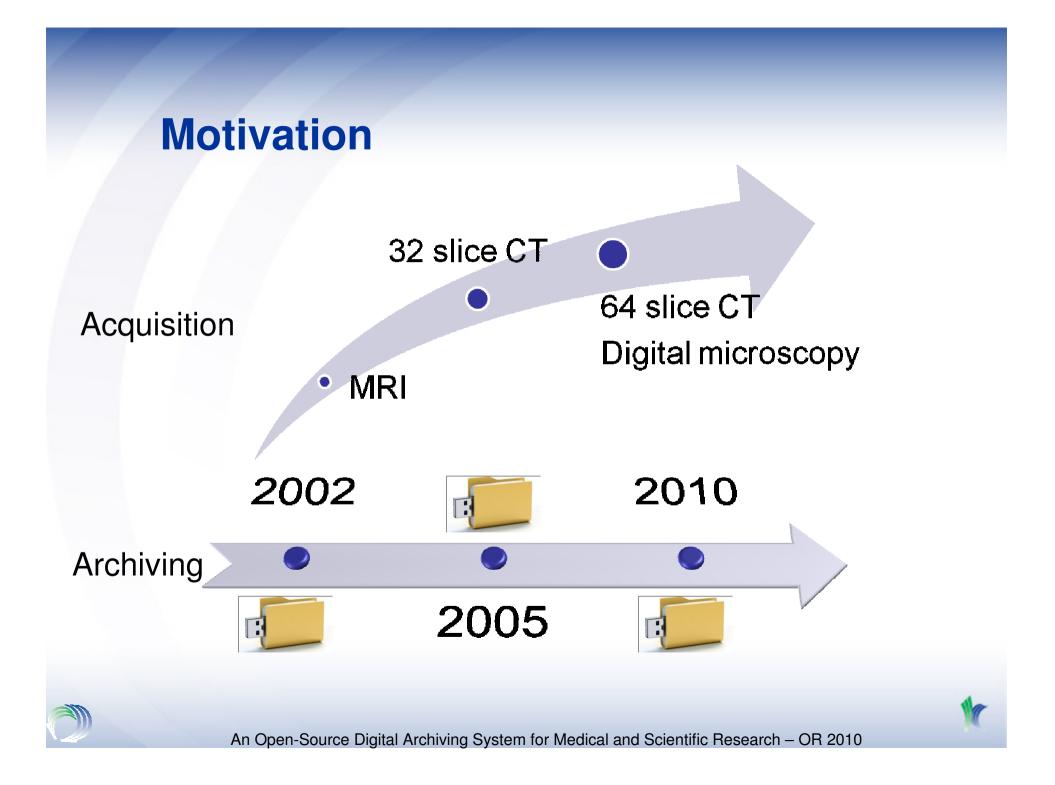


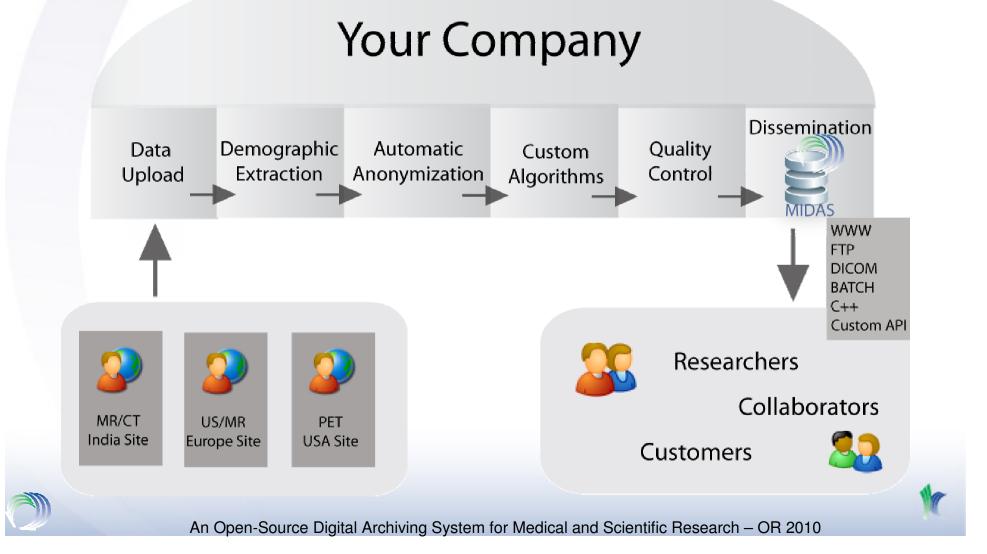
An Open-Source Digital Archiving System for Medical and Scientific Research

Open Repositories 2010

Julien Jomier *Kitware Inc.*



Data Collection, Management and Distribution



Outline

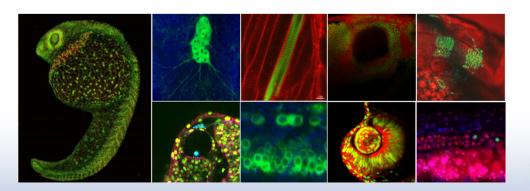
- Challenges in archiving scientific and medical datasets
- Solutions for digital archiving of scientific datasets
- The MIDAS system
- Open repositories using MIDAS
- Future work





Archiving Challenges

- Datasets are massive
- Datasets are heterogenous
- Metadata are embedded in the image format
- Metadata are not standardized
- Data provenance is equally important
- Uploading and Downloading data
- Processing
- Visualization



Ĉ

An Open-Source Digital Archiving System for Medical and Scientific Research – OR 2010

Scientific Archiving Solutions

- Hard drive/Tape Backup
- Revision control systems
- Picture Archiving and Communication System (PACS)
- XNAT: Medical (DICOM) images only
- NBIA: NCI digital archive
- DSpace/EPrint: Publications
- Others...





MIDAS

- Web-based Multimedia Digital Archiving System
- Written in PHP with PostGreSQL backend
- Modular and highly customizable framework
- Open Source (BSD)
- Based on DSpace
- API: REST, C++
- Grid computing
- Online visualization

Home Browse Search Login About					
Welcome to MIDAS!					
MIDAS is a collection of server, client, and stand-alone tools for data archiving, analysis, and access,					
 Hests public and private collections of data 					
 Handles massive collections and images 					
 Manages image and non-image files and meta-data 					
 Supports over 28 types of medical and non-medical image data formats 					
 Integrates BatchWake technology for server-side and stand-alone batch processing 					
 Builds upon open source code and open standards 					
For more information and to establish your own ANDAS repository, contact Kitware.					
For more information and to establish your own means repository, contact weekers.					
Common keywords					
landmarks background regional.minima transforms Deformable Large Hemory Filters Software CIRS ME edge flipping decimation					
NAMIC False Discovery Rate (TK fibers myscardial infection segmentation vectors and matrices Non-destructive inspection					
Inserfunction Hatab regionprops					
Recently uploaded dataset					
interior of the second s					
T1-MPRage (Bullitt E.)					



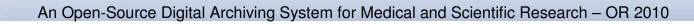
MIDAS Features

- OAI-PMH
- Handle server
- Advanced search
- Upload Filters
- Upload Templates



- WebDAV support (upload/download)
- Comments and Ratings
- MIDAScpp (C++ API)
- MIDASDesktop





OPEN REPOSITORIES USING MIDAS





Digital Archiving with MIDAS

Optical Society of America

- Interactive Science Publishing
- Enhance traditional publishing with 3D rendering
- Sponsored by the National Library of Medicine
- Low-resolution < 10s
- High-resolution = background
- Open Datasets
- http://midas.osa.org

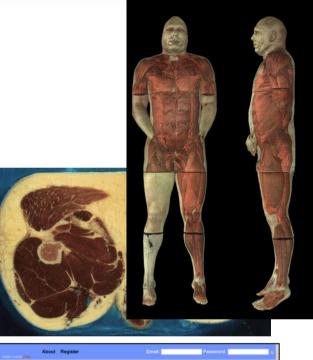


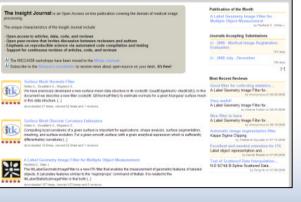


An Open-Source Digital Archiving System for Medical and Scientific Research - OR 2010

National Library of Medicine

- Insight Journal
 - Open-Access Journal
 - 1600+ registered users
 - 775 reviews
 - 400 Publications
- www.insight-journal.org
- Disseminating datasets:
 - Visible Human
 - Army research labs

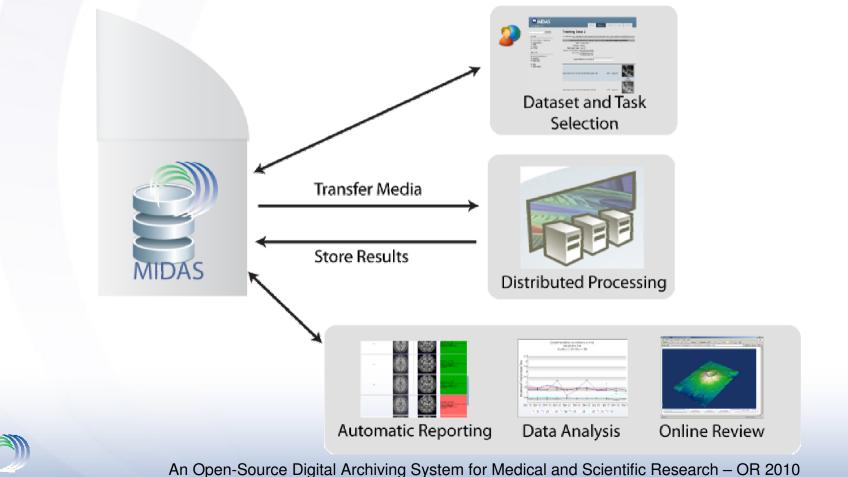






National Cancer Institute

- High Throughput for small animal imaging
- >1TB data every day



Give-A-Scan Project

- Lung Cancer Alliance and Kitware
- First patient-powered open access database
- 10 patients with advanced lung cancer
 - computer tomography scans
 - # pack years
 - respiratory functions
 - patient information
- http://www.giveascan.org

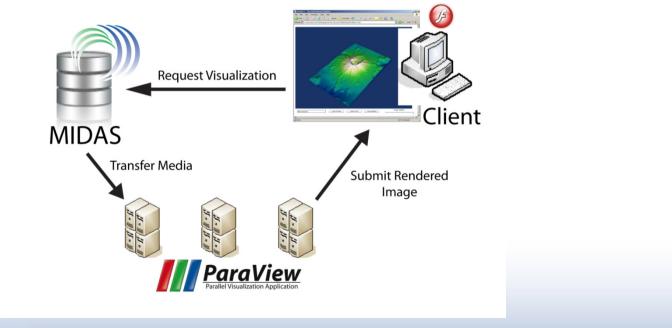
LUNG ALLIANCE	Home	Browse	Search	About	Community		
Search scans Navigation Search scans Lung Cancer Datasets Browse Communities & Collections Image Gallery Titles Autors Date Keywords	Lung Cancer Datasets Community size: 160 Lung cancer is the the leading cause of cancer deaths workdwide. Little progress has been made in increasing the percentage of cases diagnosed at an early curable stage, or in reducing its high mortality rate. Historically lung cancer research has not been funded at levels appropriate to its public health impact. Researchers have also been deterred by the lack of accessible clinical data. In answer to this urgent need and to accelerate research, Lung Cancer Alliance, with the assistance of Kitware, Inc. developed a pilot program to test the feasibility of placing patient donated scans and clinical data on a website that would be accessible by researchers around the world. This collection of the first patient donated, additional sets will be added.						
Related Sites	9 communities within this community						
Lung Cancer Alliance	A T d	P0003 A female lung cancer patient with 83 pack years of expoure was diagnosed with stage IIIA HS This study contains a 3.75mm slice thickness, 240 mA CT scan acquired with contrast enhance diagnosis. This community contains 1 collection and 1 item. last update on 2009-11-19 13:57:49-05 (11/					
	P0004 A 40 year old female subject with a family history of lung cancer was CT screened for lung cancer reconstruction is also provided. This community contains 1 collection and 3 items. last update on 2009-11-19 13:58:41-05 [14						
		his community	contains 1 col	lection and 3	ittems. last update	e on 2009-11-19 13:58:41-05 (14	
		JUOS Female lung cancer patient with 14.5 pack years of expsoure was diagnosed with stage IIIA N id guit smoking prior to the lung cancer diagnosis. This study contains several 5.0mm slice th guired on a Siemens Definition scanner with contrast enhancement at the time[]					
	T	This community contains 1 collection and 5 items. last update on 2009-11-19 13:54:24-05 (10					





Future MIDAS developments

- Current release 2.6 (last week)
- Medical datasets sharing policies
- OSA/Microsoft processing infrastructure
- Visualization on the web





Conclusion

- MIDAS: open-source data archiving/management framework
- Open-source has shown how it can impact Science, Open-data is next
- Data (and associated metadata) is key to scientific Research





Acknowledgments

- The National Institutes of Health
 STTR 5R41NS059095-02
 STTR 2R42NS059095-03
- The National Library of Medicine
- The Optical Society of America
- The University of North Carolina at Chapel Hill

http://www.kitware.com/midas







An Open-Source Digital Archiving System for Medical and Scientific Research

Thank you!

julien.jomier@kitware.com