

pySPIDER and viSPIDER

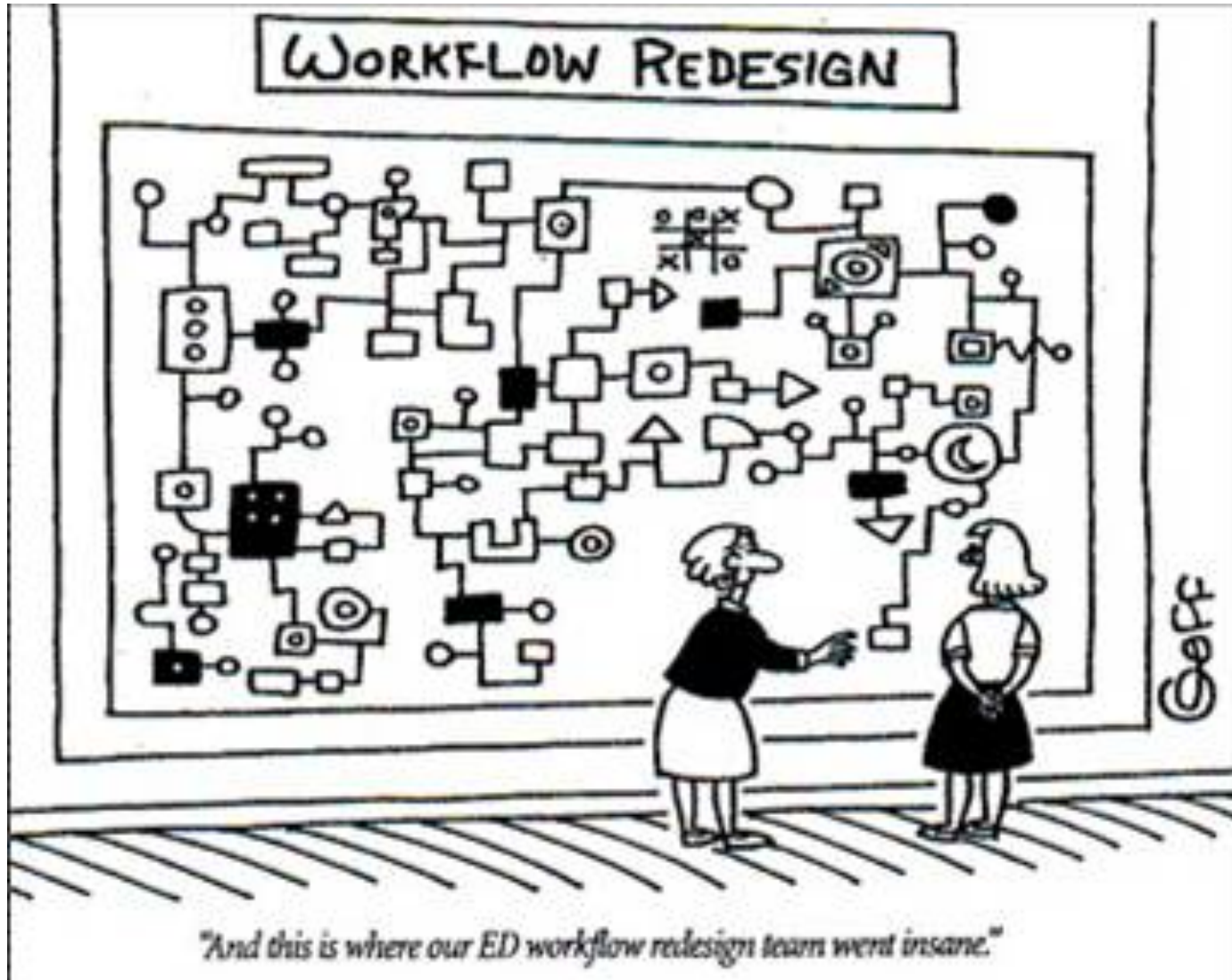
New tools for SPIDER programming and interoperability with other packages

Robert Langlois (HHMI & Columbia)

Jonathan Ginsburg (Yeshiva)

Joachim Frank (HHMI & Columbia)

What is a workflow?



SPIDER History

- Developed in 1978
- With its own scripting language

Simplify the Scripting Language

- CryoEM software packages use Python
 - EMAN2, Sparx, Appion, IPLT
- Extensive libraries
 - PyQt4, Scipy, Numpy, Scikits
- Database SQL Wrappers

Other Reasons to use Python

- Expressive language
 - Generators, keyword arguments, ...
- Advanced coding tools
 - Syntax highlighting, auto-completion, ...
 - Code validity checking
- Documentation tools
 - Sphinx

pySPIDER Design

- SPIDER Session Module
 - Communication with SPIDER
 - Interfaces for SPIDER types: in core files, templates, etc.
- SPIDER Command Module
 - Updated interface to Spider commands
 - New SPIDER commands
 - Version control – missing commands, bugs

SPIDER Command as Python Function

```
def bl(session, image_size, background=2.0, outputfile=None, **extra):
    ''' Creates an image/volume with a specified background

    alias: blank

    `Original Spider (BL) <http://www.wadsworth.org/spider\_doc/spider/docs/man/bl.html>`_

    :Parameters:

        session : Session
            Current spider session
        image_size : (float,float)
            Size of the image
        background : float
            Background color of the image (Default: 2.0)
        outputfile : str
            Filename of the output image (Default: None / e.g. create an incore file)
        extra : dict
            Unused key word arguments

    :Returns:

        outputfile : str
            Int or str describing the output file
    ...

    _logger.debug("Creating blank image of size %s in file %s: "%(str(image_size), str(outputfile)))
    if outputfile is None: outputfile = session.temp_incore_image(hook=session.de)
    session.invoke('bl', spider_fd(outputfile), spider_coord_tuple(image_size), 'N', spider_tuple(background))
    if session.extra_check: ensure_output_file_exists(session, outputfile)
    return outputfile
```

With Python SPIDER commands

- pySPIDER batch files
 - MPI alignment, refinement
 - EMAN2/Sparx reconstruction
- viSPIDER visual workflow editor
 - Program lab view like applications
 - Hierarchical workflows
 - Module for Vistrails

pySPIDER Batch File

```
def batch(args, param_file, output, **extra):
    ''' Main entry point for multireference supervised classification
    '''

    from vispider.core import spider

    spi = spider.open_session(args, **extra) # Start a s
    spider_params.read_parameters_to_dict(param_file, extra) # Add CTF p
    angle_doc, angle_num = spi.vo_ea(**extra) # Generate
    defocus_dict = read_defocus(**extra) # Read defc
    output = spi.replace_ext(output) # Ensure ou
    root, base = os.path.dirname(output), os.path.basename(output) # Split out
    for inputfile in args:
        subdir = os.path.join(root, "ref_"+os.path.splitext(os.path.basename(inputfile))[0])
        try: os.mkdir(subdir)
        except: pass
        for id, vals in defocus_dict.iteritems():
            defocus = vals.defocus
            soutput = os.path.join(subdir, spider_utility.spider_filename(base, id))
            project(spi, inputfile, angle_doc, angle_num, soutput, defocus=defocus, **extra)

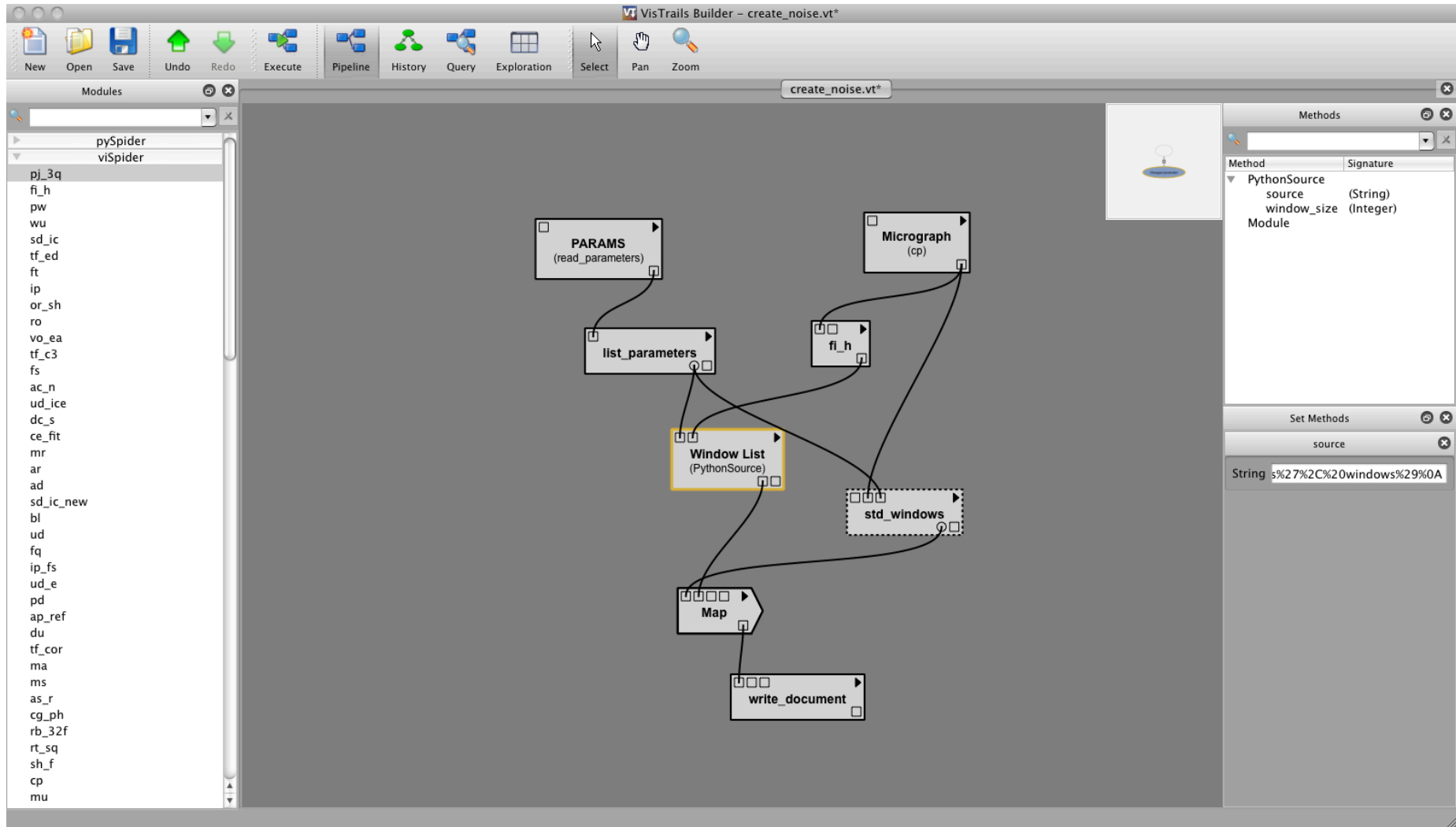
def read_defocus(defocus_file="", defocus_header="id:0,defocus:1", **extra):
    '''Read defocus document file.

    defocus_dict = format.read(defocus_file, header=defocus_header, numeric=True)
    return format_utility.map_object_list(defocus_dict)
```

With Python SPIDER commands

- pySPIDER batch files
 - MPI alignment, refinement
 - EMAN2/Sparx reconstruction
- viSPIDER visual workflow editor
 - Program LabView-like applications
 - Hierarchical workflows
 - Module for Vistrails

viSPIDER Example

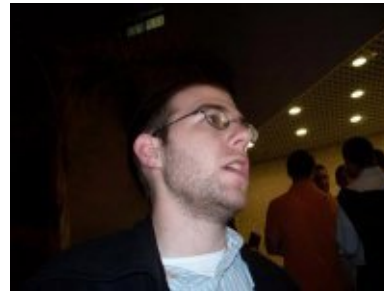


Future Work

- Graphics Integration into Vistrails
 - Display images
 - Display volumes
- Visual batch file generator
 - MPI implementations
- Database for Meta-data
 - Spider document files

Acknowledgements

- Joachim Frank (PI)
- Original Work – Summer Intern
 - Jonathan Ginsburg
- Ideas and *Tireless* Quality Assessment
 - Amedee
 - Yaser
 - Jesper
- New data and input on programs
 - Nam (new particle selection challenge)
 - Bo (ap-cover)
 - Nash
 - Hstau
- Support (Rebooting nodes)
 - Harry



Tornubari Barinee
Administrative Assistant



Bob Grassucci
Senior Microscopist



Bo Chen
Graduate Research



Sanchoita Das
Post-doc



Amedee des Georges
Post-doc



Cheri Hampton
Post-doc



Yaser Hashem
Post-doc



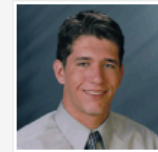
Nam Ho
Graduate Research



Amy Jobe
Graduate Research



Harry Kao
Computer Analyst



Drew Kennedy
Graduate Student



Robert Langlois
Post-doc



Wen Li
Research Scientist



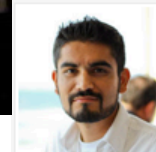
Hstau Liao
Research Scientist



Jesper Pollesen
Post-doc



Nadia Severina
Research Technician



Gyanesh Nash Sharma
Post-doc



Bingxin Shen
Post-doc



Ming Sun
Graduate Student



Melissa Thomas
Graphics & Website