GlobusWORLD Conference
April 21, 2016

Paving the Road from Instruments to the RCC's Midway HPC Cluster: **XDM**, the XROMM Data Management Platform

H. Birali Runesha

runesha@uchicago.edu

THE UNIVERSITY OF CHICAGO

Research Computing Center

RCC

A unit in the Office of the Vice President for Research and National Laboratories

Serving a wide array of research disciplines with various computation and data management needs

Compute-Intensive

Analytics-Intensive

Data-Intensive

Collaborative-Intensive

THE UNIVERSITY OF CHICAGO | Research Computing Center

"Provide access to hardware (computing, storage, and visualization resources), software and advanced technical support"

Hardware

- **High End Computing**
  Tightly and Loosely coupled nodes
- **Research Storage**
  Data Management and Backup
- **Special Hardware**
  GPU, PHI, Shared memory, Network

Software

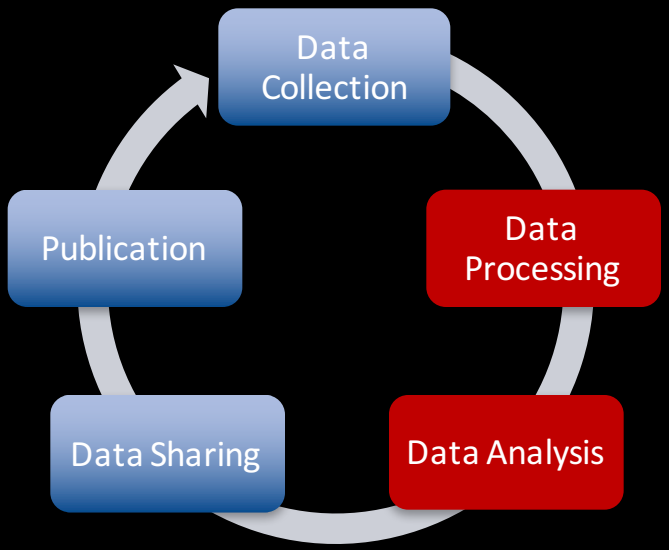
- Commercial software Licensing
- Public/Community codes
- Homegrown codes

Technical Support

- Consulting
  Data Visualization
- Education and Training
  Grants
- Application Software Development

RCC

The deluge of research data is a reality at RCC and a challenge for researchers



- Instrumentation creates a flood of data
- Data storage and compute resources are highly distributed
- Storing of data on inadequate media with no consistent back up mechanism
- Volume, Velocity and Veracity of data
- Metadata and data preservation
- HIPAA, Sensitive data
- Lack of standards
- Campus cyberinfrastructure-Network
- Etc.

We will talk about data from experiments and observations

Xenon1T dark matter detector in the Gran Sasso Underground Laboratory in Italy. Data will be stored and analyzed at RCC

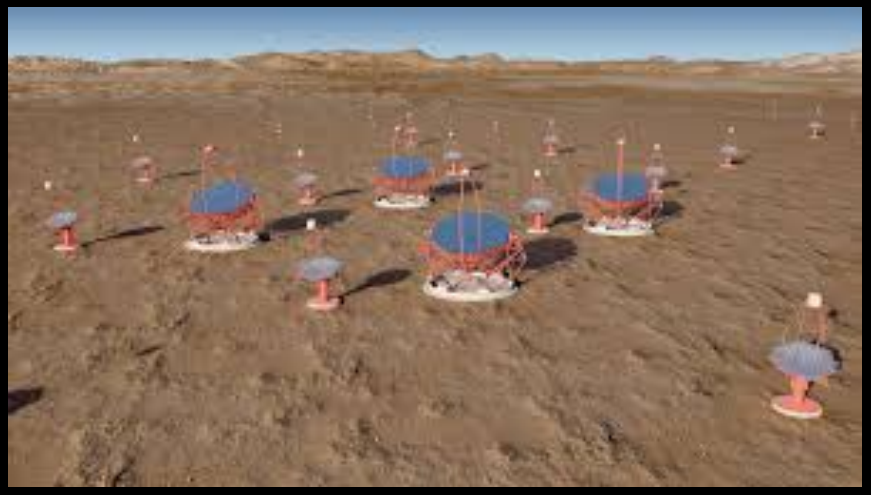
A new, 3.3 ton Xenon detector (XENON1T), 100 times more sensitive than the current leading experiment to focus on searching for dark matter.

During calibration it will produce up to 30TB of raw data per day and will need ~2 PB per year

The South Pole Telescope (SPT)

The third generation detector SPT-3G will be installed in 2016 and will be 20 times more sensitive than the previous detector. 600TB of data per year.

The Cherenkov Telescope Array (CTA) will improve on the sensitivity of the current generation of telescopes by a factor of 10 and will produce as much as 200TB of data per year



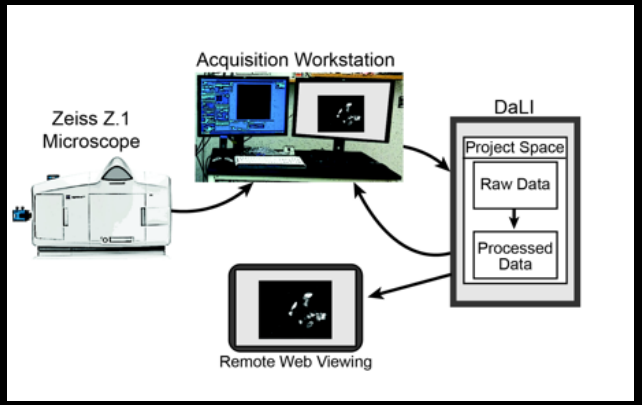
Biodiversity in the deciduous forests of Eastern Asia and Eastern North America

These forests were anciently connected and have a shared evolutionary and ecological history offering a unique opportunity to study the drivers of biodiversity across geographical space and through evolutionary time. Current data volumes will increase significantly over the next three years to >100TB

Microscopy

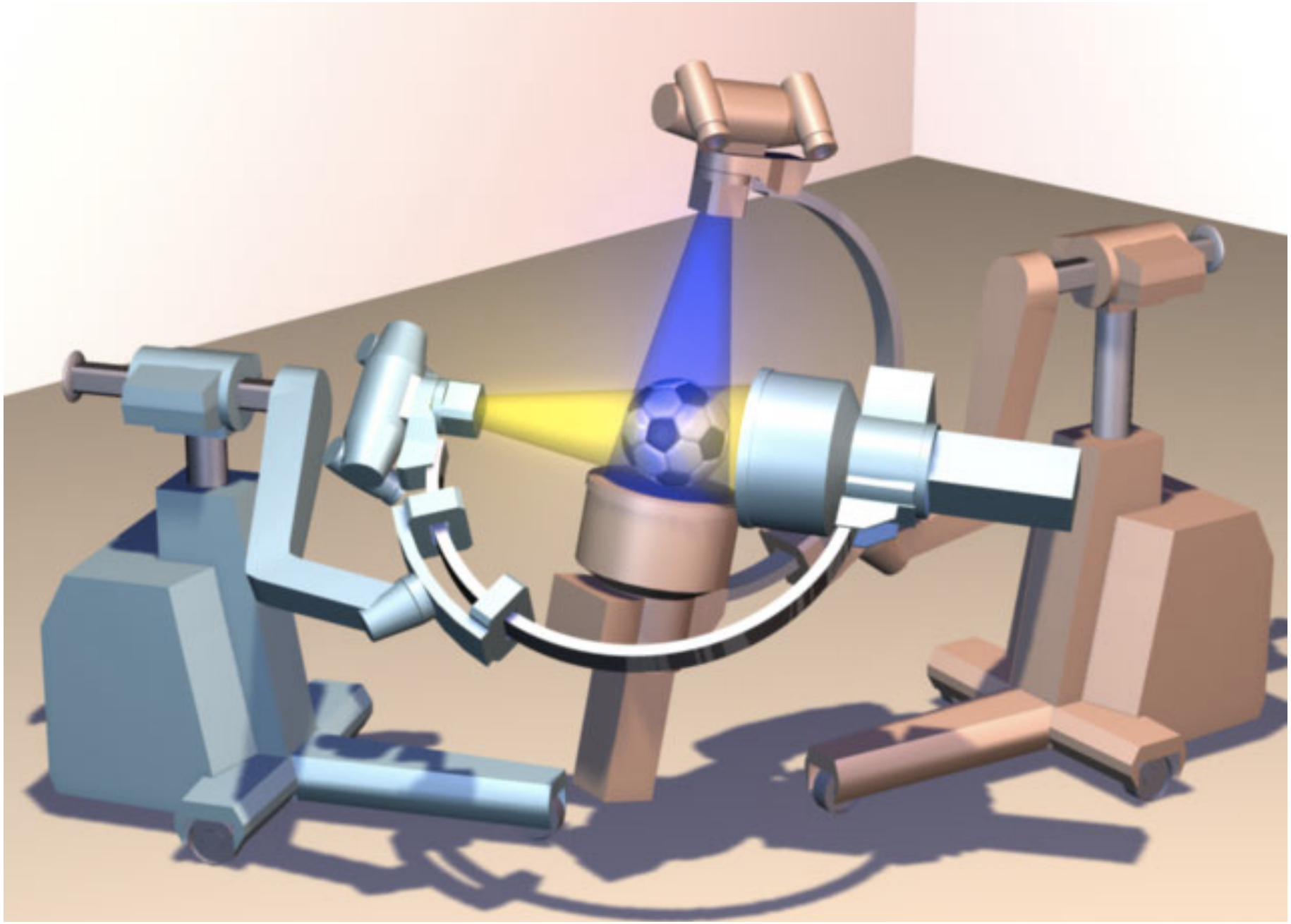
Advances in scanned imaging modalities such as confocal and light sheet microscopy are constantly increasing the speed, resolution, and sophistication of image measurements

software for the real time acquisition, processing, and analysis of images acquired during light sheet microscopy (LSM)

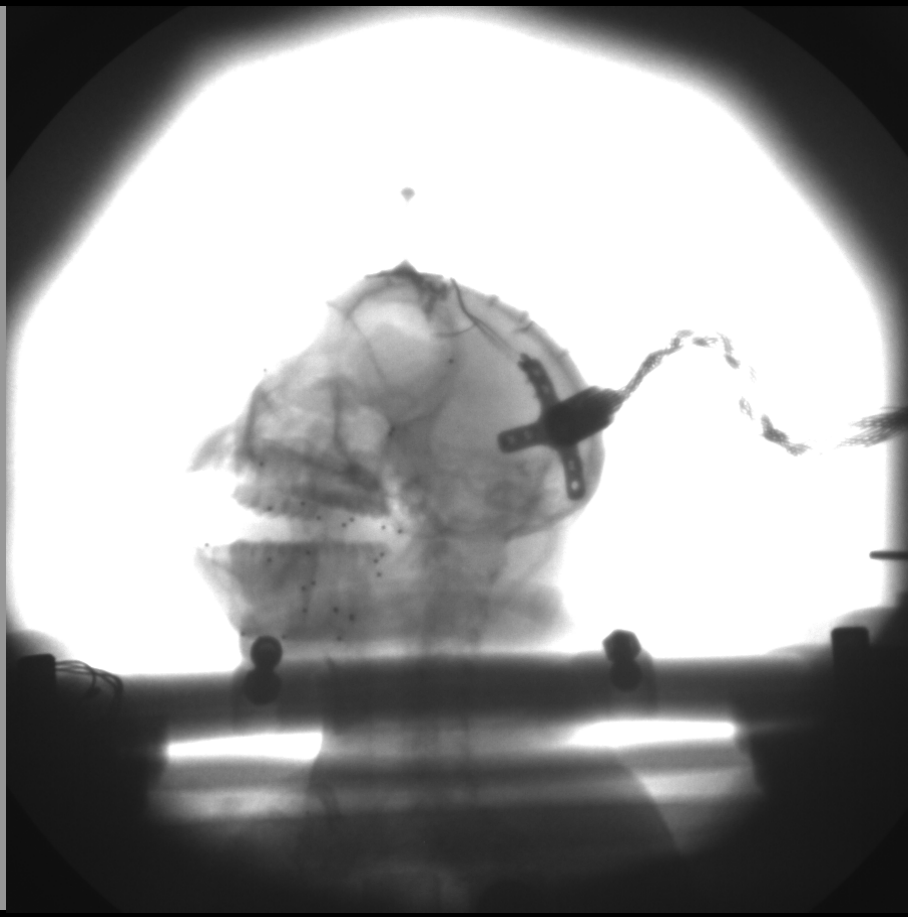


The X-ray Reconstruction of Moving Morphology (XROMM) Lab

Biplanar Digital Videofluoroscopy Instrument for X-ray Reconstruction of Moving Morphology (XROMM)



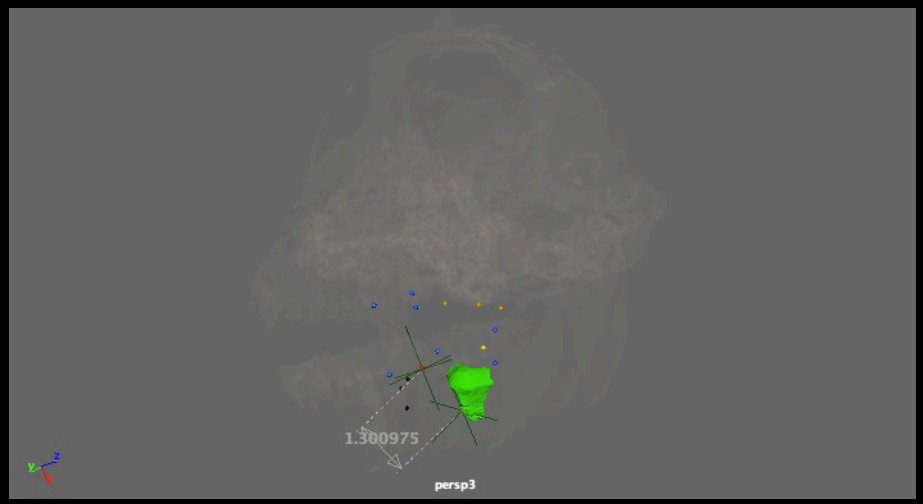
The XROMM Lab

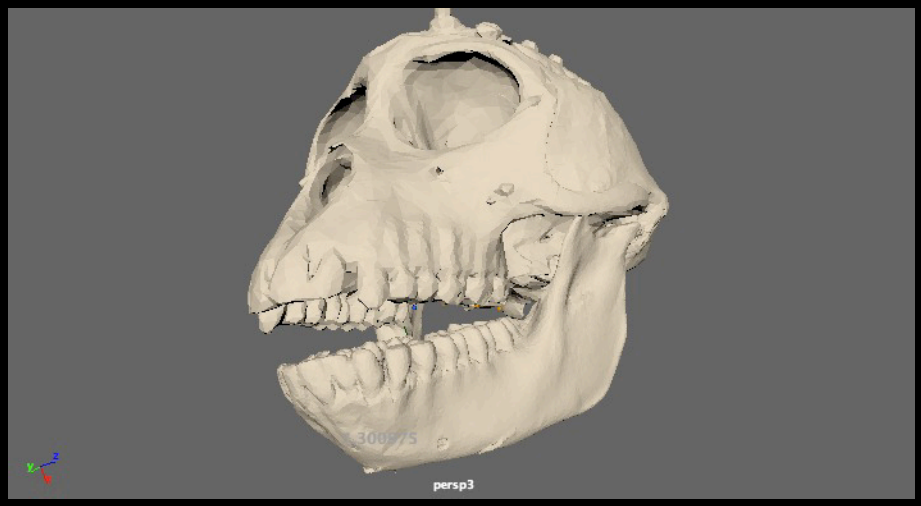


Prof. Callum Ross et al.

The XROMM Lab

Study of the organization, function, and evolution of vertebrate feeding systems using *in vivo* measurement of 3D jaw kinematics, muscle activity, bone strain, and cortical neuron activity, combined with computational methods for modeling bone deformation, muscle architecture dynamics, and motor control.

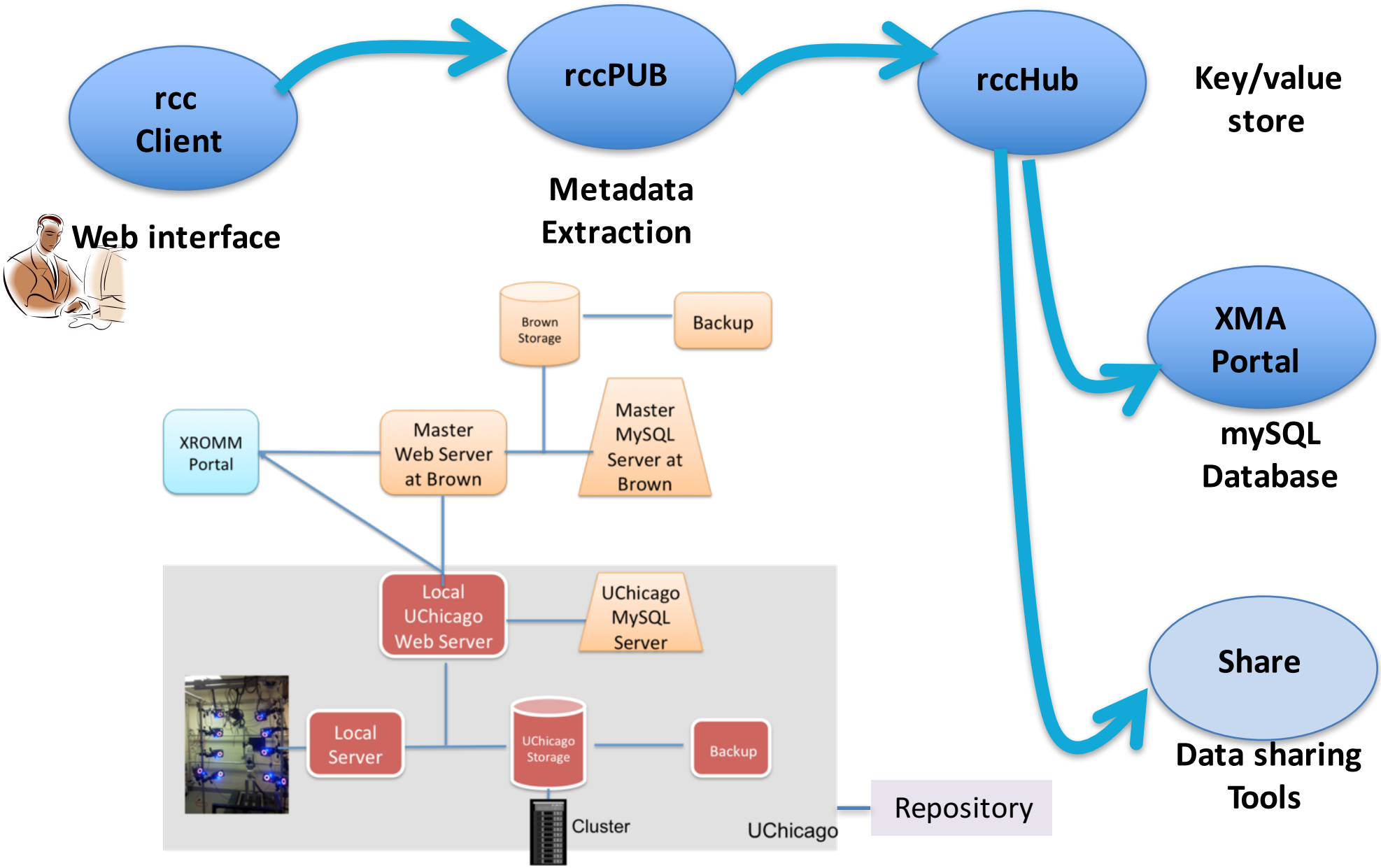




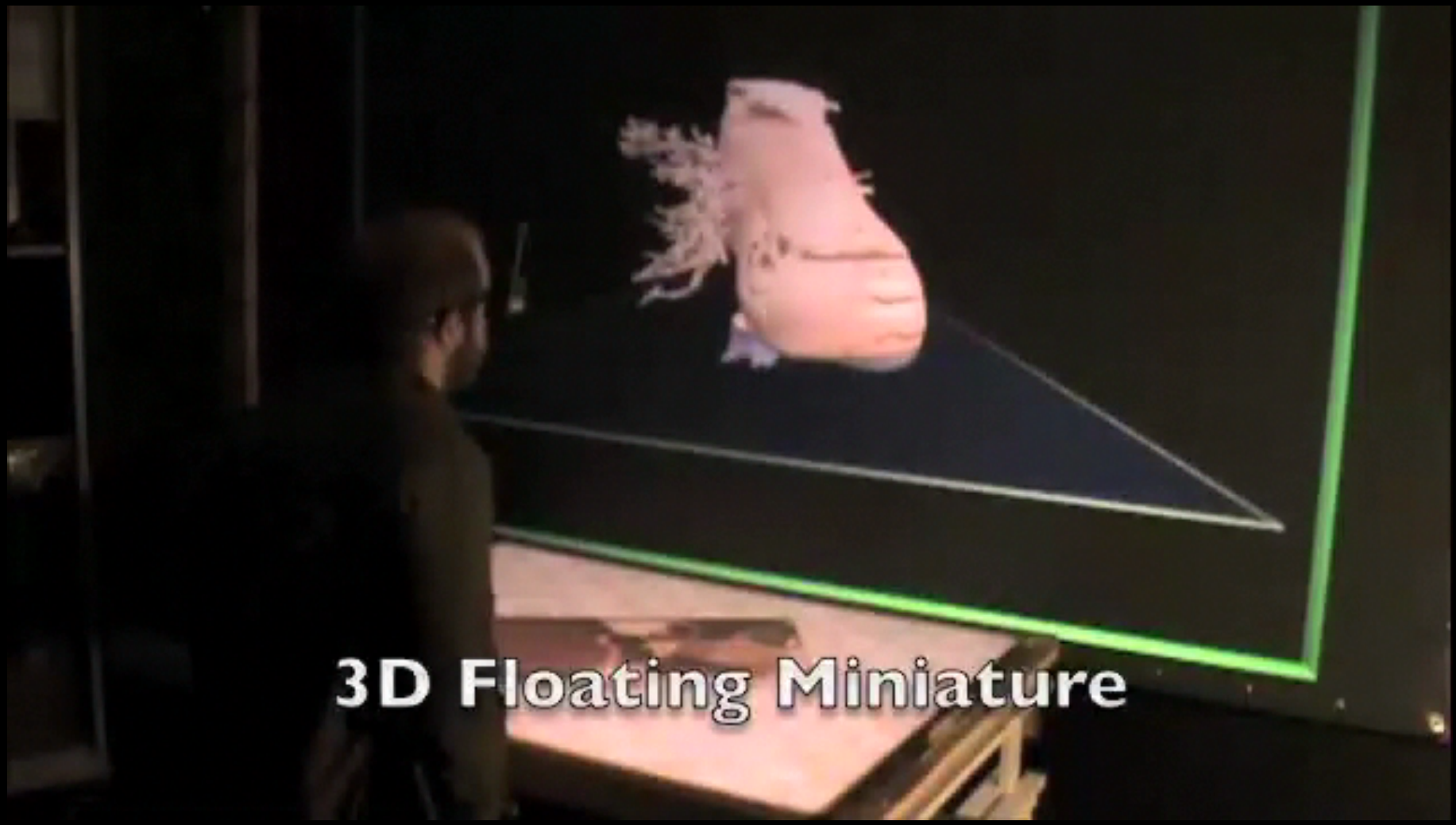
examines the role of primate sensorimotor cortex in control of ingestion, chewing, and swallowing, involves collection of up to 16 channels of physiological data, synchronously with video and neural data.

Prof. Callum Ross et al.

XROMM Data Management platform



Navigation through Complex Anatomical **and Design by dragging**

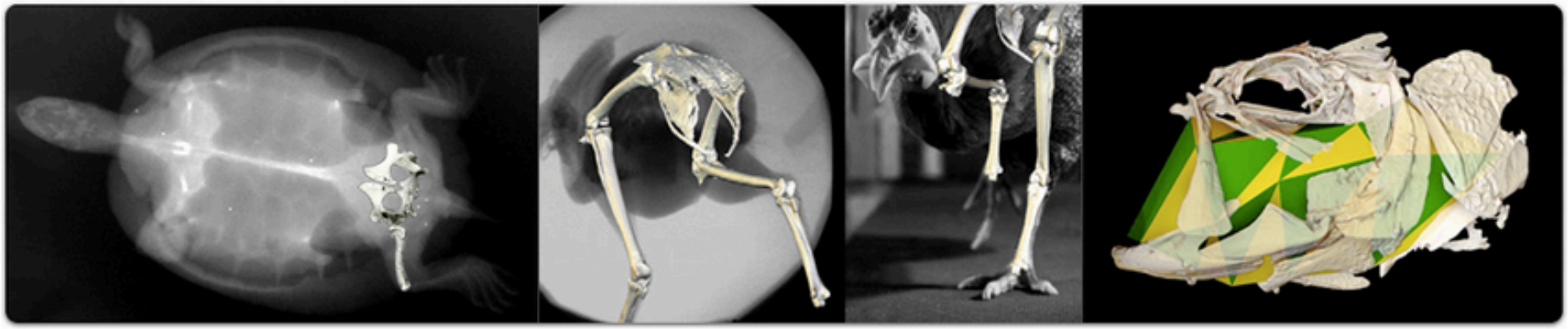


http://www.nlm.nih.gov/research/visible/visible_human.html

XROMM Data Management Platform

- Selection of experimental files for upload through a user interface
- automatic extraction of useful instrument metadata
- Key/value storage of files
- View and query of data
- Populate database
- Automatically upload data to the national XROMM repository

Organize, *Store* and *Share* X-ray Motion Data with XMA Portal



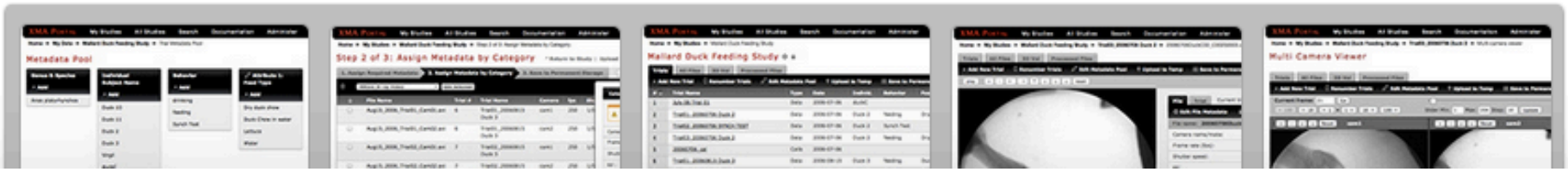
The **X-ray Motion Analysis Portal** is a web environment for management of XROMM data. Non-logged-in users should go to All Studies to explore the organization of the XMA Portal. Click on a Public Study to view video data, and click on Browse (Metadata) to explore the organization and contents of a non-public study.

Use XMA Portal to store and share:

- X-ray videos
- Calibration images
- CT scan data
- Metadata (individuals, treatments, annotations)
- Processed data files
- Access your data from anywhere

Tools for Data Management:

- Metadata Pool for organizing species, individuals, behaviors and treatments
- Multi Camera Viewer for viewing synchronized videos
- Annotation fields for tagging trials and files for later analysis
- Nearly lossless jpg compression for faster video download
- Interface with companion program, XMA Lab, for X-ray Motion Analysis
- Share a whole study or just a few files with specific users or the general public



Thank you

rcc.uchicago.edu