

# Publishing multi-dimensional image data



Mike Rossner, Ph.D.

Executive Director, The Rockefeller University Press

rossner@rockefeller.edu





#### Mike Rossner, Ph.D.

Executive Director, The Rockefeller University Press rossner@rockefeller.edu











"The future of scientific publishing lies in providing access to the data underlying a publication and giving the reader the ability to interact with those data."

Colloquium on Rethinking the Future of Scientific Communication, March 9, 2012, Stanford University Libraries

## Data need context

Publication

"Publication represents a determination by the scientist authors, reviewers, and editors that a body of work should be delivered to the community for dissemination and consideration. Following this well-established principle, data associated with experiments reported in a publication should be publicly available."

Hill and Swedlow – Response to OSTP RFI on Public Access to Digital Data, January, 2012

"As a condition of publication, scientific journals should progressively enforce requirements for traceable and usable data available through an article, when they are intrinsic to the arguments in that article. This should be in line with the practical limits for that field of research. Materials should be uploaded to a repository before publication of the article, though their release may be subject to a temporary embargo."

Science As An Open Enterprise - June 2012 The Royal Society

- <u>NIH (2003)</u> "The NIH expects and supports the timely release and sharing of final research data from NIHsupported studies for use by other researchers."
- <u>NSF (2010)</u> "Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants."

"The MRC expects valuable data arising from MRC-funded research to be made available to the scientific community with as few restrictions as possible so as to maximize the value of the data for research and for eventual patient and public benefit. Such data must be shared in a timely and responsible manner."

UK Medical Research Council policy on research data-sharing

"Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property."

**RCUK Common Principles on Data Policy** 

"In accordance with important international organisations involved in funding and performing research, the Alliance supports the long-term preservation of, and the principle of open access to, data from publicly funded research."

Alliance of German Science Organisations

"...digitally formatted scientific data resulting from...research supported...by Federal funding should be stored and publicly accessible to search, retrieve, and analyze."

U.S. Office of Science and Technology Policy Memorandum for the Heads of Executive Departments and Agencies February 22, 2013

Implementation? Enforcement?

Funding?

**Publication!** 

#### Journal Mandates on Public Access to Structured Data Sets

Reliance on public repositories

- Protein Data Bank
- Genbank
- GEO (Gene Expression Omnibus)
- Peptidome / Pride



jcb-dataviewer.rupress.org



## The challenges of presenting modern microscopy data

• <u>Multiple file formats</u> – >125

Interpreter: Bio-Formats.

• <u>Multidimensional</u> – x, y, z, color, time

## <u>The challenges of presenting modern</u> <u>microscopy data</u>

#### <u>Number of images</u>

Single z stack of 60 sections monitored over 200 timepoints = 12,000 images (512 x 512 pixels)

There are low resolution ways of representing this, but they badly compress the original data and have limited interactivity.

High content screens – 10<sup>6</sup> images

#### • Size of images

 $(3 \times 10^5)$  pixels x 10<sup>6</sup> pixels = ~300 gigapixels = ~20 Gb 26,000 images, 1.5 mm x 0.6 mm, 16 million dpi

## **JCB**Data Viewer

jcb-dataviewer.rupress.org

- Browser-based application for viewing original image files - from various types of microscopes and gel-documentation systems - associated with JCB articles.
- The first browser-based system for viewing and analyzing multi-dimensional microscope image data.

## **JCB**Data Viewer

jcb-dataviewer.rupress.org

## <u>Authors</u>

- Present original data as acquired. [Link from published paper]
- Share data that were not possible to share previously.

## **JCB**Data Viewer

jcb-dataviewer.rupress.org

## <u>Users</u>

- See data they could not see previously.
- Interact with the data within the browser (scrolling through a z stack or time series; zoom/pan a very large image), make your own movie, and perform simple analyses (e.g. line plots, plots of HCS data).
- Download the data in a standardized format for complex analyses.

#### Home Page



## Supported File Types

jcb-dataviewer.rupress.org/jcb/page/imageformats/			🏠 🔻 😋 🔀 🕶 Google	
JCB Data Viewer	Search the JCB DataViewer	Search	ome 🚯 About 🔤 Contact	JCB   log
About	Supported File Types	;		
Instructions for Use	Format	Extensions 🗢	Comments	
Supported File Types	Adobe Photoshop PSD	.psd		
FAQ	Alicona 3D	.al3d		
Terms of Use	Amersham Biosciences GEL	.gel		
	Amira Mesh	.am, .amiramesh, .grey, .hx, .labels		
	Analyze 7.5	.img, .hdr		
	Animated PNG	.png		
	Aperio SVS TIFF	.SVS		
	ARF	.arf		
	AVI (Audio Video Interleave)	.avi	Uncompressed (raw) or compressed with select lossless compression codecs (rle, Apple RPZA (rpza))	
	Bio-Rad PIC	.pic,.xml,.raw	Excluding multi-plane "stitched" files and annotated ROI images	
	Bitplane Imaris	.ims		
	Bitplane Imaris 3 (TIFF)	.ims		
	Bitplane Imaris 5.5 (HDF)	.ims		
	Cellomics	.c01		
	Compix Simple-PCI	.cxd		
	DeltaVision	.dv, .r3d, .r3d_d3d, .dv.log		
	DICOM	.dicom,.dic,.dcm,.jp2,.j2ki, .j2kr,.raw	Uncompressed only	
	EPS (Encapsulated PostScript)	.eps,.epsi,.ps		
	Evotec/PerkinElmer Opera Flex	.flex		
	FEI	.img		
	FITS (Flexible Image Transport	.fits		

#### Integration with the literature



#### Main data page / minimal viewer



## **Full Viewer**



## Individual channel view







#### **Split Channel View**



#### Line Plots



#### High Content Screening



#### High Content Screening



#### High Content Screening







The Journal of Cell Biology – Executive Editor. Liz Williams | Rockefeller University Press | Home | About | Contact | Current Issue Copyright @ 2007-2013 Glencee Software Inc. All rights reserved. Bull, Powered and Run by Glencee Software. Inc. THE ROCKEFELLER UNIVERSITY PRESS





The Journal of Cell Biology - Executive Editor: Liz Williams | Rockefeller University Press | Home | About | Contact | Current Issue Copyright @ 2007-2013 @encoes Software hor. All rights reserved. Built, Powered and Run by Glencee Software, hor. THE ROCKEFELLER UNIVERSITY PRESS Home page - http://jcb-dataviewer.rupress.org

Link from HW to 5 dimensional image - <u>http://jcb.rupress.org/content/186/1/11.full</u>

4 dimensional image - <u>http://jcb-dataviewer.rupress.org/jcb/browse/105/134/</u>

Multiple images -

http://jcb-dataviewer.rupress.org/jcb/browse/5468/16161/

Blots - http://jcb-dataviewer.rupress.org/jcb/img\_detail/133657/15610/

HCS display - Thorpe Data -

http://jcb-dataviewer.rupress.org/jcb/browse/4608/S1/

HCS display - Rohn Data -

http://jcb-dataviewer.rupress.org/jcb/browse/4609/S2/

Very large image -

http://v.jcb-dataviewer.glencoesoftware.com/webclient/img\_detail/201/



jcb-dataviewer.rupress.org

- Standard for publication of image data.
- Precursor to an international repository of original, published image data