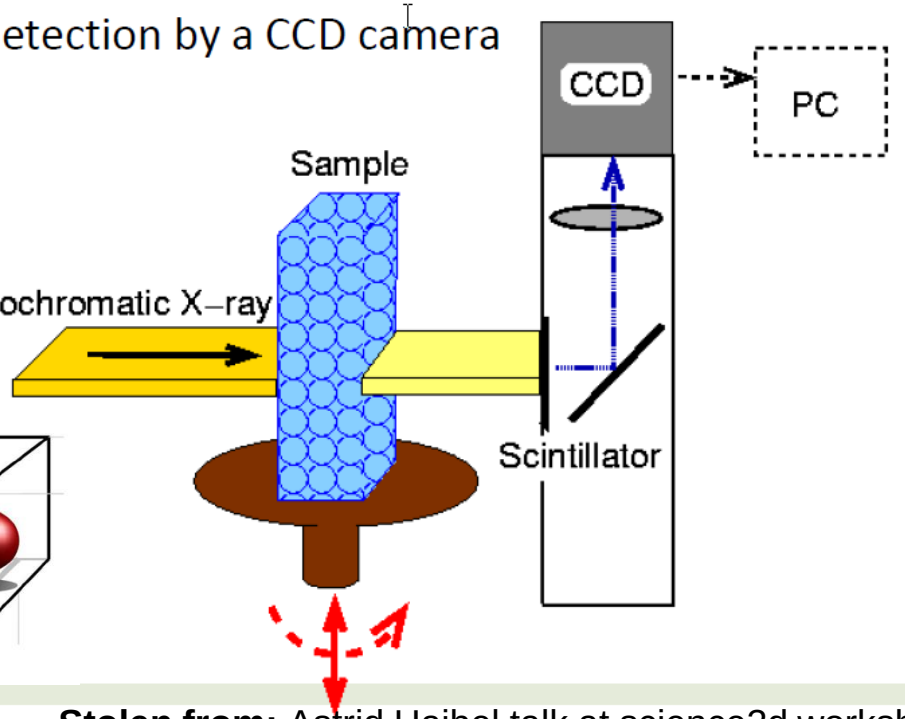
Synchrotron CT
- Measurement Principle -

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

- **highly intensive**, (quasi-) **parallel**, **monochromatic**, **partially coherent beam** penetrates the sample
- local attenuation and/or phase shift depends on materials density, thickness and inner structures
- measuring of many different angular positions (0°-180°, 900 -2400 images)
- conversion of the X-ray light into visible light by a scintillator screen
- magnification by a microscope optics and detection by a CCD camera
- reconstruction of the 3D image







Stolen from: Astrid Haibel talk at science3d workshop

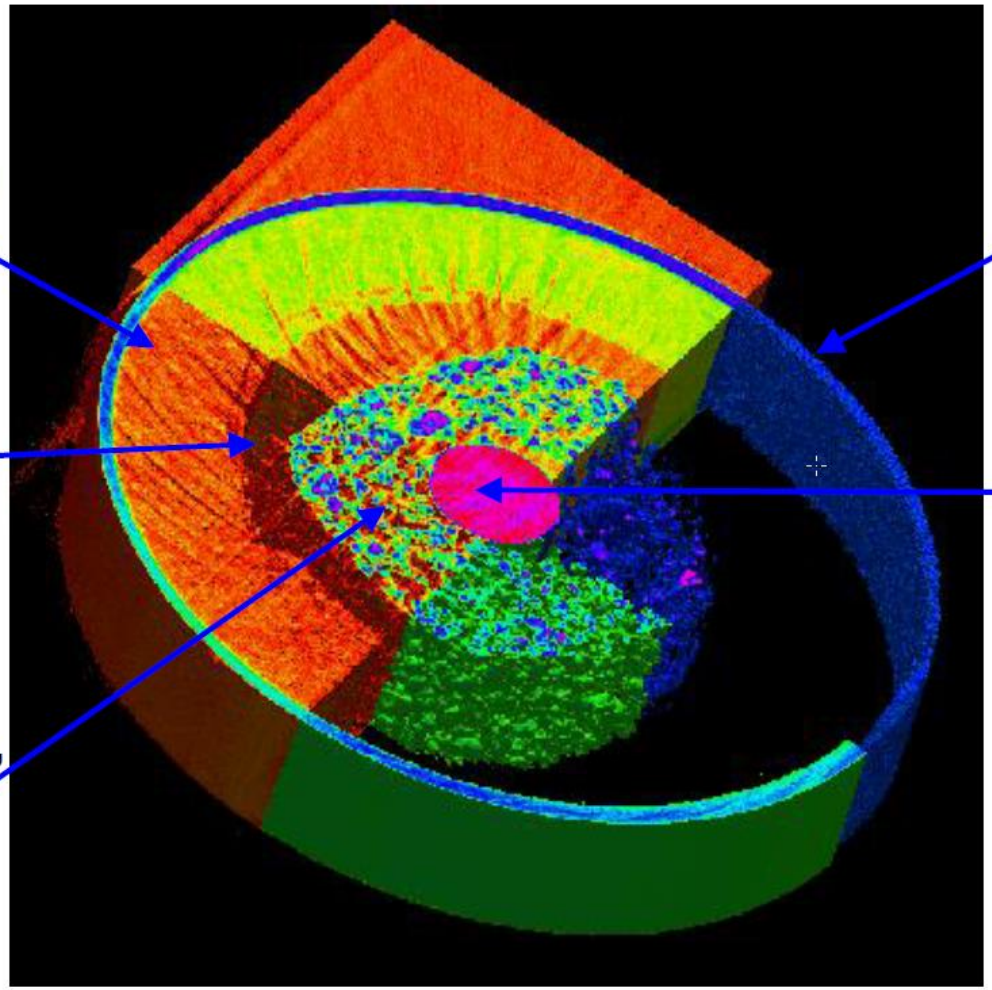
SCIENCE3D

Applications

Application: In situ tomography at alkaline battery LR61

Investigation of zinc particle dissolution during discharging

BESSY



manganese oxide (cathode)

steel tube

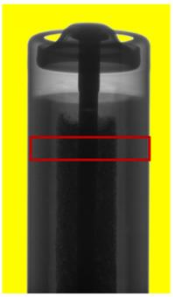
separation fleece

arrester

zinc gel (67% zinc, 33% electrolyte) (anode)

alkaline battery

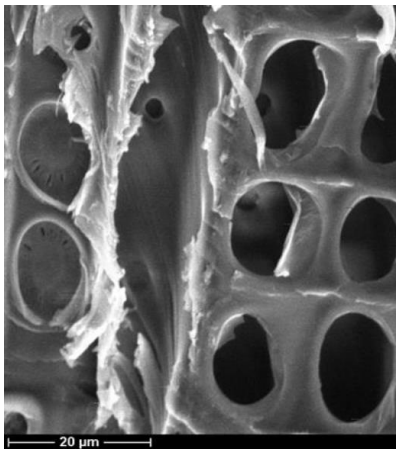
radiographic image

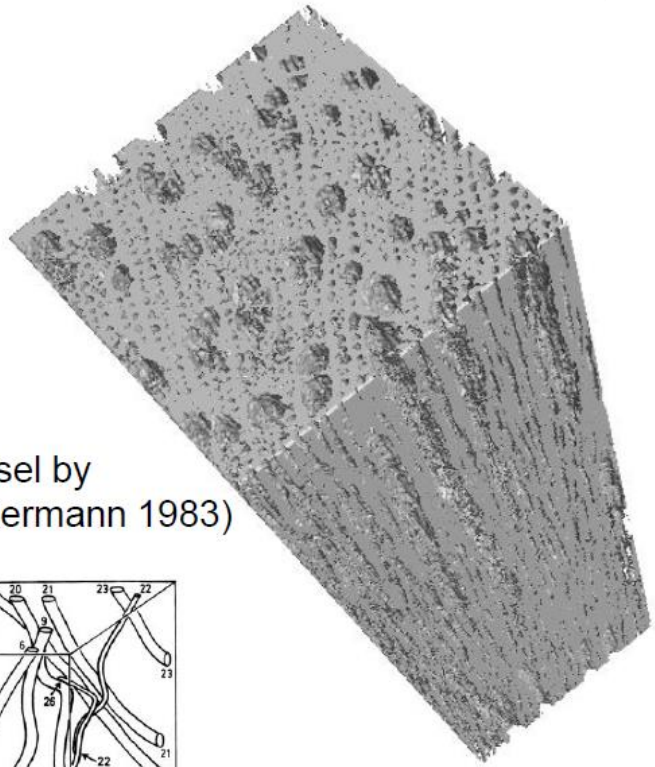


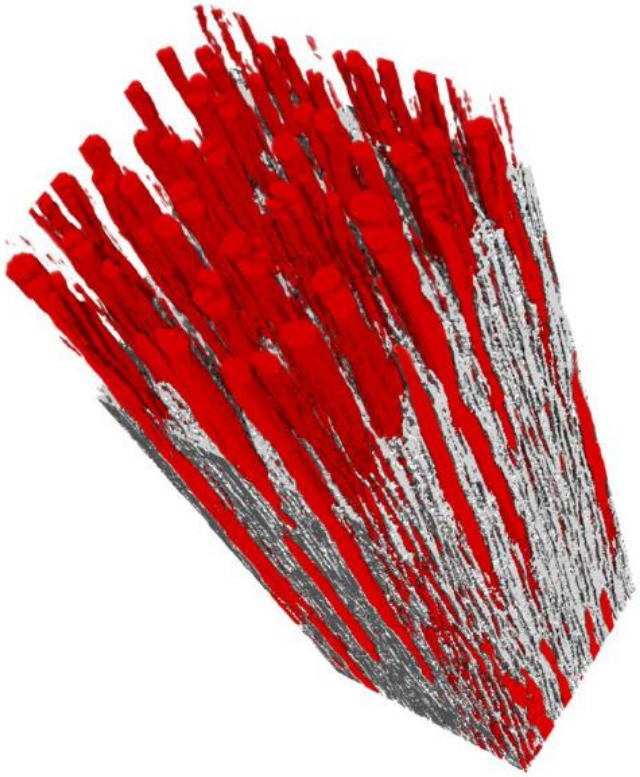
Stolen from: Astrid Haibel talk at science3d workshop

poplar wood SRμCT imaging

HNE Eberswalde
Hochschule für nachhaltige Entwicklung (FH)
Eberswalde University for Sustainable Development
University of Applied Sciences







Reconstruction of course of vessel by light microscopy analysis (Zimmermann 1983)

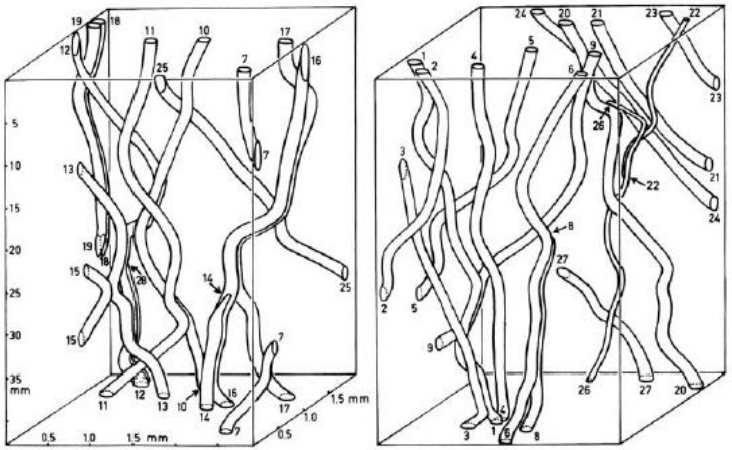
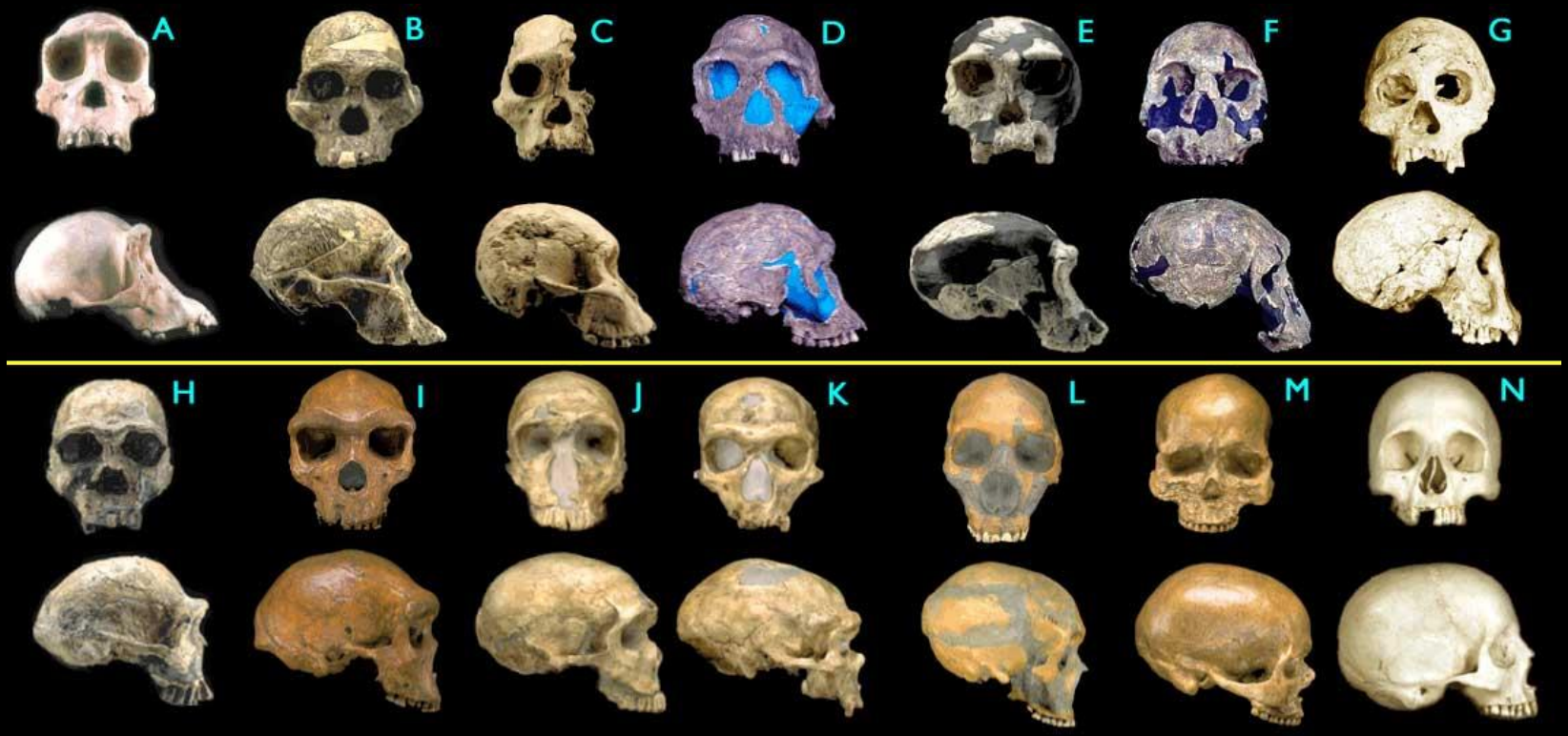


Fig. 2.2. Course of vessels in a piece of wood of *Cedrela fissilis*. Individual vessels have been arbitrarily separated into two blocks for clarity. Vessels are numbered where they exit from the block. *Arrows* show vessel ends. Note that the axial scale is foreshortened ten times. (Zimmermann and Brown 1971)

3D reconstruction of the vessel system

Stolen from: Silke Lautner talk at science3d workshop

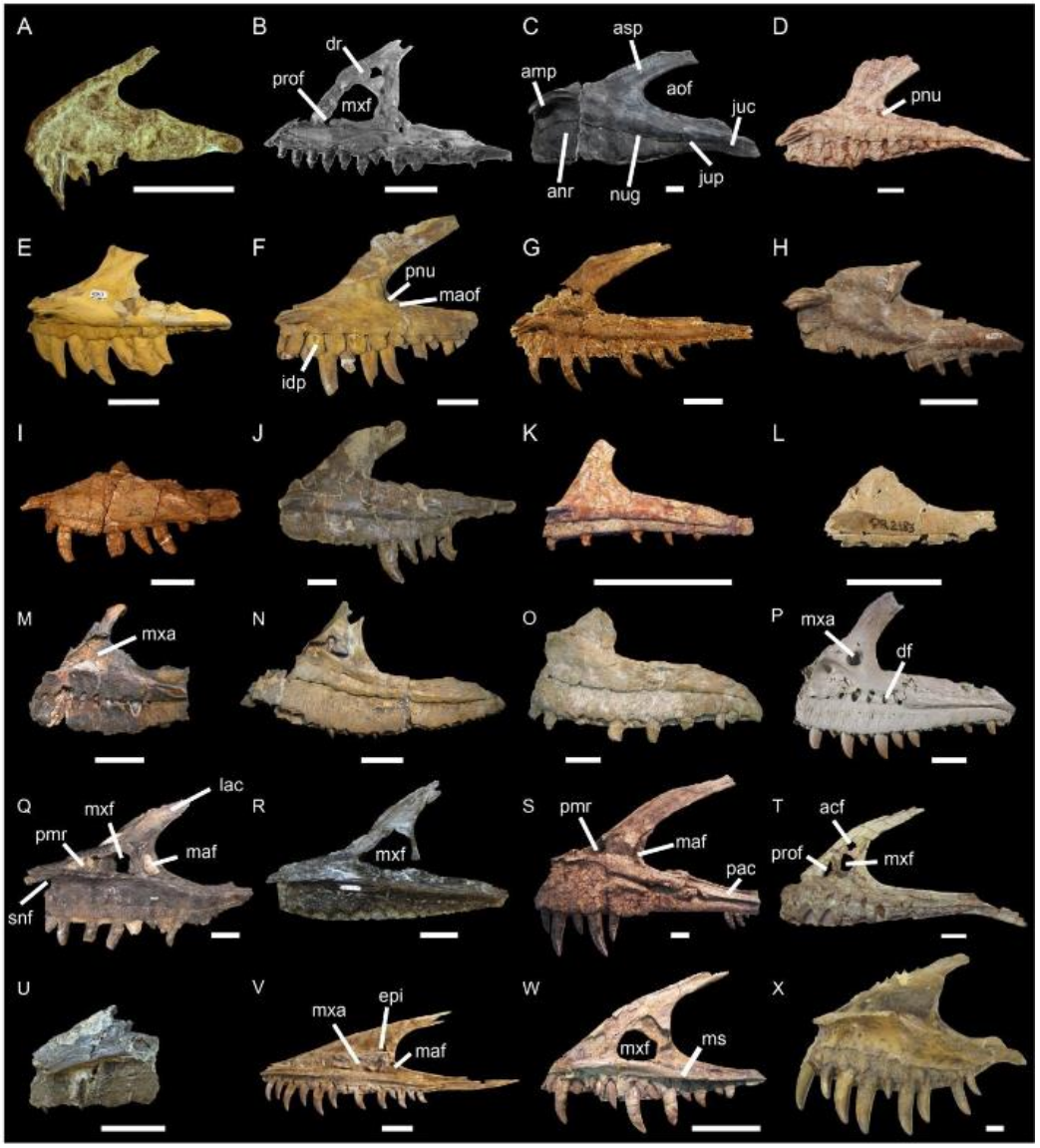
Evolution



From chimpanzee over *Australopithecus africanus* to *Homo sapiens*

SCIENCE3D

Embryos of dinosaurs



Note 3 – Maxillae of nonavian theropods in medial view. **A.** Embryo of *Torvosaurus* sp. (ML 1188, reconstructed). **B.** Hatchling of *Allosaurus* (IPFUB Gui Th 4; Rauhut &

http://www.nature.com/srep/2013/130530/srep01924/full/srep01924.html

http://felipe-elias-portfolio.blogspot.com

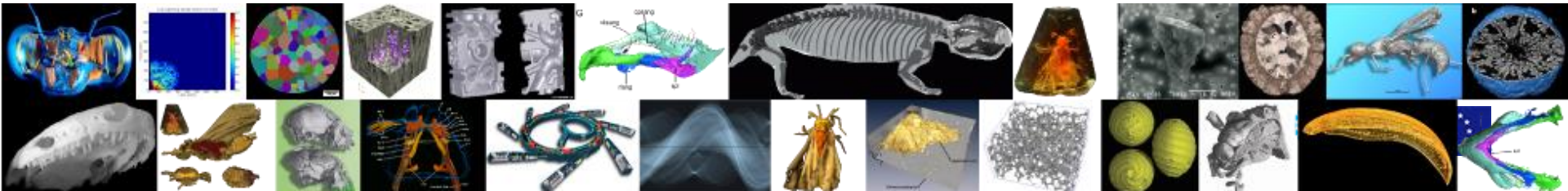
Very applied science

- Not too far from topics typically taught in universities & schools
  - Comparative & evolutionary biology
  - 3-D models to look at on a sub-µm-scale – if you have the right tools
  - 3-D models to take to the classroom
    - 3-D printing the next revolution?

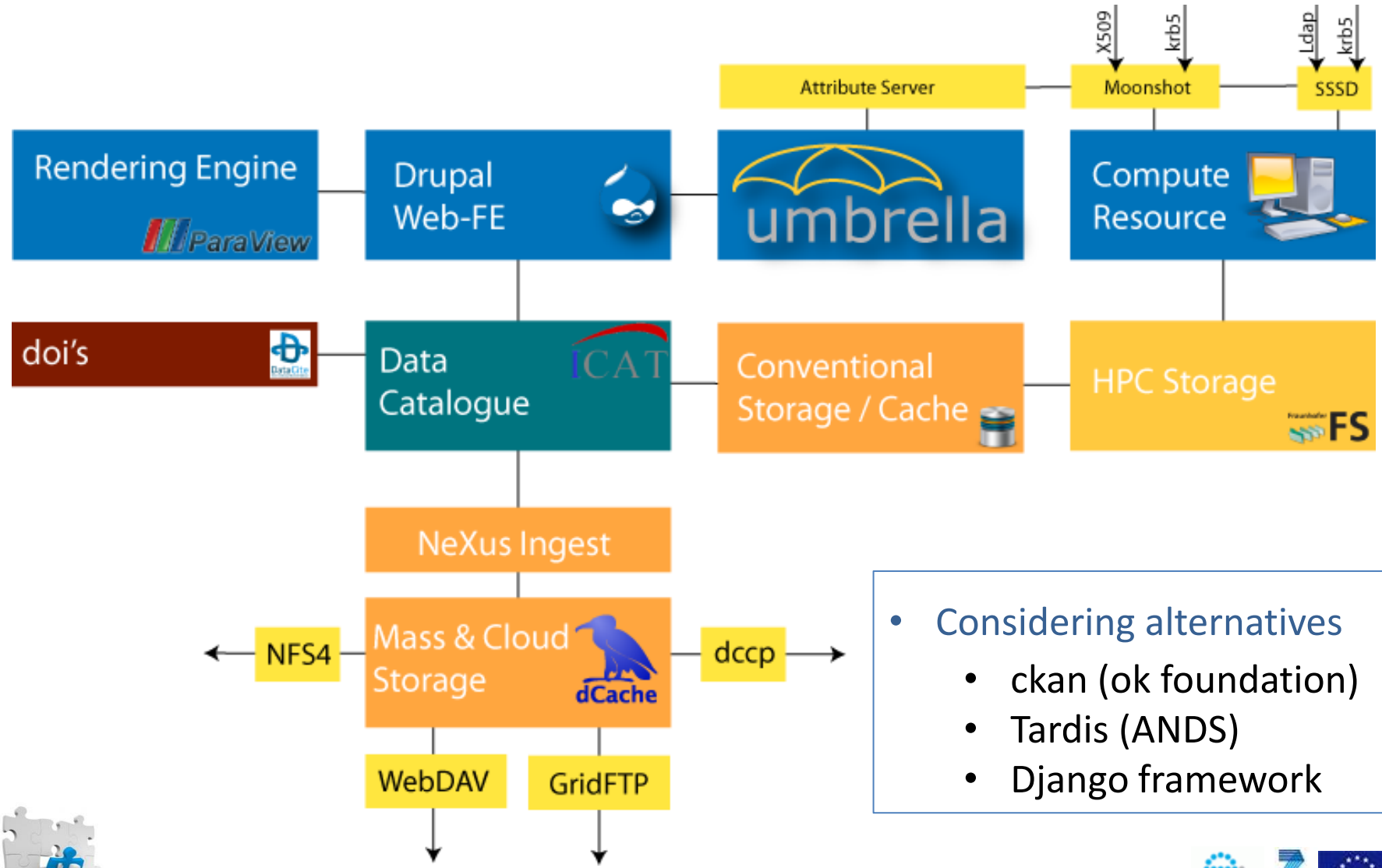
- Science3D workshop in June ‘14
  - https://indico.desy.de/event/science3d
  - Several teachers / educators
  - Presentations at a rather scholarly level
  - Have to see how it develops…

SCIENCE3D

Services

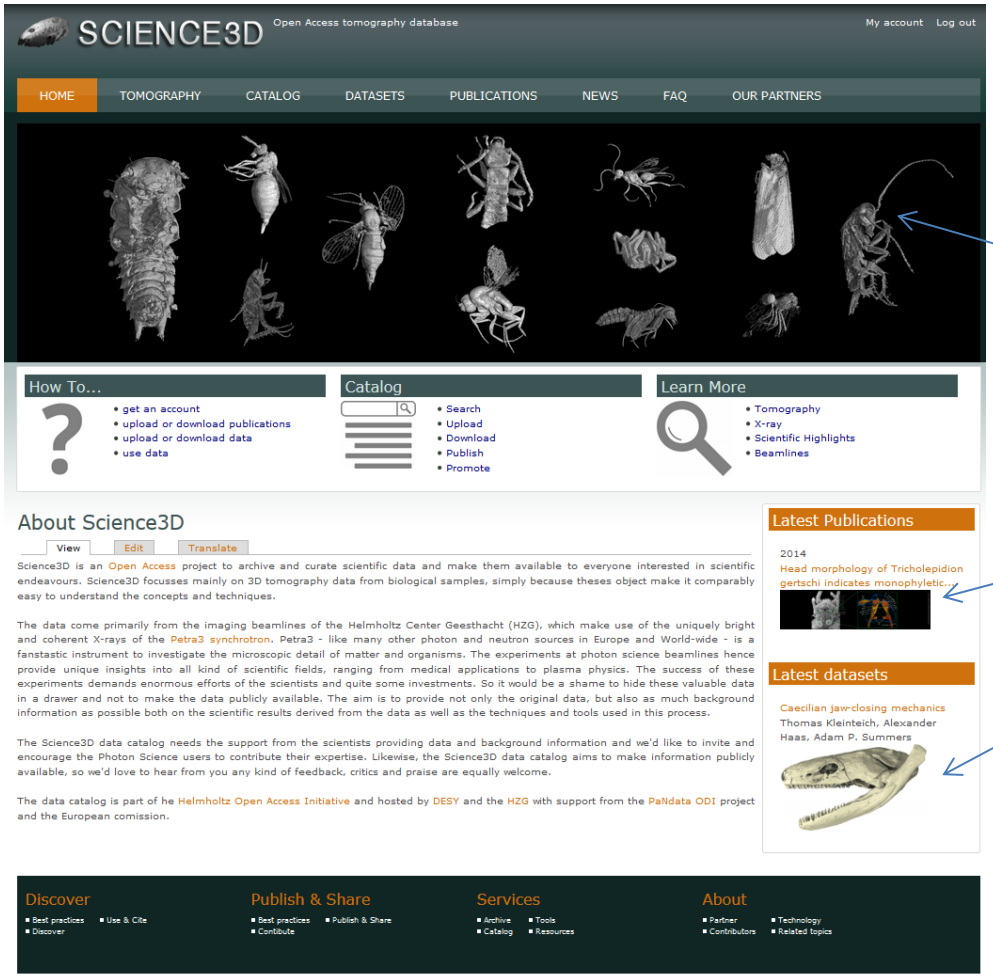


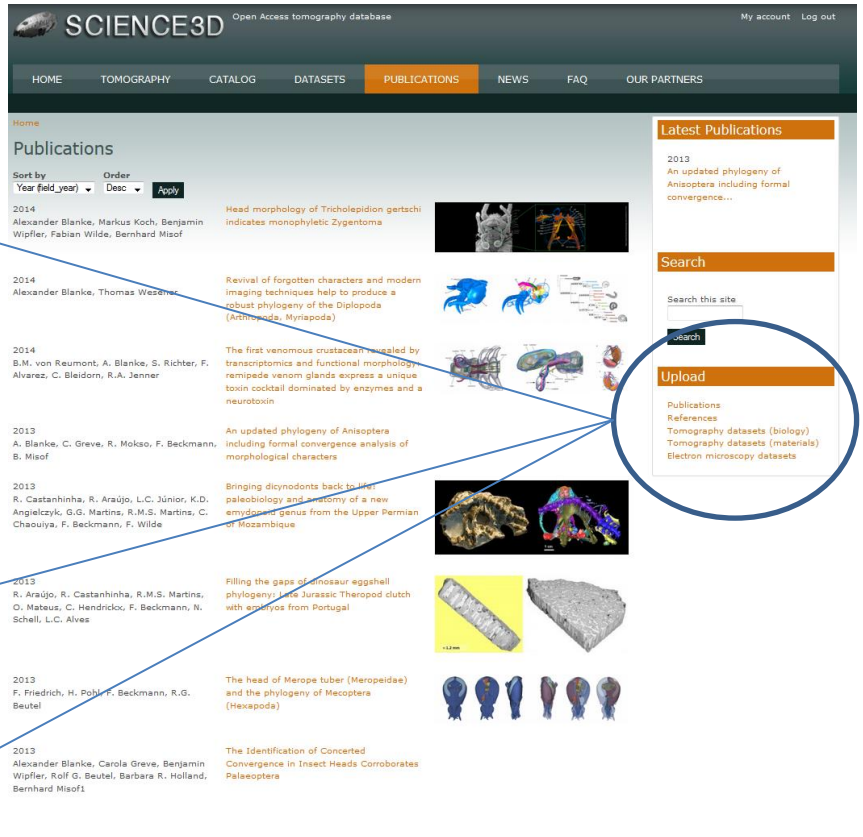
infrastructure



- Considering alternatives
  - ckan (ok foundation)
  - Tardis (ANDS)
  - Django framework

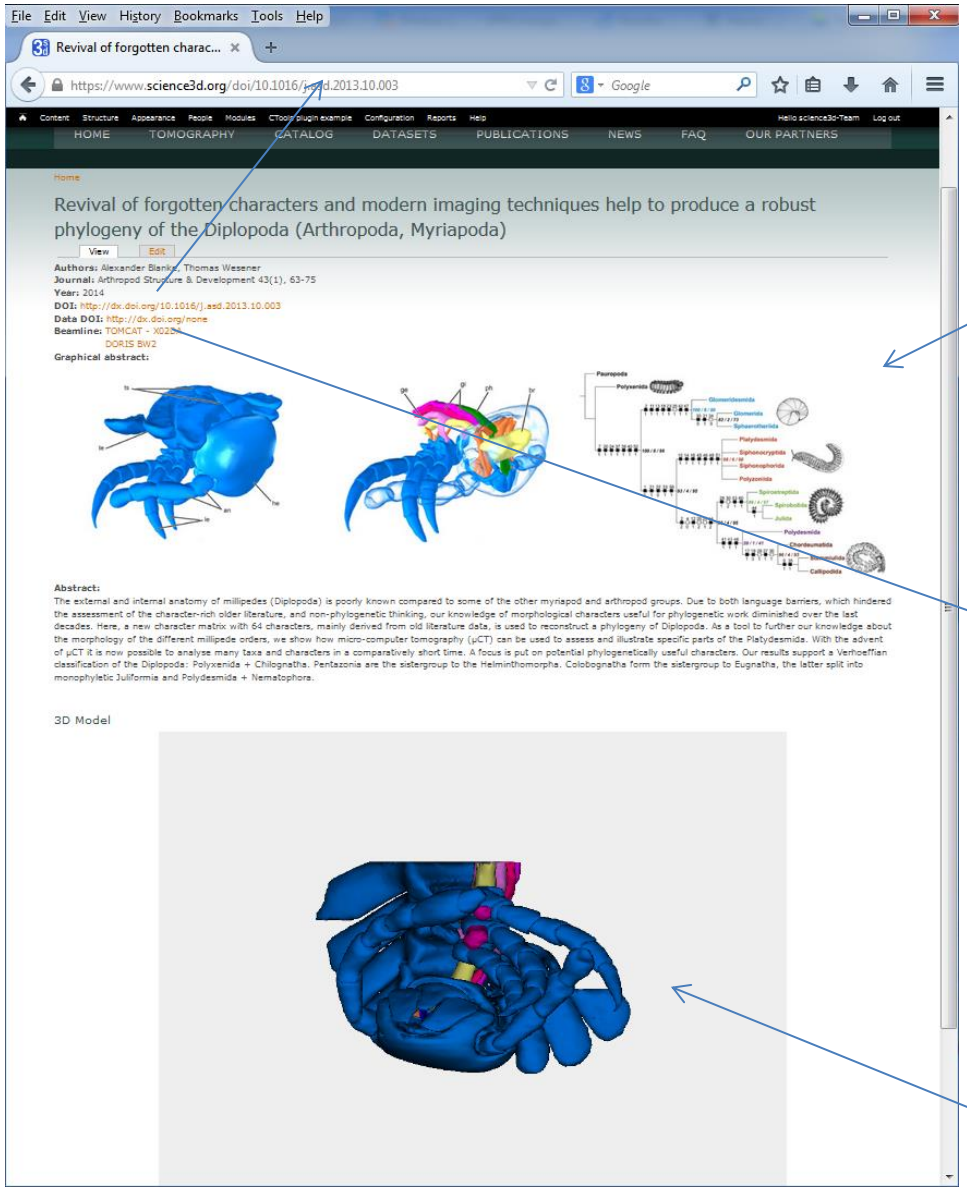
Web: science3d.org

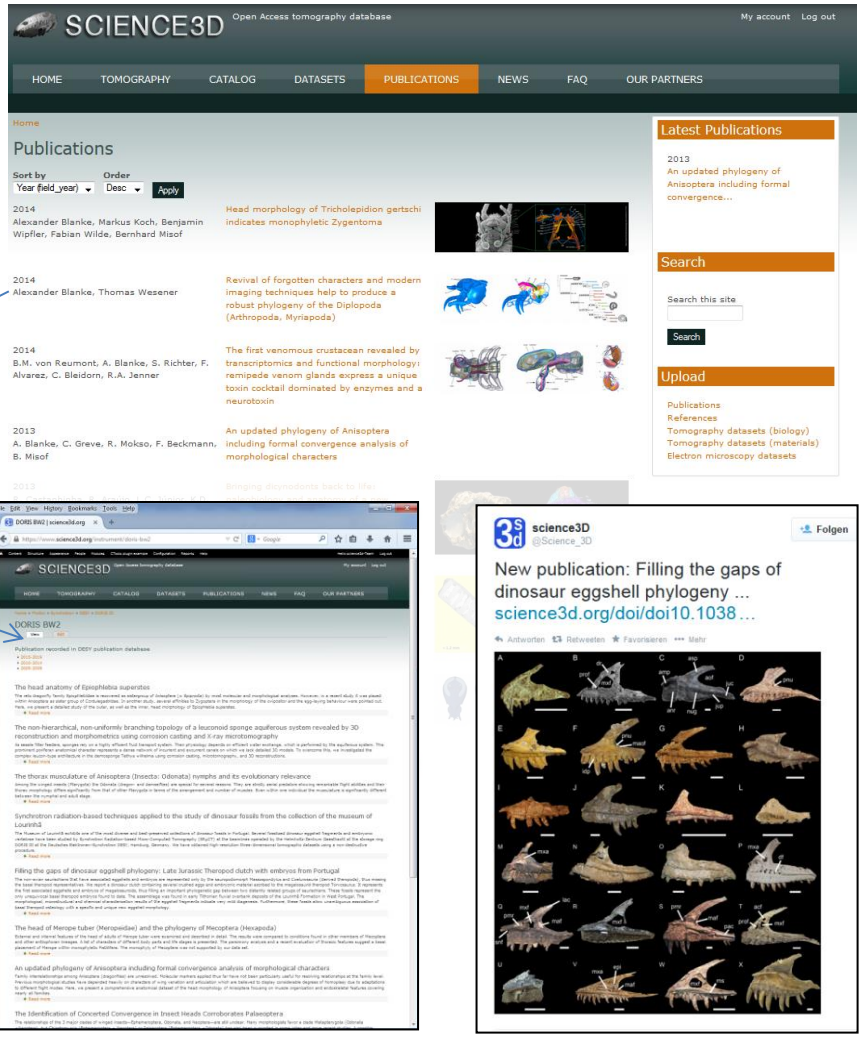




- Simple mechanisms
  - Uploading any data
  - Highlighting input

web





- 3-D models
  - Still primitive
  - Lacking 3-D pdf capabilities

Resources for scientists

- Data
  - Data upload/download
  - Long term mass storage dCache (unlimited)
  - Conventional storage for the catalog (secure, fast)
  - Fast cluster storage (for HPC wherever needed)
  - Register data with datacite
  - Convert / Provide data in maintainable formats

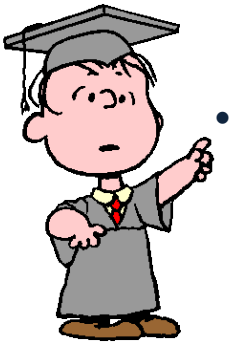
- Compute infrastructure
  - Work group server without access restrictions
  - Batch farm (~8000 cores); GRID farm (many more cores)
  - HPC cluster (~3000 cores) with fast interconnect (IB)
  - GPGPUS (nvidia M2050 → K20X) with slow/fast interconnect
  - Avizo, Matlab, IDL, Mathematica, Maple, etc
  - Fast remote graphical access

- Other Services
  - feeds/tweets/newsletters
  - 3D print (tbc)

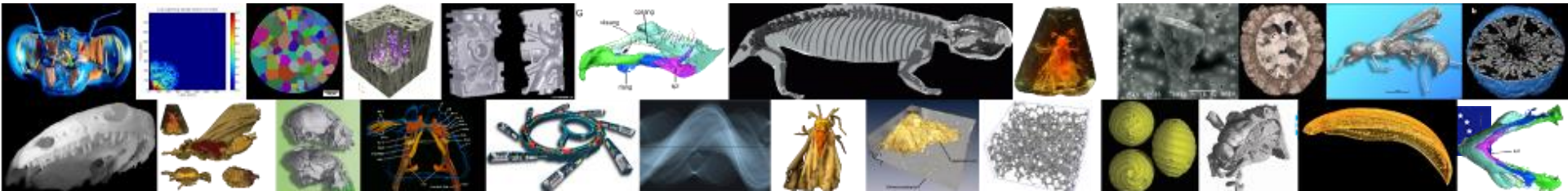
SCIENCE3D

for students & teacher

- Tomography offers tremendous insights into evolution, anatomy, materials…
  - 3D models on the screen; 3D prints for the classroom
  - Still very difficult to harvest for pupils
  - Might well be completely unrealistic
- Science 3D might offer a platform
  - for scientists and teacher/pupils/students to make contacts and possibly bring science more into classrooms
- High hurdles
  - Physics, Mathematics, Biology, Language barriers
  - Even a challenge for teachers
- Considering making (very limited) experiment time available (Schülerlabor)



SCIENCE3D



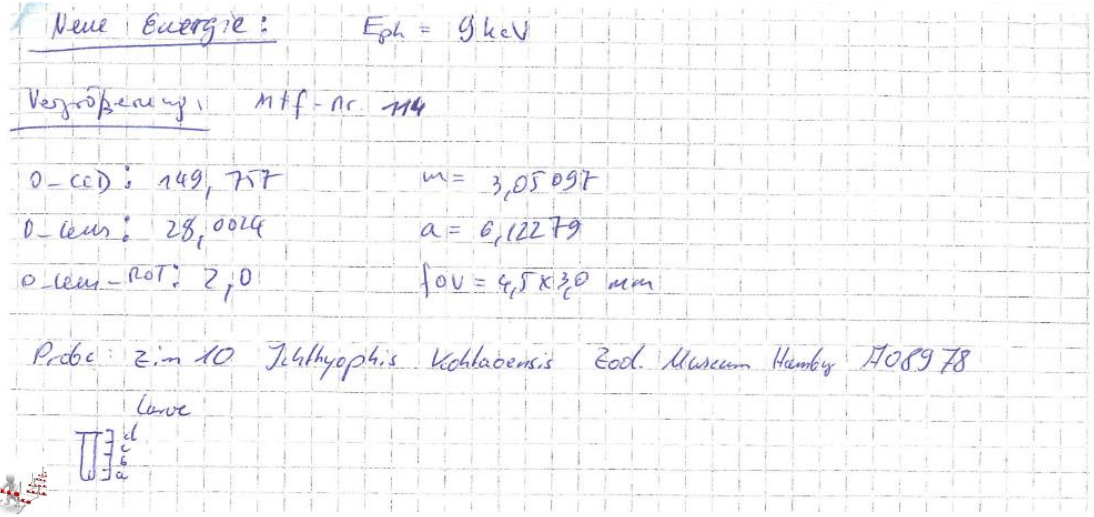
The usual problems

- Hardware
  - No problem at all
- Chicken & egg
  - But quite a bit of support from scientists
- Manpower & Time
  - Development of services (3D, workflows, platform)
    - Currently a very primitive setup
  - Data standardization & Content (background material)
  - HUGE problem – intend to apply for a H2020 grant
- Knowledge
  - We are lacking the information which data precisely where used
    - Sample preparation
    - Experiment parameters
  - We can't make data public; should be responsibility of the owner
    - Though ownership is an open debate
  - Name the sample correctly
  - Get taxonomies right

SCIENCE3D

The usual problems

- Data formats & meta-data
  - NeXus/HDF5 as a standard
    - Evolving & insufficient
    - Software support limited
  - No data in standard format
    - Complex meta-data
    - Spread over various sources
  - No automatic workflow
    - Manual process



3.4.19. NXtomo

Status:

application definition, extends *NXobject*, version 2.0

Description:

This is the application definition for x-ray or neutron tomography raw data. In t images and, of course the sample. In order to distinguish between them images

Symbols:

These symbols will be used below to coordinate datasets with the same shape.

nFrames: number of frames

xsize: number of pixels in X direction

ysize: number of pixels in Y direction

Groups cited:

NXdata, *NXdetector*, *NXentry*, *NXinstrument*, *NXmonitor*, *NXsample*, *NXsource*

Structure:

entry: *NXentry*

title: *NX_CHAR*

start_time: *NX_DATE_TIME*

end_time: *NX_DATE_TIME*

definition: *NX_CHAR*

Official NeXus NXDL schema to which this file conforms

Obligatory value: NXtomo

instrument: *NXinstrument*

(source): *NXsource*

type: *NX_CHAR*

name: *NX_CHAR*

probe: *NX_CHAR*

Any of these values: neutron | x-ray | electron

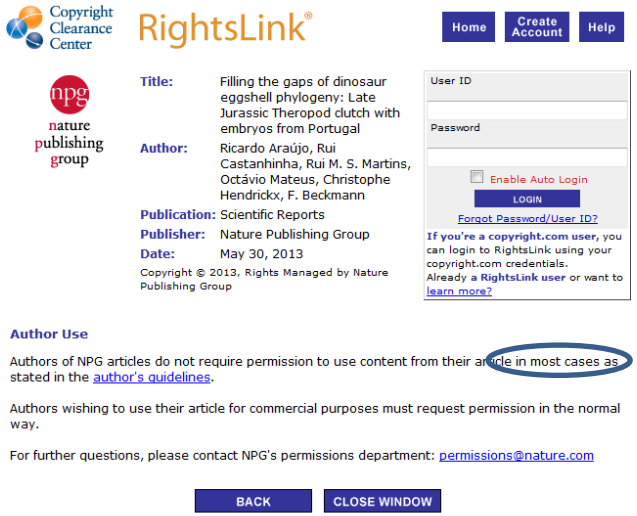
detector: *NXdetector*

data[nFrames, xsize, ysize]: *NX_INT*

image_key[nFrames]: *NX_INT*

The usual problems

- Publications & materials
  - Copyright situation almost always unclear
  - Providing resources on behalf of the author
  - Providing publications via the DESY-library services?
  - Re-use of images?
  - Re-use of supplemental materials?



SCIENCE3D

Contact: science3d@desy.de

Thank you for your attention !