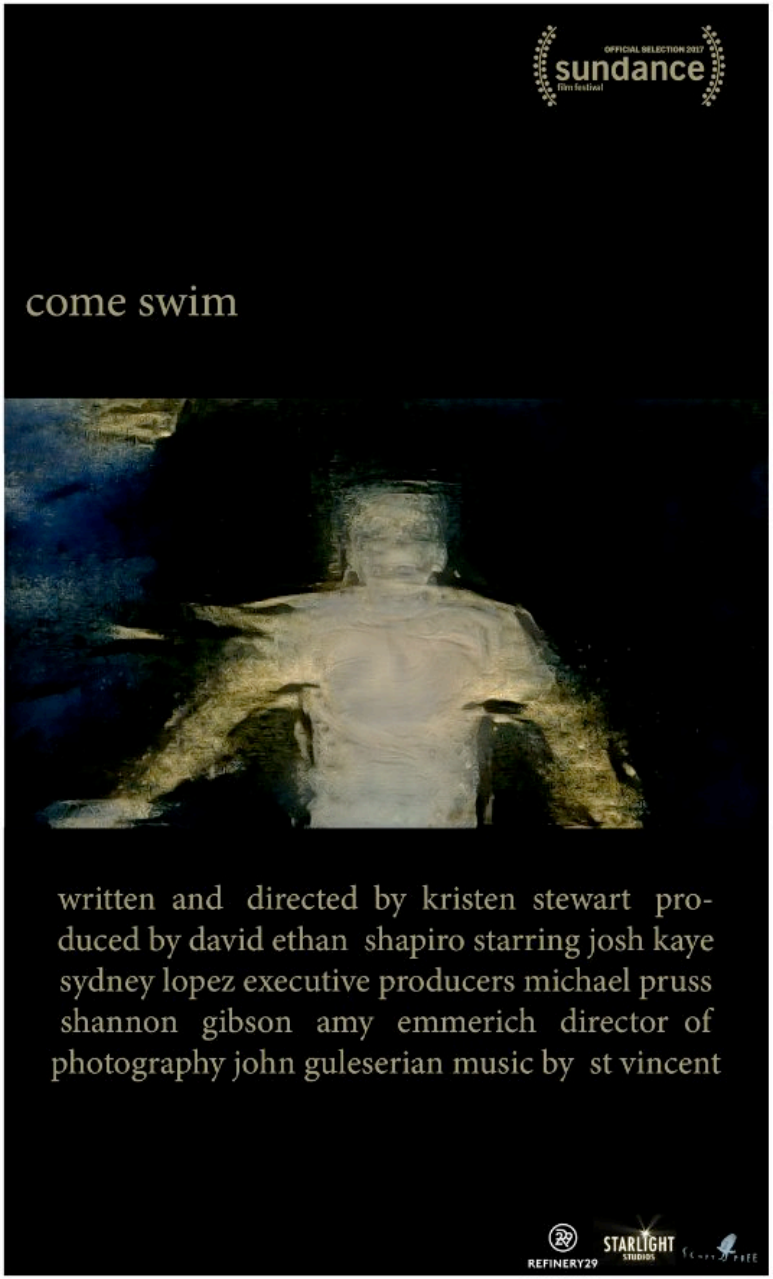
USC INFORMATION SCIENCES INSTITUTE

USC Viterbi School of Engineering

An Introduction to Workflows and WINGS

Yolanda Gil

Information Sciences Institute
and Department of Computer Science
University of Southern California
http://www.isi.edu/~gil
@yolandagil
gil@isi.edu


OFFICIAL SELECTION 2017
sundance
film festival
come swim
written and directed by kristen stewart produced by david ethan shapiro starring josh kaye sydney lopez executive producers michael pruss shannon gibson amy emmerich director of photography john guleserian music by st vincent
REFINERY29
STARLIGHT STUDIOS


2017
SUNDANCE
FILM
FESTIVAL


IMDb

https://arxiv.org/abs/1701.04928

Search

Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > cs > arXiv:1701.04928

Search or Article ID inside arXiv | All papers | Broaden your search using Semantic Scholar

(Help | Advanced search)

Computer Science > Computer Vision and Pattern Recognition

Download:

Bringing Impressionism to Life with Neural Style Transfer in *Come Swim*

Bhautik J Joshi*
Research Engineer, Adobe

Kristen Stewart
Director, *Come Swim*

David Shapiro
Producer, Starlight Studios

Figure 1: *Usage of Neural Style Transfer in Come Swim; left: content image, middle: style image, right: upsampled result. Images used with permission, (c) 2017 Starlight Studios LLC & Kristen Stewart.*

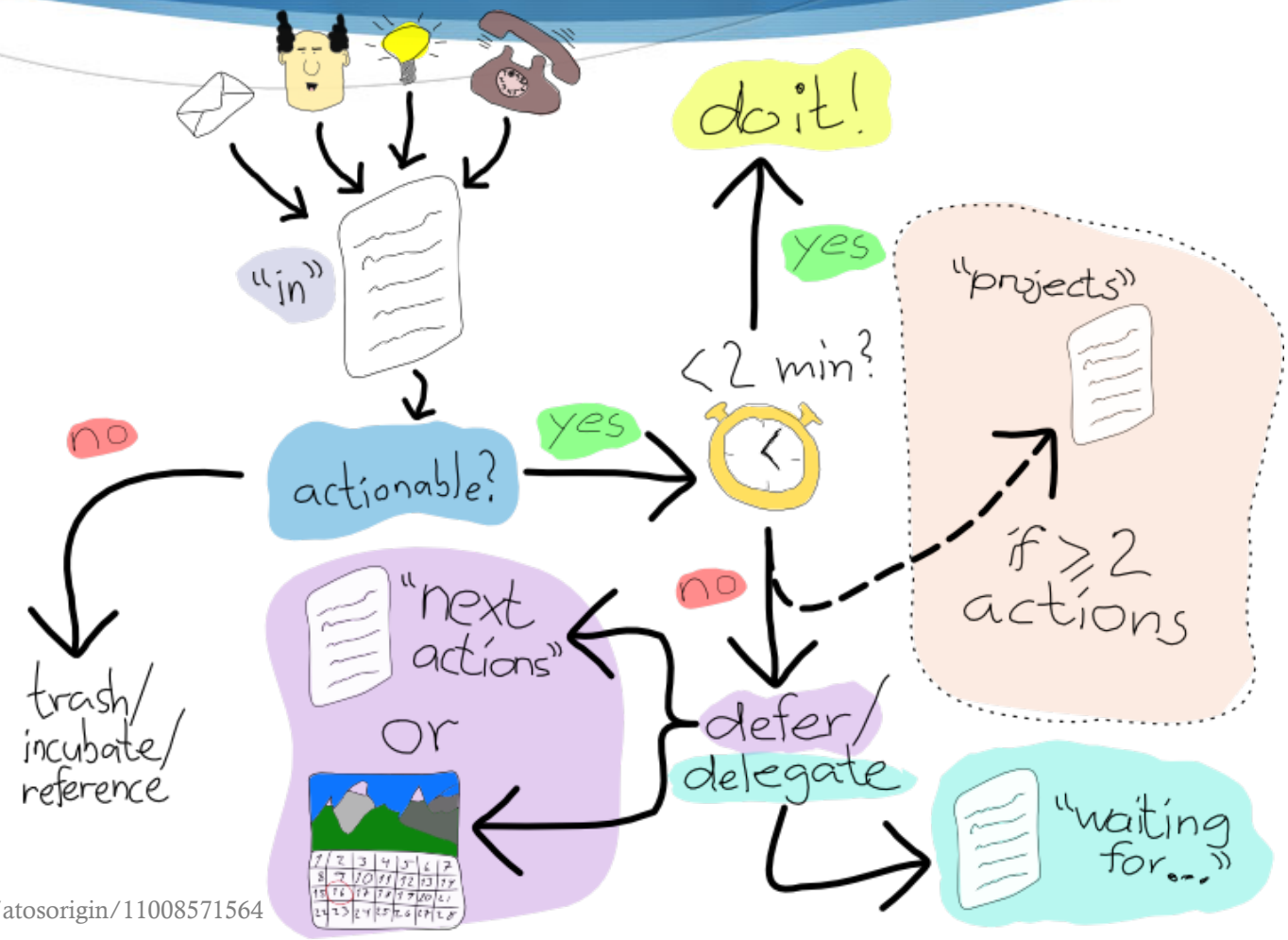
Abstract

Neural Style Transfer is a striking, recently-developed technique that uses neural networks to artistically redraw an image in the style of a source style image. This paper explores the use of this technique in a production setting, applying Neural Style Transfer to redraw key scenes in *Come Swim* in the style of the impressionistic painting that inspired the film. We document how the technique execute efficiently and predictably. In a production setting, however, a great deal of creative control is needed to tune the result, and a rigid set of algorithmic constraints run counter to the need for this creative exploration. While early investigations to better map the low-level neural net evaluations to stylistic effects are underway [Li et al. 2017], in our paper we focused on examining the higher-level parameter space for Neural Style Transfer and found a set of working shortcuts to map them to a reduced but meaningful

Topics

1. Computational workflows
2. Benefits of using workflows
3. Workflow systems
4. Semantic workflows

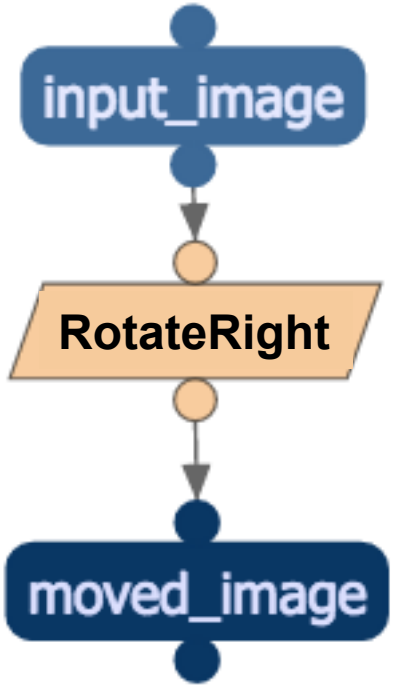
"Workflow" Is a Common Term to Denote Organized Activities

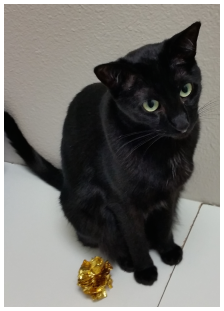


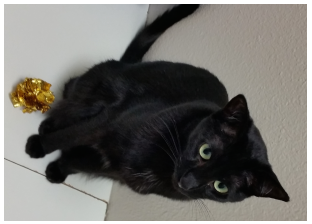
https://www.flickr.com/photos/atosorigin/11008571564

Treating Workflow Components as "Black Boxes"

- You **don't have to understand** the inner workings in order to **use the component**
- This is why we often refer to software as a **"black box"**
- You do need to understand **inputs/outputs/parameters** and the **program's function**

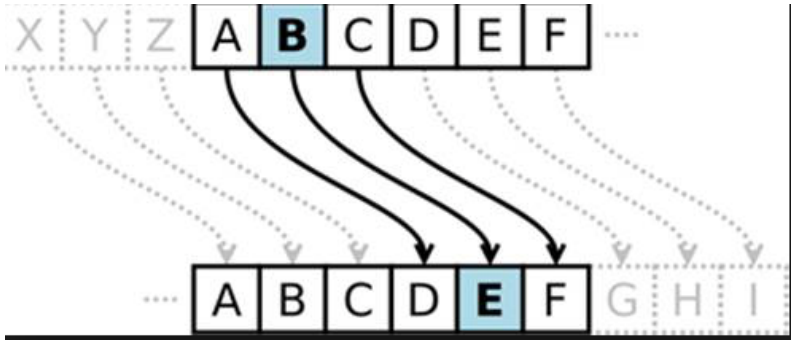




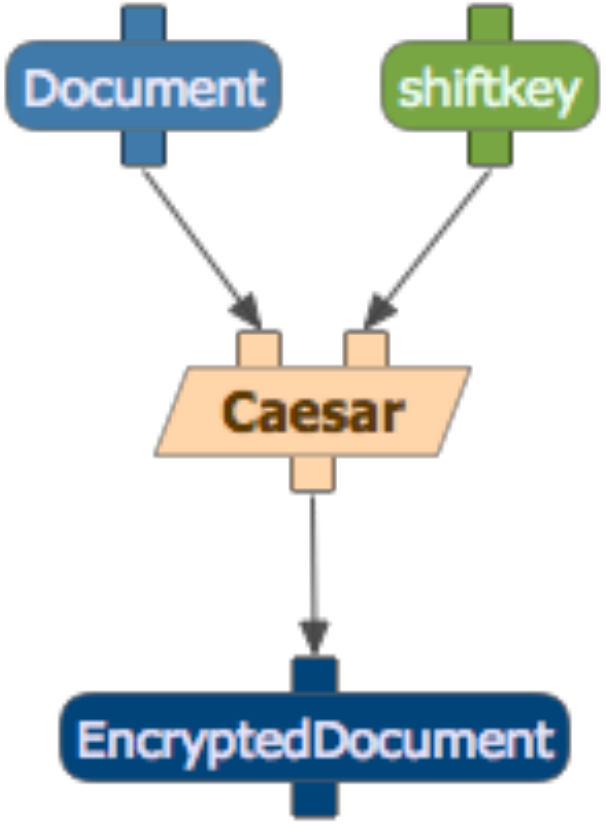


Components as Black Boxes

Original: BAD



Cipher: EDG



Software for Data Analysis

1) Commercial statistical packages

- SPSS
- SAS
- Stata
- MATLAB
- Mathematica
- …
- (Excel)

2) Open source (free) software

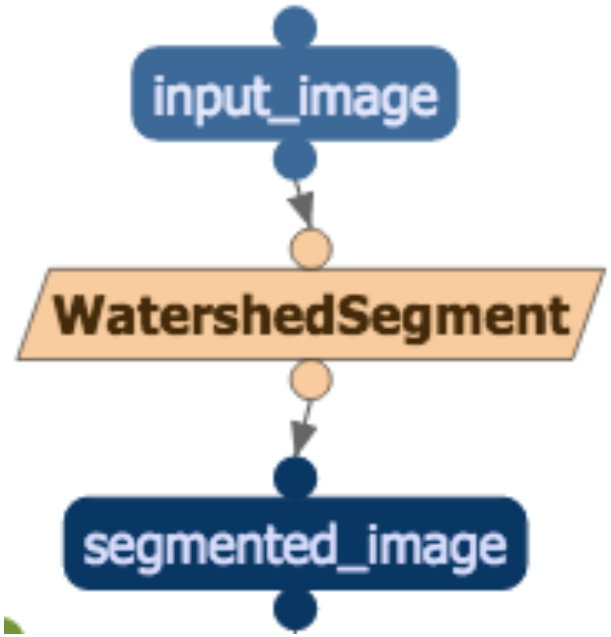
- R package
- Python libraries
- OpenCV for image processing
- NLTK for text processing
- …

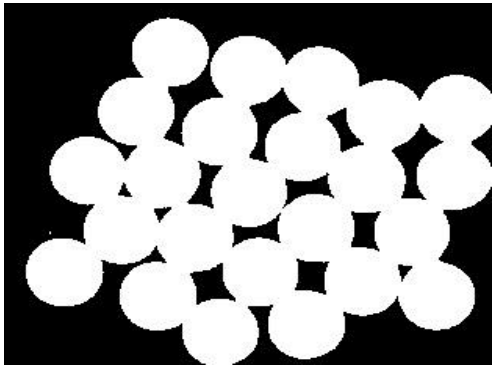
3) Custom software

- Scripts
- Functions built by
- developers
- …

- Can create workflow components by exposing a function (i.e., a command line invocation)

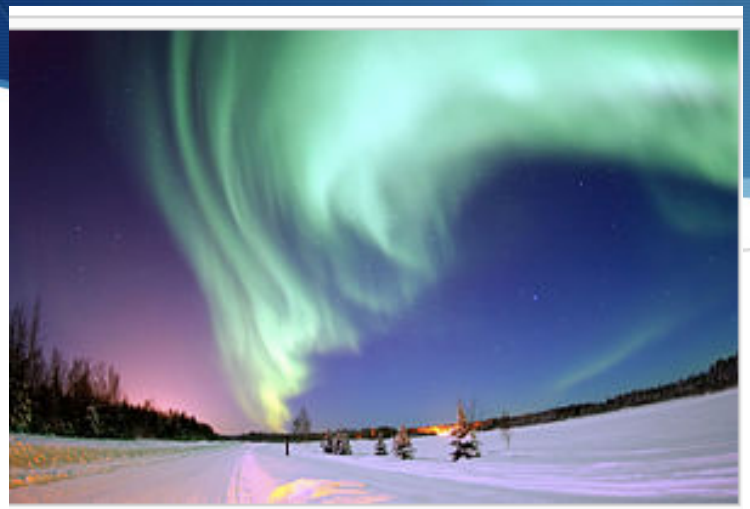
Segmentation

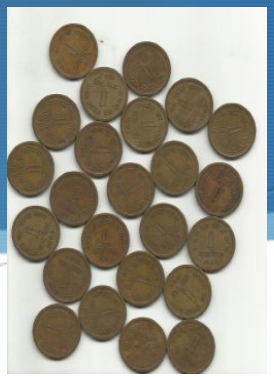


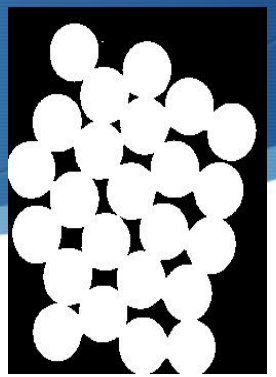


http://docs.opencv.org/master/d3/db4/tutorial_py_watershed.html

Segmentation

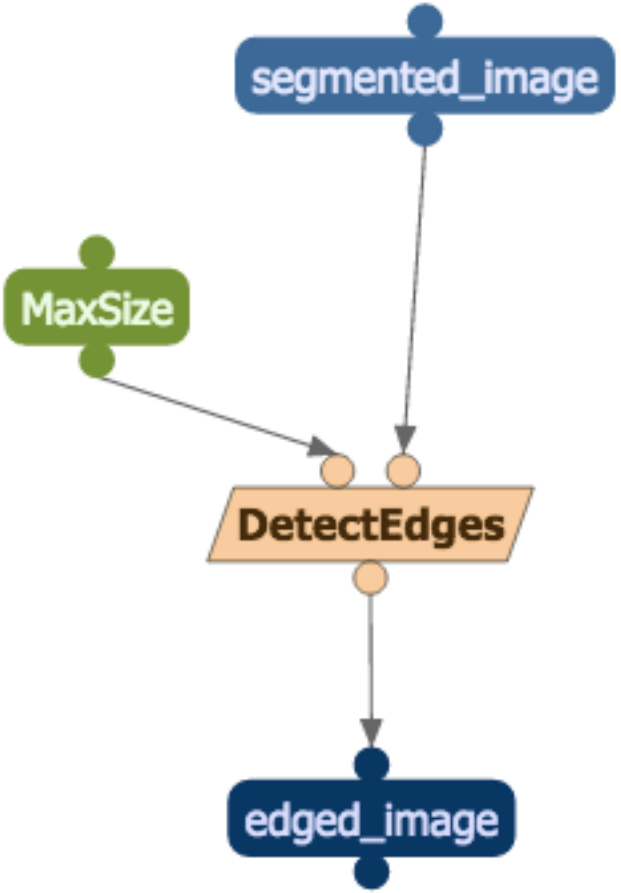






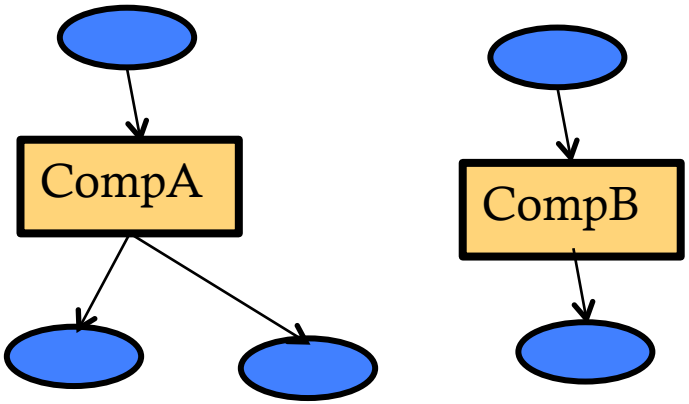
- Dividing an image into regions
  - Each region contains “similar” pixels
- Useful for detecting objects

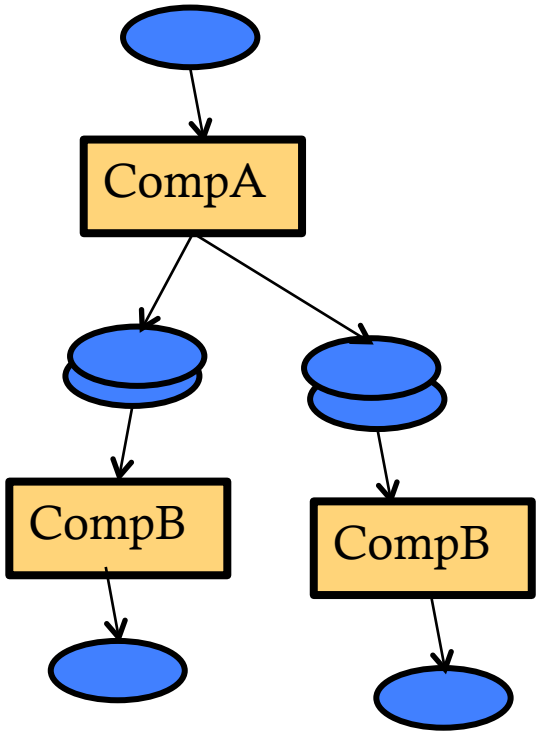
Edge Detection



https://en.wikipedia.org/wiki/Edge_detection#/media/File:%C3%84%C3%A4retuvastuse_n%C3%A4ide.png

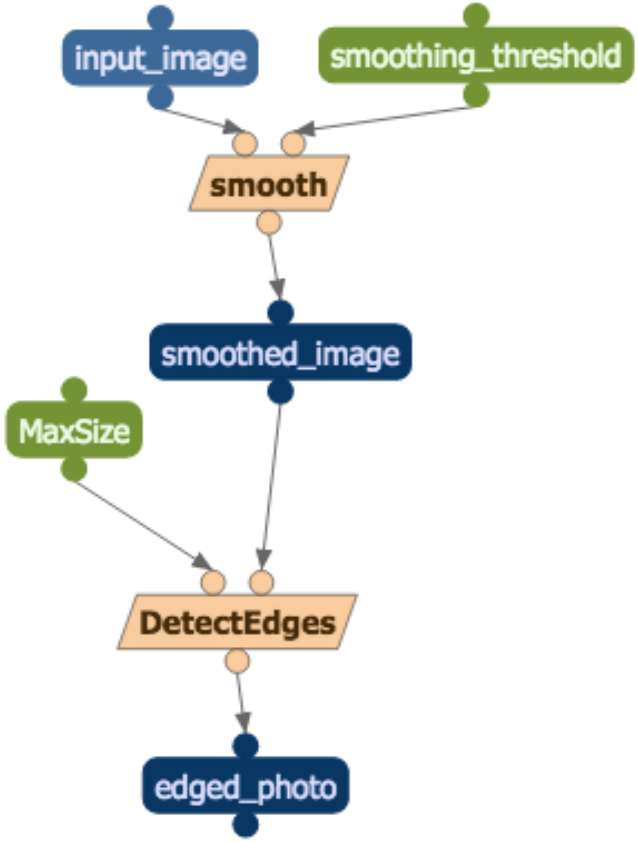
Composing Functions





Computational Workflows

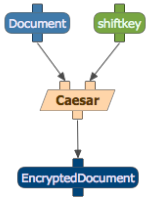
- A computational workflow is a composition of functions implemented as software components

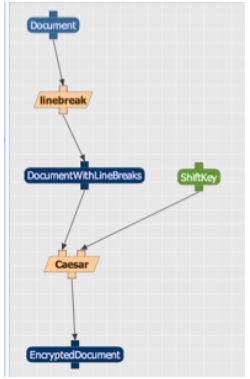


Topics

1. Computational workflows
2. Benefits of using workflows
3. Workflow systems
4. Semantic workflows

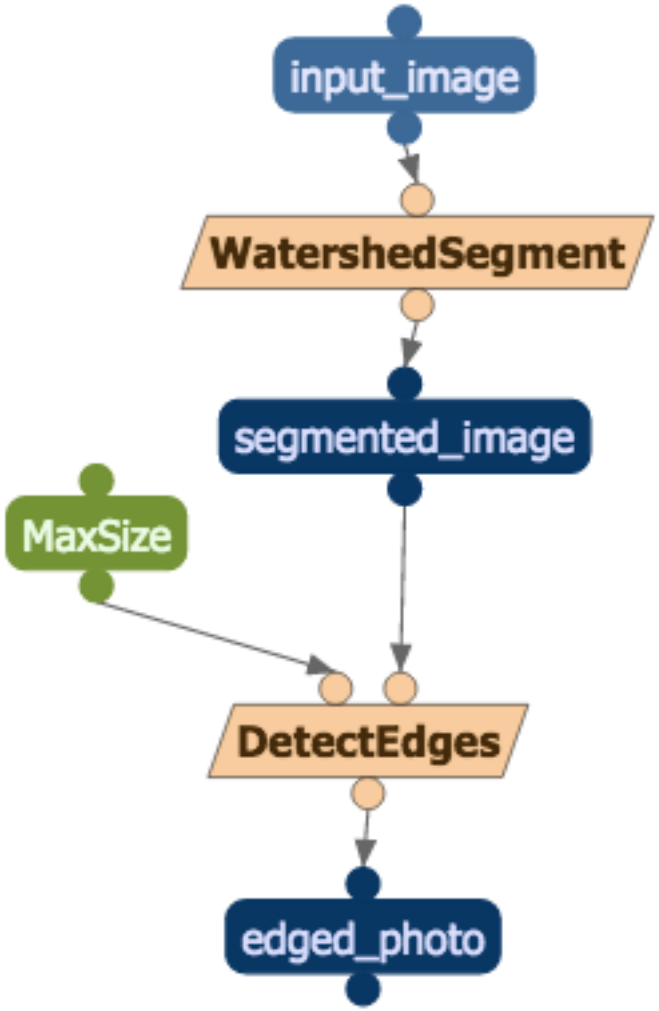
Simple Programming Paradigm

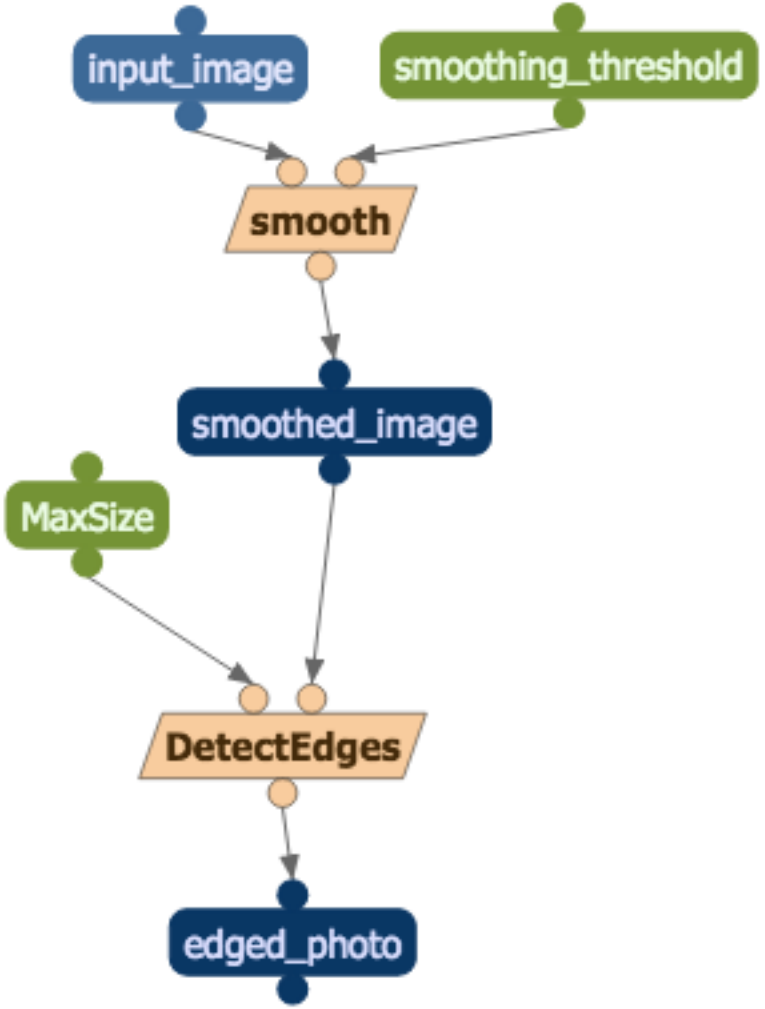




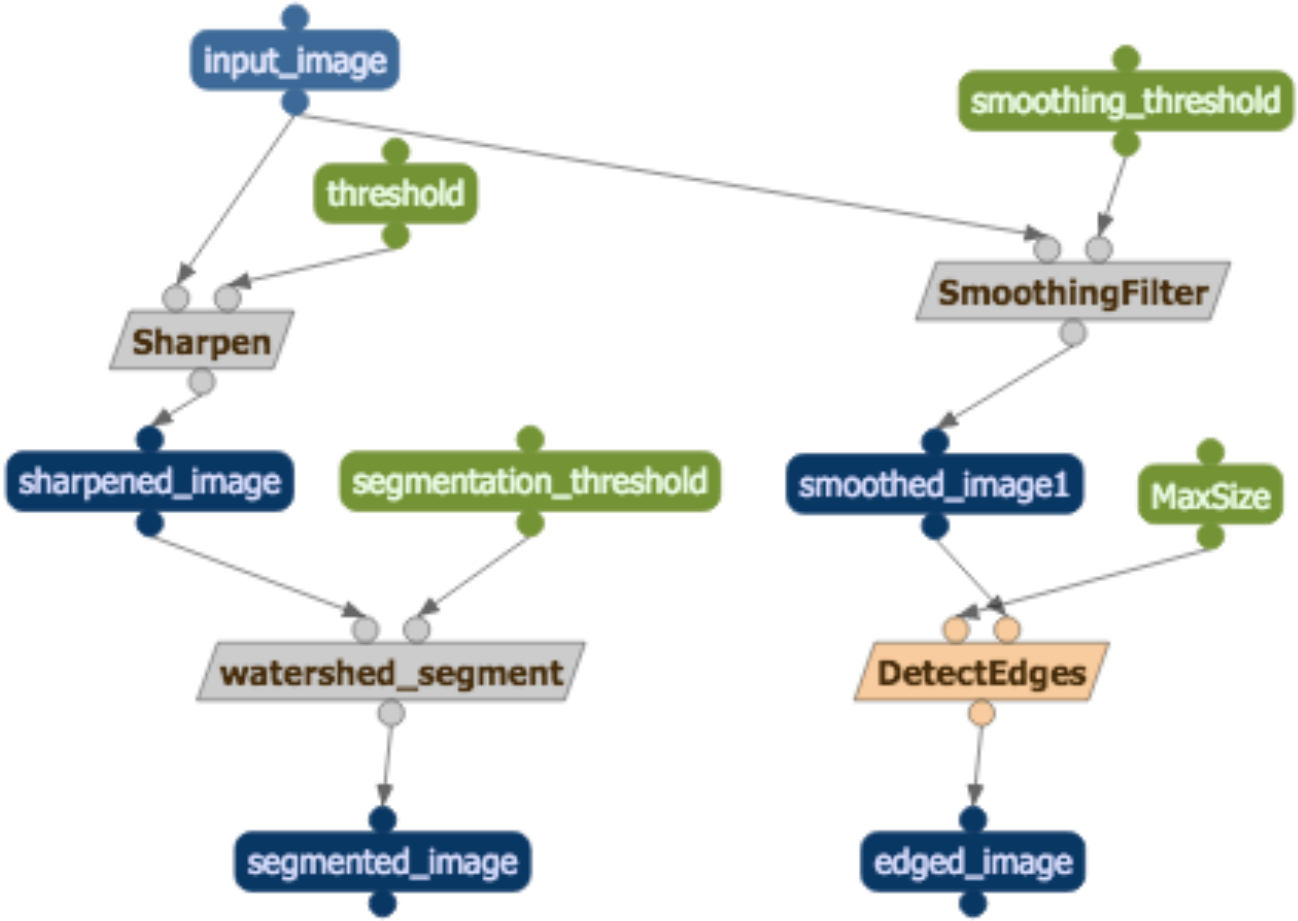
https://www.flickr.com/photos/pasukaru76/5571502641
https://www.flickr.com/photos/georgivar/5535049084

Many Compositions to Try

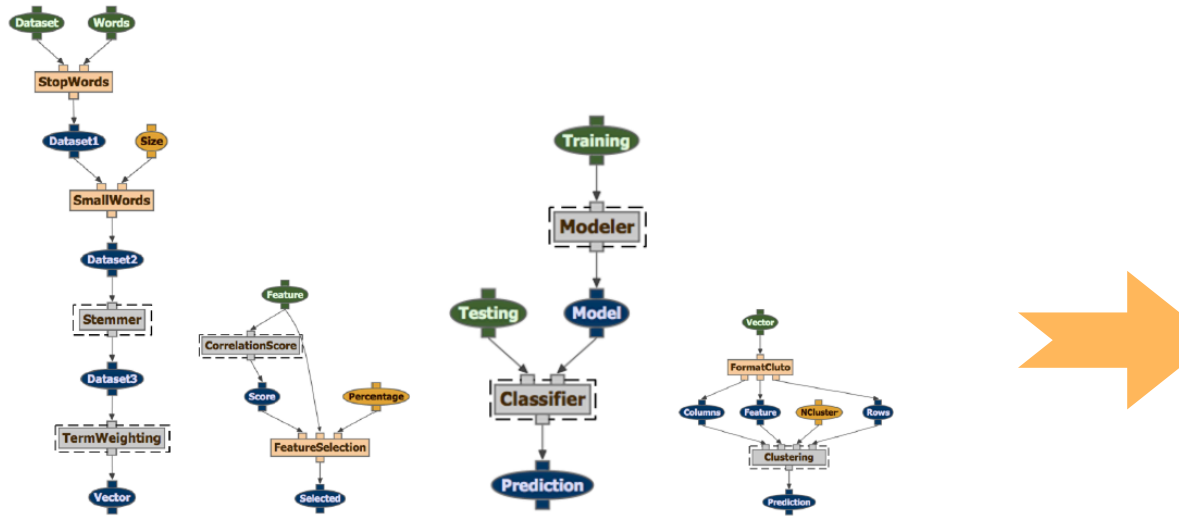


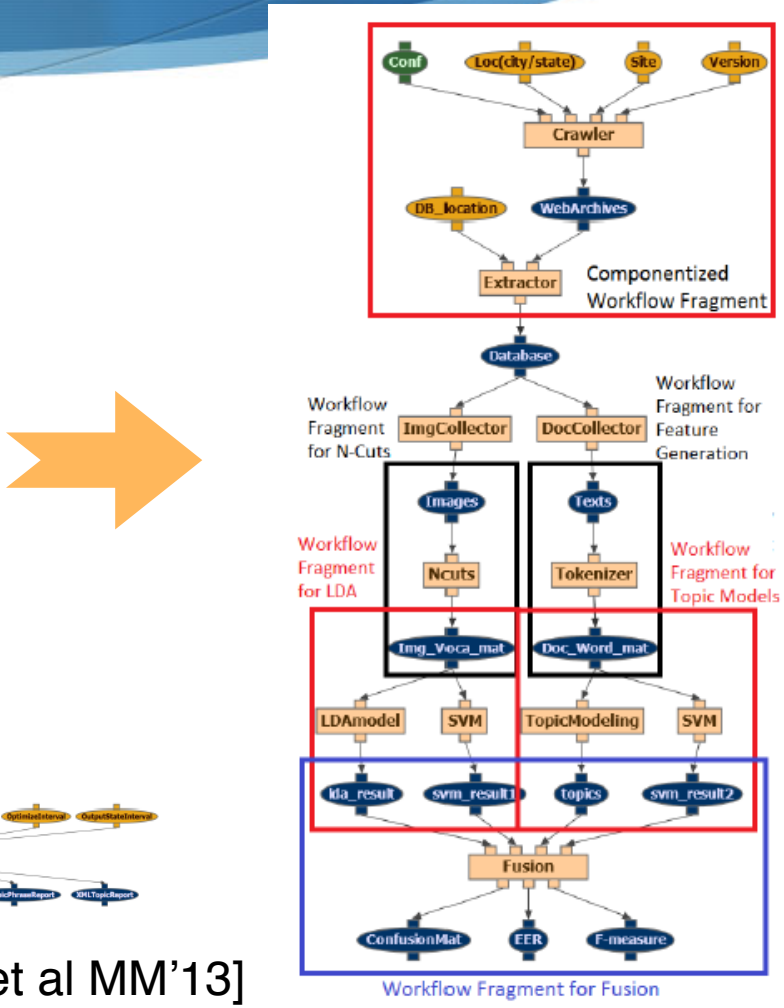


More Complex Compositions



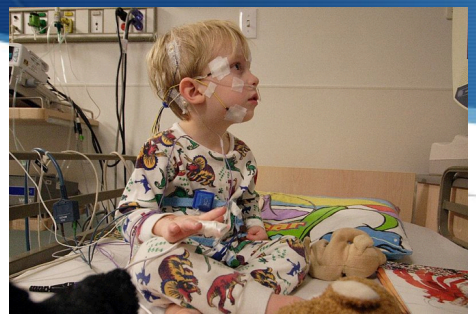
Modular Assembly





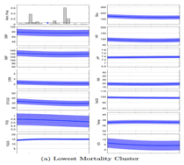
[Sethi et al MM'13]

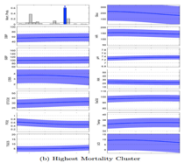
Facilitating Communication Across Data Science Expertise Areas



Describe problem

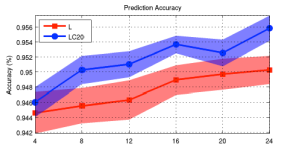
Provide data



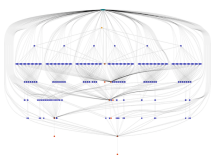


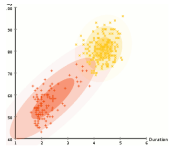
Show results

Point out issues



Show more results

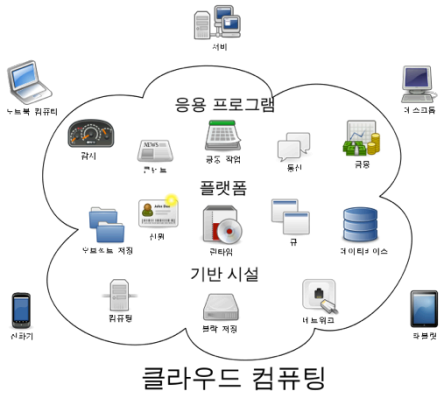




CODE

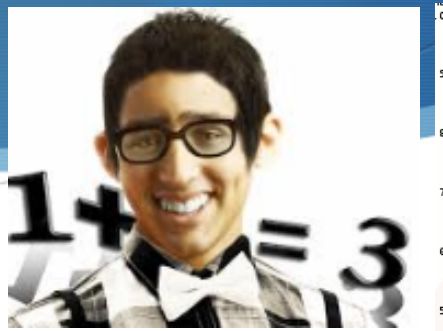
Data Science Teams:
1) The Domain Experts

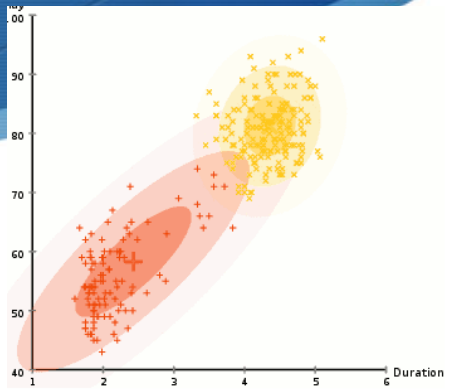
Data Science Teams:
2) Semantics and Data Integration Masters

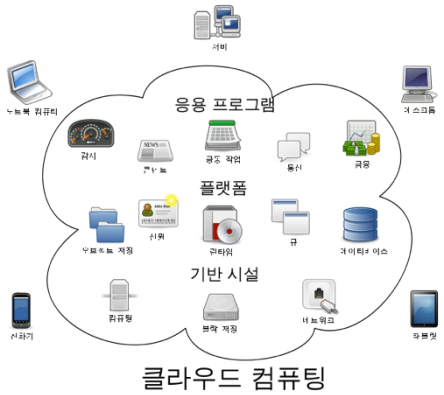


클라우드 컴퓨팅

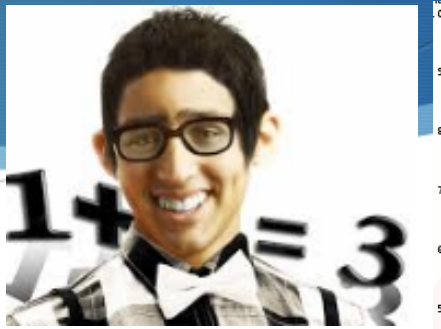
Data Science Teams:
3) The Math Whizes

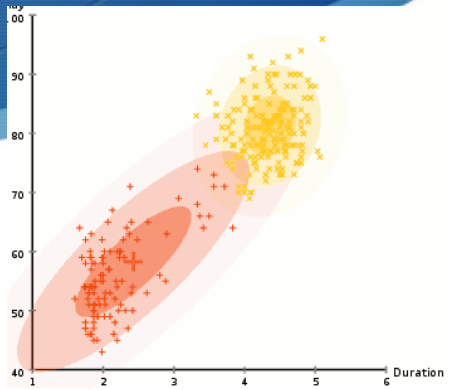


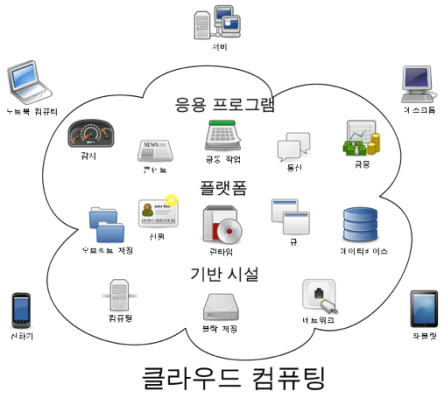


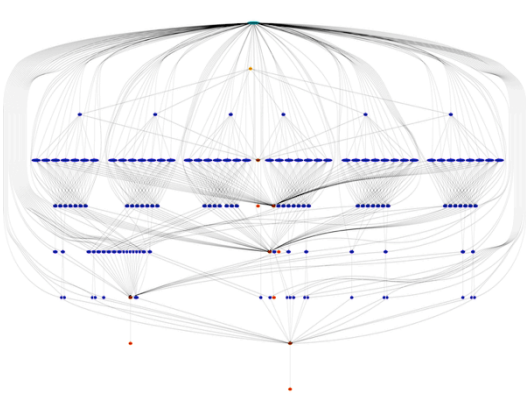


Data Science Teams:
4) The Scalability Hackers

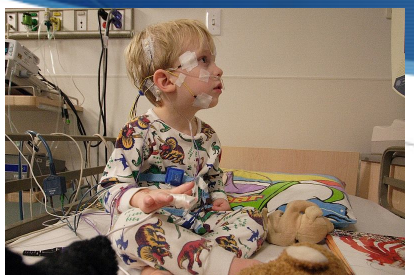




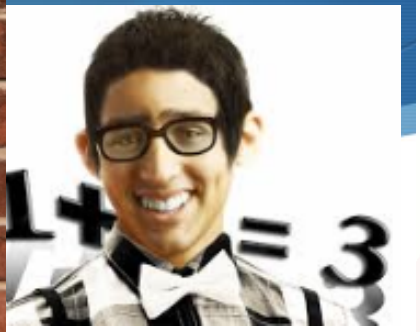


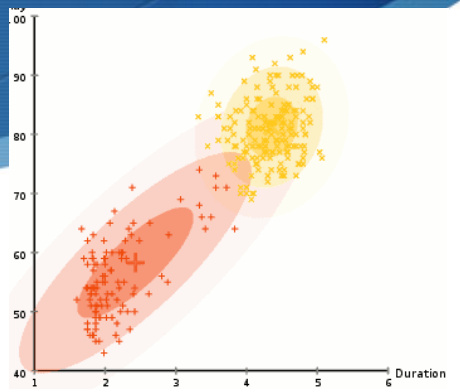


Distinct Expertise in Data Science



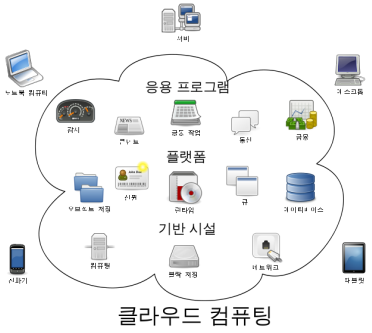
Domain knowledge





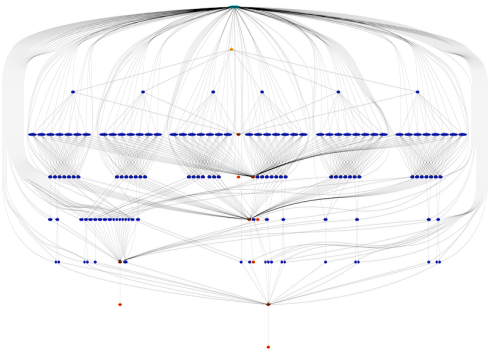
Statistics, data mining

Semantics and data integration

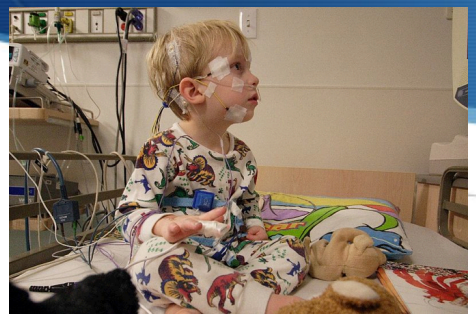


클라우드 컴퓨팅

Large-Scale Data Processing

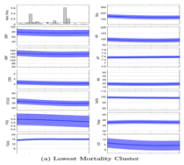


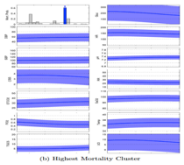
Facilitating Communication Across Data Science Expertise Areas



Describe problem

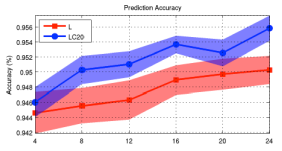
Provide data



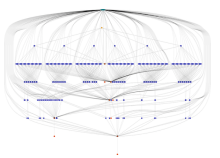


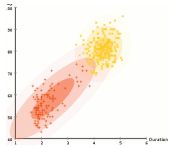
Show results

Point out issues



Show more results

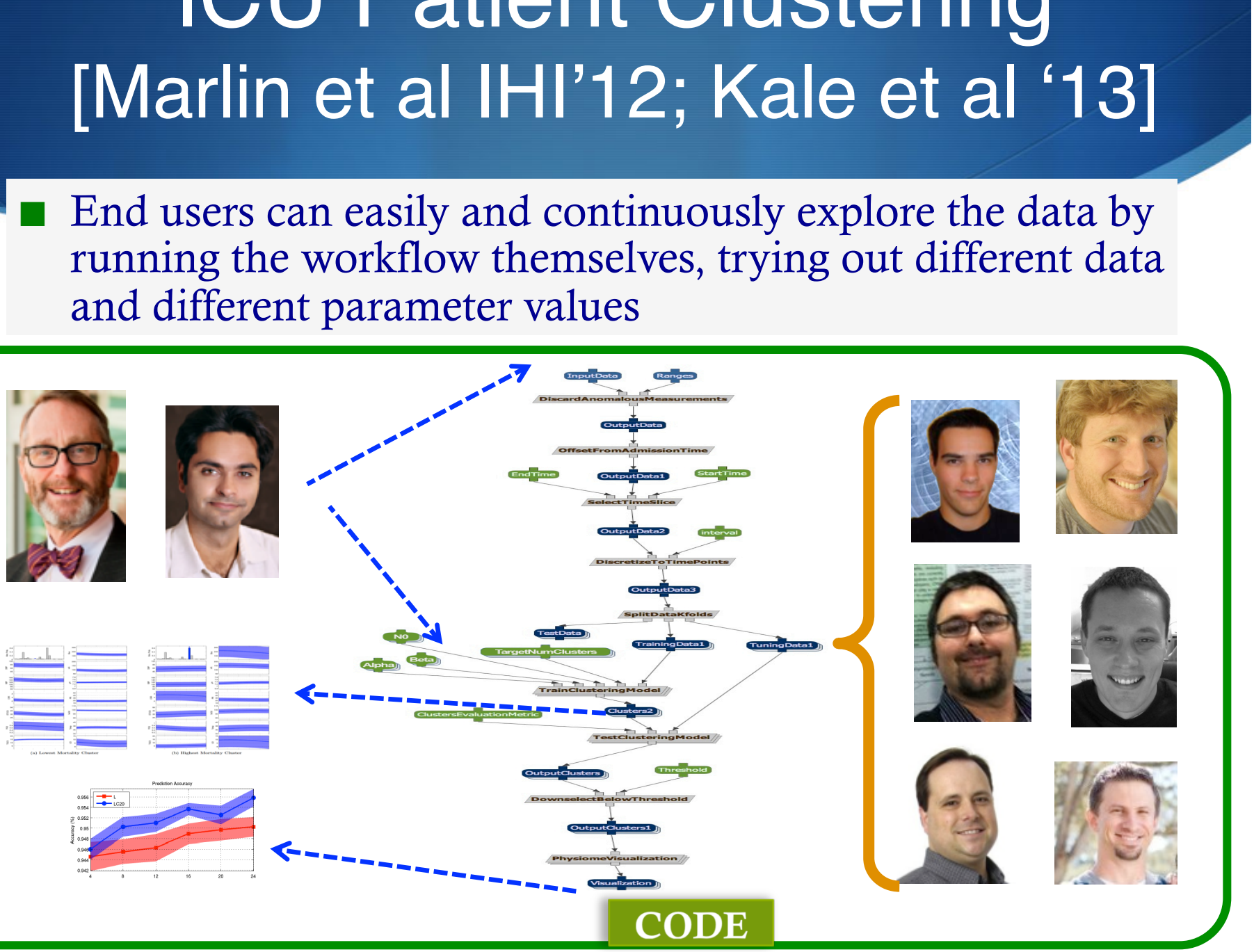




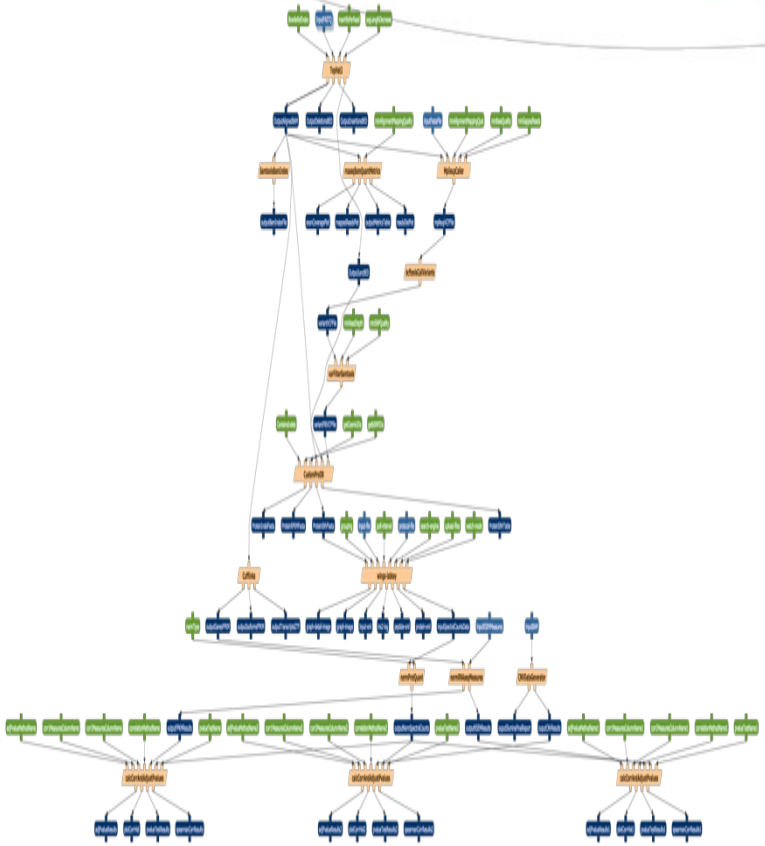
CODE

ICU Patient Clustering [Marlin et al IHI'12; Kale et al '13]

- End users can easily and continuously explore the data by running the workflow themselves, trying out different data and different parameter values



Benefits of Using Workflows



- Simple programming paradigm
- Modular assembly
- Facilitating communication across data science expertise areas
- Composing heterogeneous code
- Data preparation steps
- Data visualization steps
- Provenance and reproducibility
- Large-scale processing

Topics

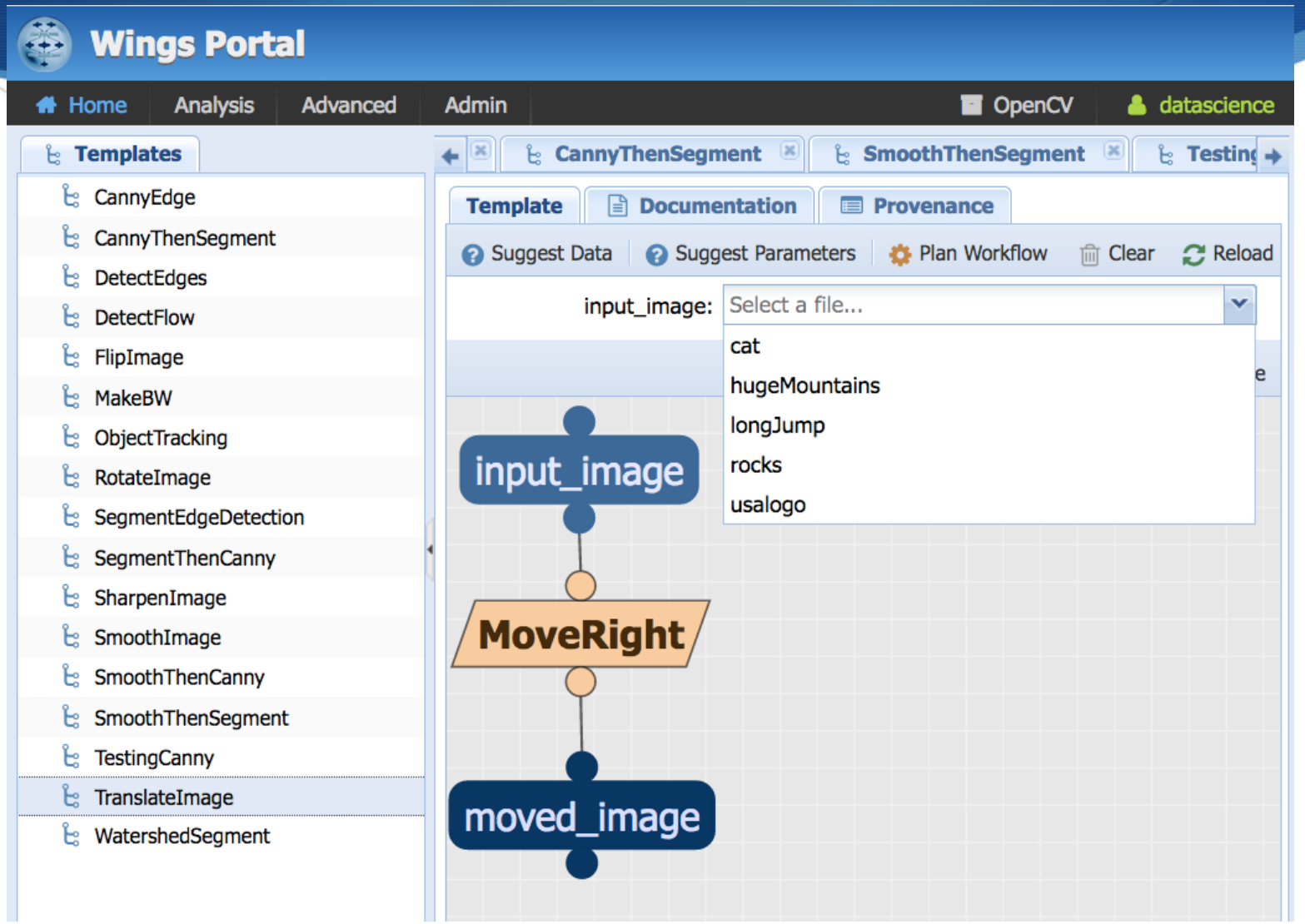
1. Computational workflows
2. Benefits of using workflows
3. Workflow systems
4. Semantic workflows

Workflow Systems

- Many choices
  - Academic prototypes
  - Operational open source
  - Commercial
- Each has different capabilities
  - Scalable computations
  - Domain components
  - Data visualizations

WINGS

http://www.wings-workflows.org

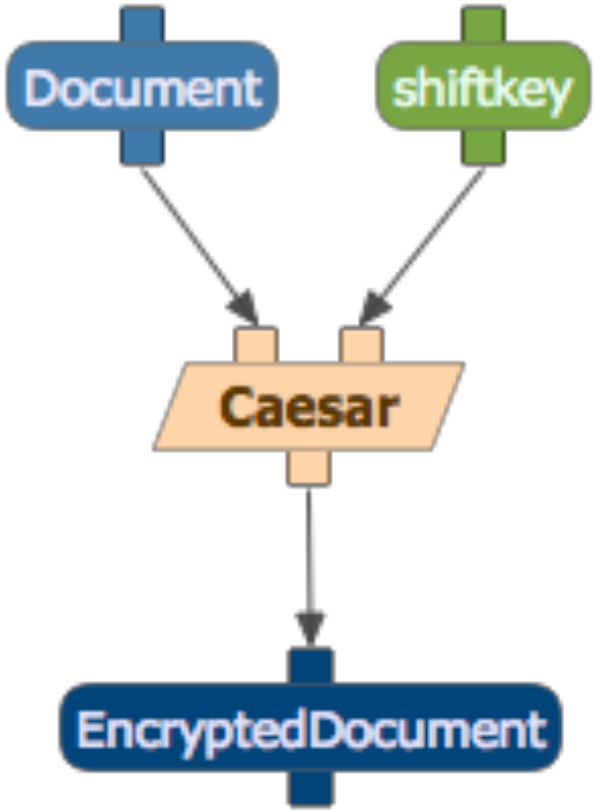


Topics

1. Computational workflows
2. Benefits of using workflows
3. Workflow systems
4. Semantic workflows

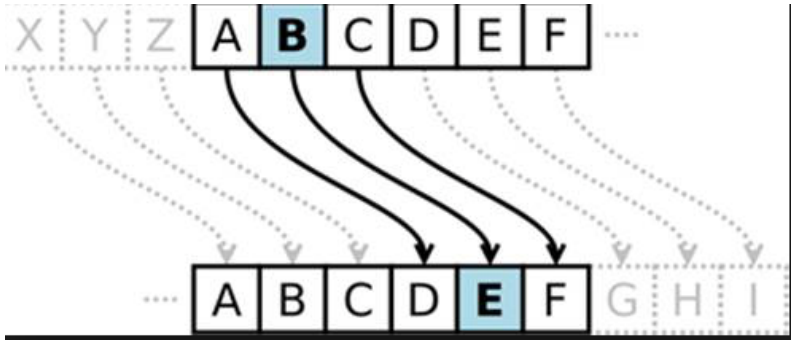
Workflows Treat Software Components as "Black Boxes"

- You **don't have to understand** the inner workings in order to **use the component**
- This is why we often refer to software as a **"black box"**
- You do need to understand **inputs/outputs/parameters** and the **program's function**

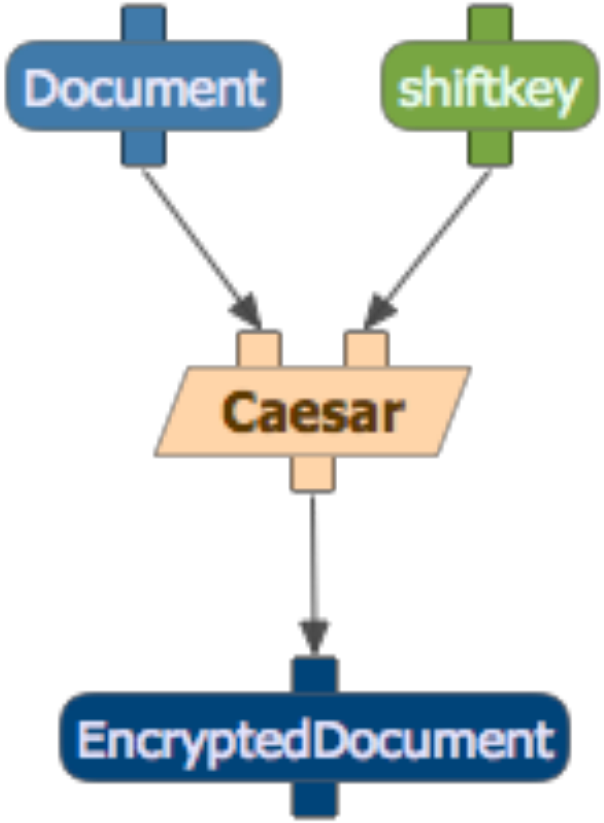


Sometimes There Are Important Constraints

Original: BAD



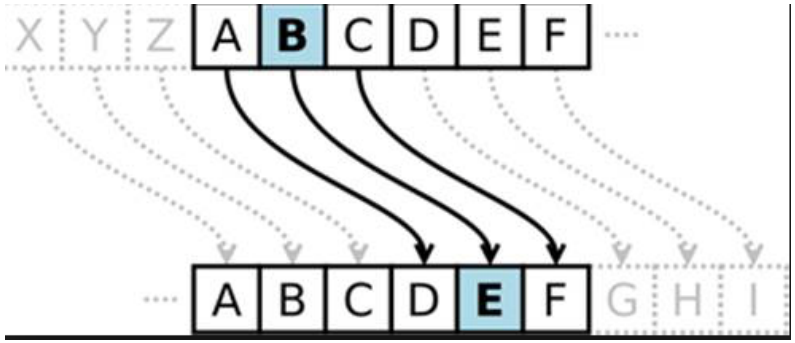
Cipher: EDG

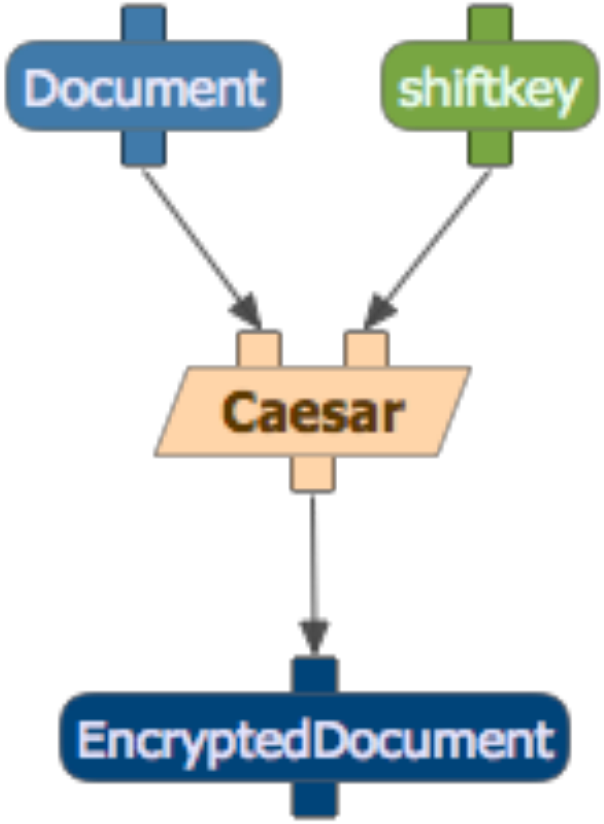


Sometimes There Are Important Constraints

Shiftkey should not be 26

Original: BAD



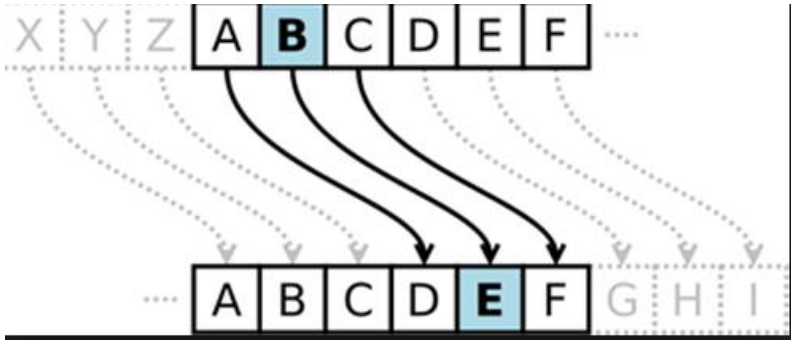


Cipher: EDG

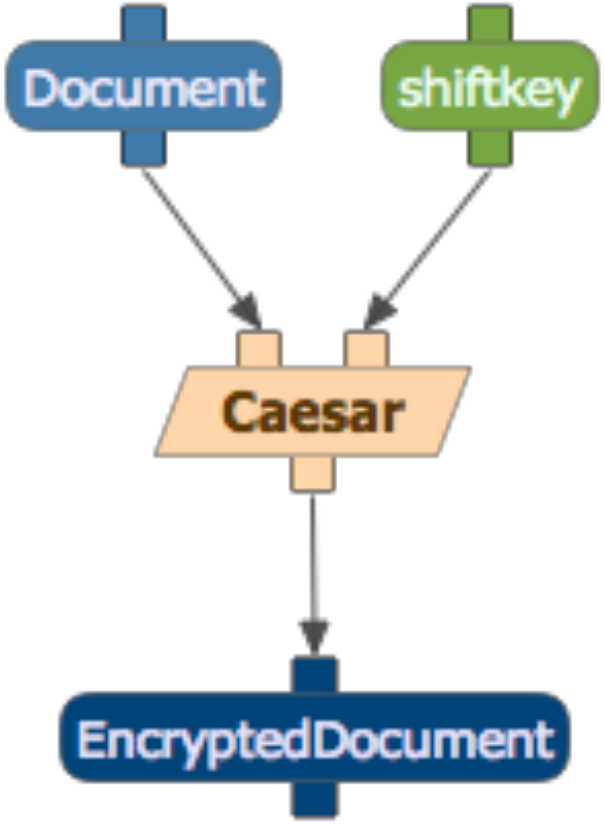
Sometimes There Are Important Constraints

Shiftkey should not be 26, or 0

Original: BAD



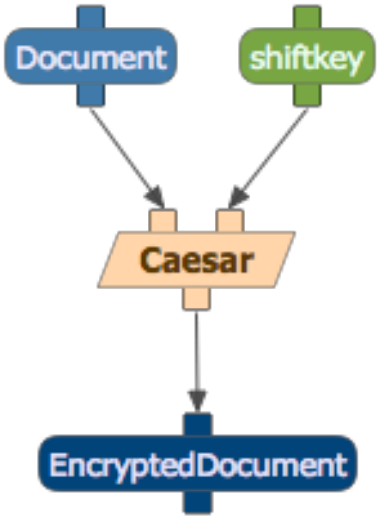
Cipher: EDG



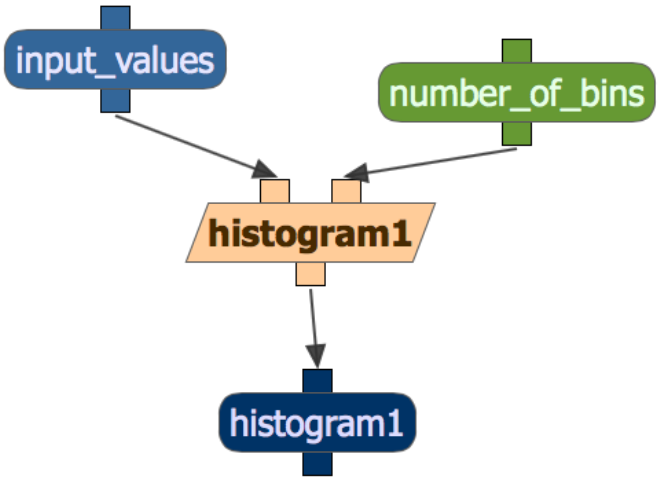
Semantic Components

- Semantic components include semantic constraints that express logic rules about their input or output data, parameters, or other uses of the underlying code

Validation rule: Shiftkey cannot be zero



Parameter setting rule: Sturges' rule suggests that number_of_bins be set to $\log_2$(number of data points) +1



Semantic Constraints To Set Up Parameter Values

Fall-2016-Courses

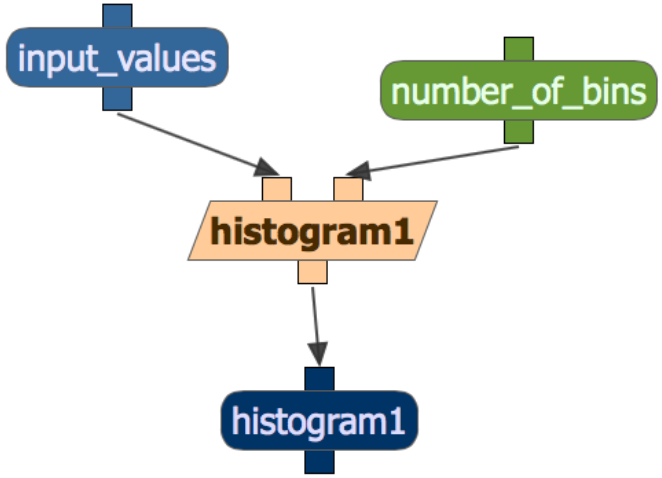
Metadata

Number of datapoints: 479
Semester: Fall
Year: 2016

Data

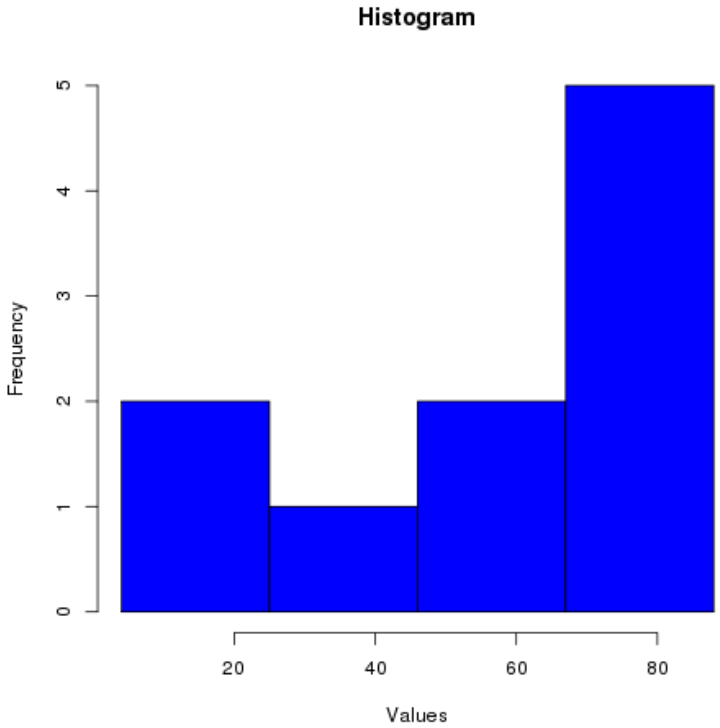
Course ID	Enrollment	Instructor	Open
549	20	A. Gold	Yes
533	50	D. Garcia	Open
556	25	P. Peters	Y
521	100	J. Smith	Open

Sturges' rule:
bins = $\log_2$(number of data points) +1

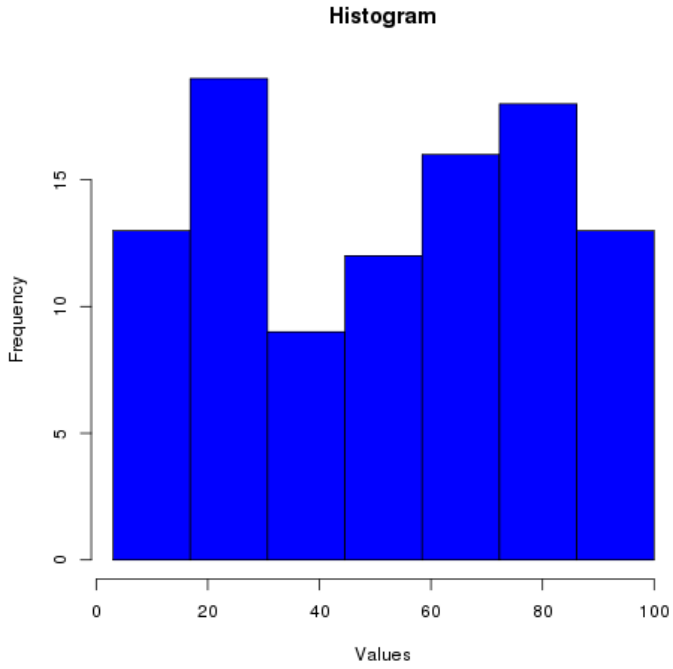


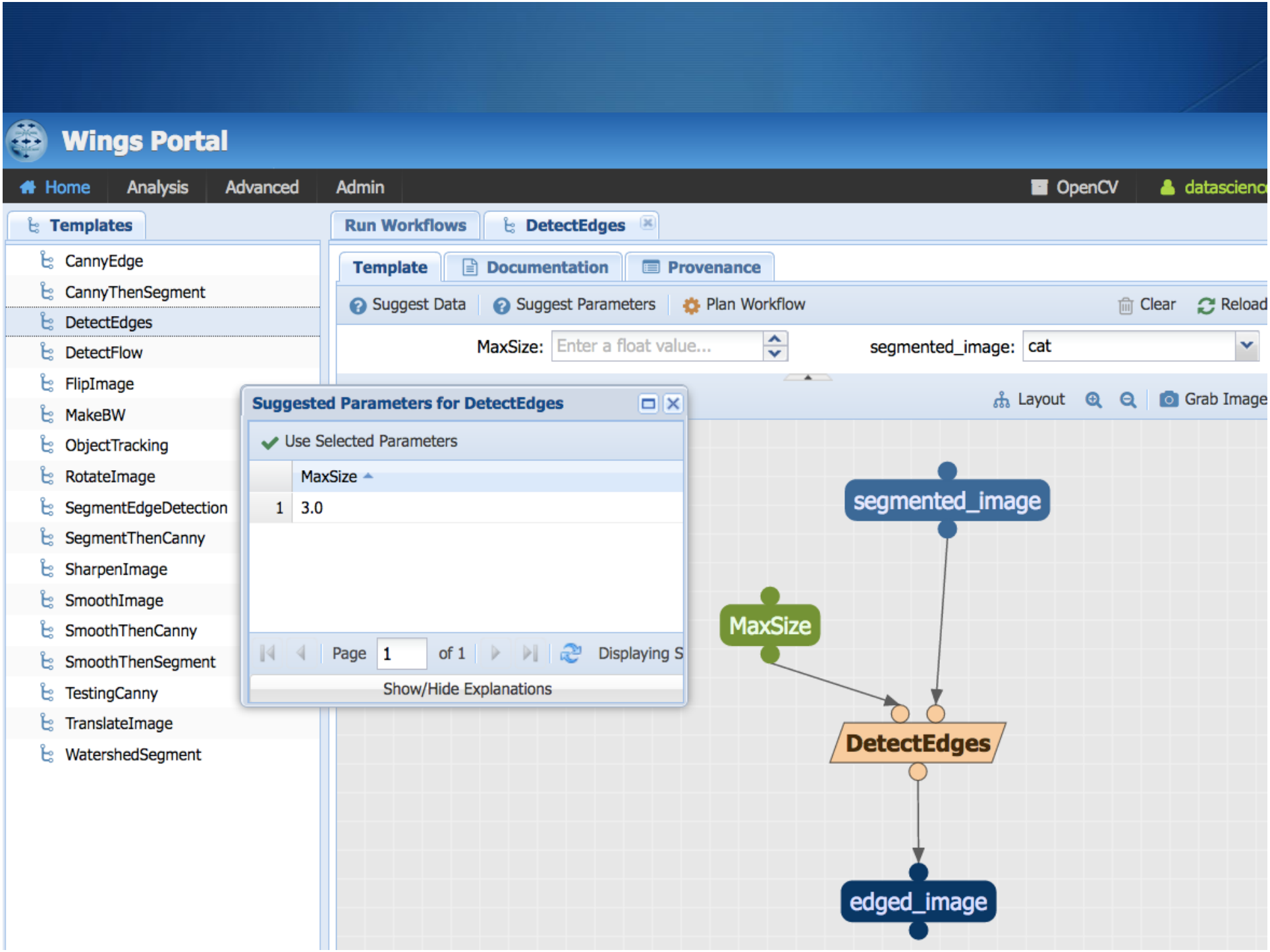
WINGS Customizes Workflows through Semantic Constraints

The data file for this histogram has 10 data points, so WINGS automatically proposes 4 bins:

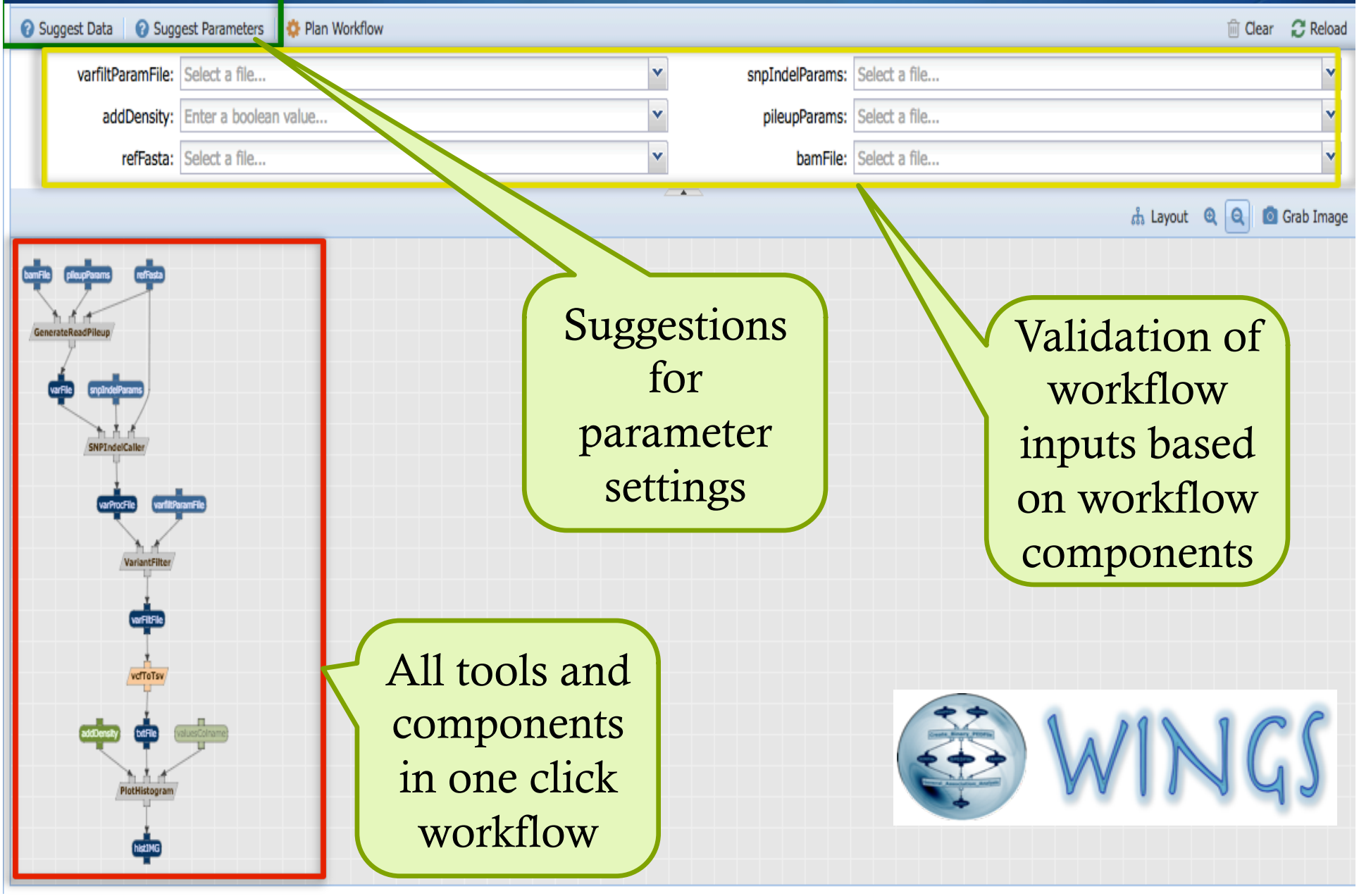


The data file for this histogram has 100 data points, so WINGS automatically proposes 7 bins:




Wings Portal
Home
Analysis
Advanced
Admin
OpenCV
Templates
CannyEdge
CannyThenSegment
DetectEdges
DetectFlow
FlipImage
MakeBW
ObjectTracking
RotateImage
SegmentEdgeDetection
SegmentThenCanny
SharpenImage
SmoothImage
SmoothThenCanny
SmoothThenSegment
TestingCanny
TranslateImage
WatershedSegment
Run Workflows
DetectEdges
Template
Documentation
Provenance
Suggest Data
Suggest Parameters
Plan Workflow
Clear
Reload
MaxSize:
Enter a float value...
segmented_image:
cat
Suggested Parameters for DetectEdges
Use Selected Parameters
MaxSize
1
3.0
Page
1
of 1
Show/Hide Explanations
Layout
Grab Image
segmented_image
MaxSize
DetectEdges
edged_image

Summary: Semantic Workflows



Conclusions

- Workflows are complex compositions of advanced data analysis software
- Workflows make it easy for domain experts to use advanced analytic methods
- Many workflow systems available
- WINGS has additional capabilities to enforce important semantic constraints

http://www.wings-workflows.org