CA16212- INDEPTH Kickoff Meeting, Clermont-Ferrand, March 12-14th, 2018

WrapScienceJ : an Integrative Multipurpose Platform for Data Analysis on a Distributed Infrastructure

Rémy Malgouyres
Université Clermont Auvergne, *LIMOS*, France

remy.malgouyres@uca.fr
http://malgouyres.org
http://malgouyres.org/wrapScienceJ

INDEPTH Open Data and Protocol Comparison Mission

Huge amounts of Data **Distributed on Heterogeneous Storage**

Possibility to Query and **Correlate any Subset** of the Data

---

Specification of **What is an Acquisition Protocol** for Comparison

My Server Can Make Coffee Approach

- Configure One Server for Many Tasks
- Combines Data Stored on One/Few Server(s)
- Infrastructure Doesn't Scale
- Software Maintenance Doesn't Scale

Service Oriented Architecture *SOA* Approach

- Each Service Makes only One Task
- Services Deploy Independently (Light Server Configuration)
- Infrastructure Scales Conveniently at Low Cost
- Software Maintenance Does Scale at Low Cost

Who Uses Which Approach in the *IT* World ?

My Server Can Make Coffee Approach

- Small Size Businesses
- Secondary Sector
- High Tech *Labs* in Universities

Service Oriented Architecture *SOA* Approach

- The *GAFA*'s
- Multinationals, Airline/Railways Companies
- *Uber*, *BlaBlaCar*, *Airbnb*...

Why is That ?

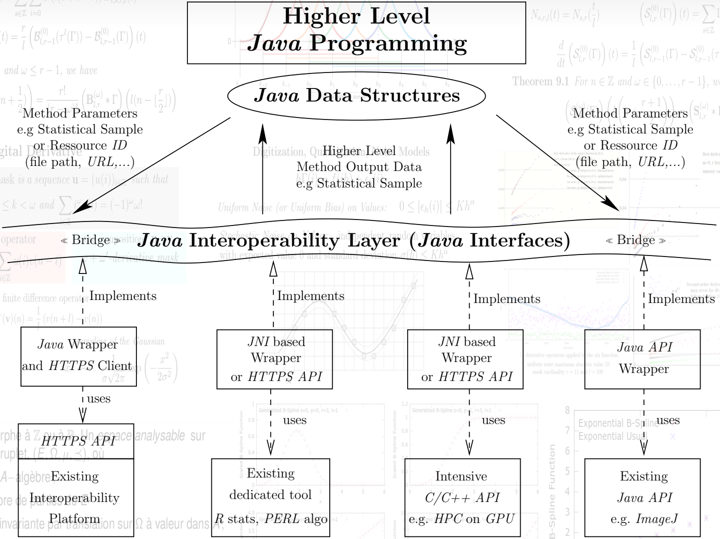
My Server Can Make Coffee Approach

- We don't have the Knowhow
- We can afford to waste
- We *Publish or Perish*

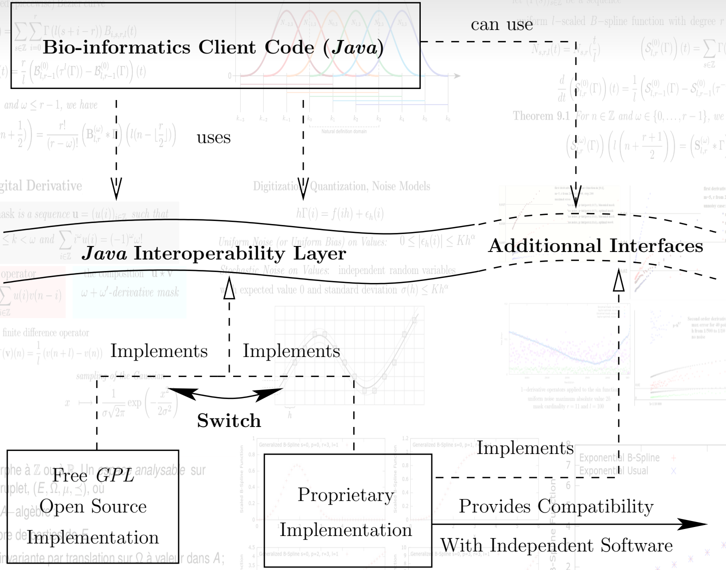
Service Oriented Architecture *SOA* Approach

- One Billion *R&D* isn't that much and it's money well spent
- We *Bring Down the Costs or Perish*

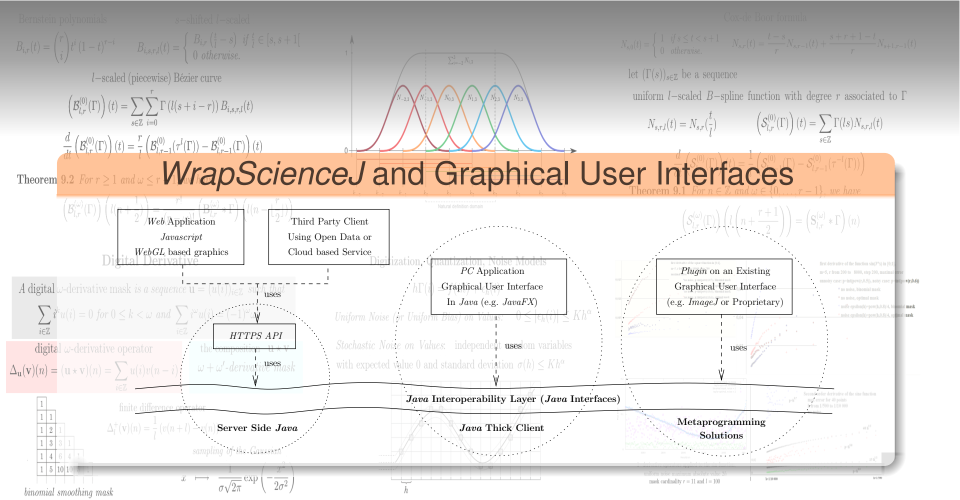
Overview of the *WrapScienceJ* Inversion of Control Structure



WrapScienceJ Honey Pot Function

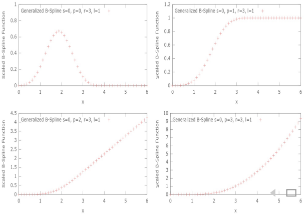


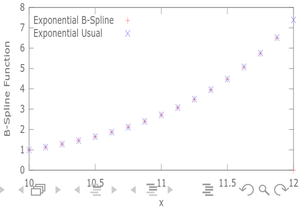
WrapScienceJ and Graphical User Interfaces



A est un anneau isomorphe à $\mathbb{Z}$ ou à $\mathbb{R}$. Un *espace analysable* sur $(A,+,.,\preceq)$ est un quadruplet, $(E,\Omega,\mu,\preceq)$, où

1. $(E,+,.,.)$ est une A−algèbre ;
2. Ω est une σ−algèbre de parties de E
3. μ est une mesure invariante par translation sur Ω à valeur dans A ;
4. et $\preceq$ est un ordre complet, compatible avec la multiplication externe.

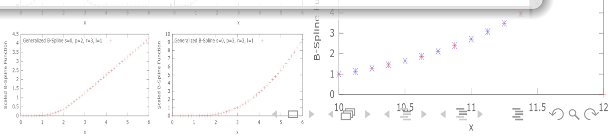




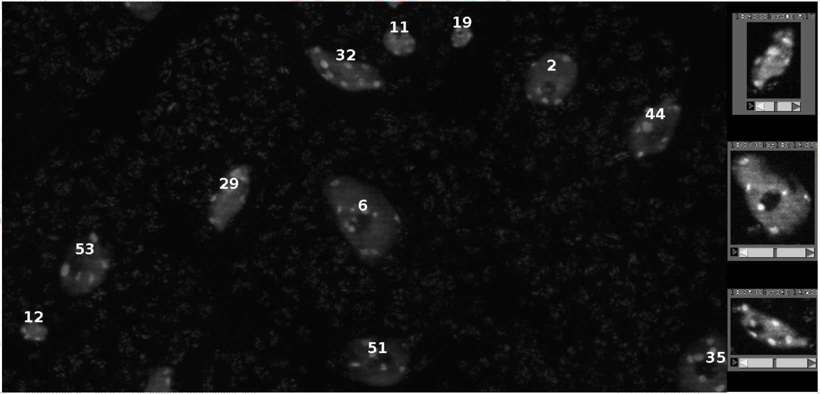
WrapScience API Style and Design Patterns

```
VoxelDouble voxelEdgesLength = image.setTitle("Original Image")
                                    .retrieveMetaData()
                                    .getImageCalibration()
                                    .getVoxelLength();

ImageCore binarizedImage = image.duplicate()
                                .getImageBlur()
                                .getGaussianBlurCalibrated(0.6, 0.6, 0.6, voxelEdgesLength)
                                .getImageConvolved(ConvolutionNormalizationPolicy
                                                   .Gray8 Scale MaximizeContrast)
                                .setTitle("Thresholded Image (" + ThresholdingOption.Otsu + " Method)")
                                .getImageThresholding()
                                .thresholdImageAndBinarize(ThresholdingOption.Otsu, true);

ImageCore labeledImage = binarizedImage.duplicate()
                                       .setTitle("Labeled Image(" + ThresholdingOption.Otsu + ")")
                                       .getImageConnectedComponents()
                                       .getLabeledComponents(
                                           labelingPolicy, // Full 3D or Slice by slice 2D
                                           white, // Foreground
                                           false, // Remove components on the border
                                           componentVolumeThreshold, // Lower threshold on components v
                                           true // Set a uniform random color on each component
                                       )
                                       .getImage();

RenderToolFactoryIJ3D.getInstance().getRenderTool().display(labeledImage);
binarizedImage.getPreferedRenderTool().display(binarizedImage);
GlobalOptions.getDefaultRenderTool().display(image);
```

2. Ω est une σ−algèbre de parties de E
3. μ est une mesure invariante par translation sur Ω à valeur dans A;
4. et $\preceq$ est un ordre complet, compatible avec la multiplication externe.

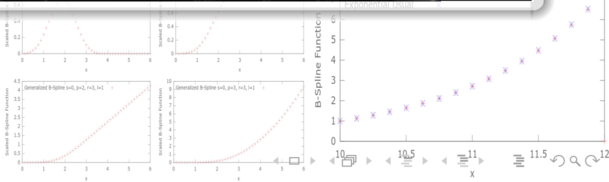


WrapScienceJ Application : Segmentation and Auto Crop

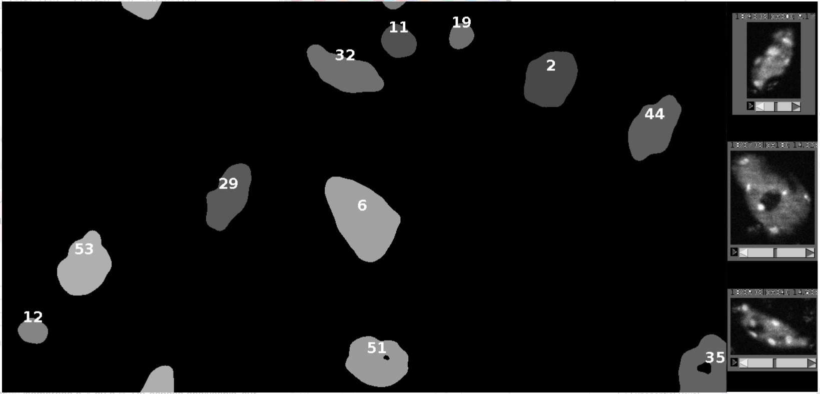


A est un anneau isomorphe à $\mathbb{Z}$ ou à $\mathbb{R}$. Un *espace analysable* sur $(A, +, ., \preceq)$ est un quadruplet, $(E, \Omega, \mu, \preceq)$, où

1. $(E, +, ., .)$ est une $A-$algèbre ;
2. Ω est une $\sigma-$algèbre de parties de E
3. μ est une mesure invariante par translation sur Ω à valeur dans A ;
4. et $\preceq$ est un ordre complet, compatible avec la multiplication externe.

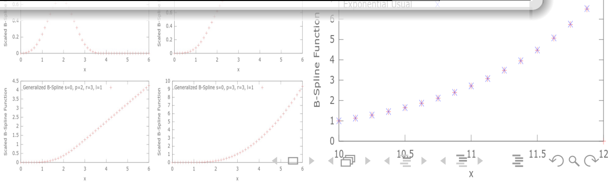


WrapScienceJ Application : Segmentation and Auto Crop



A est un anneau isomorphe à $\mathbb{Z}$ ou à $\mathbb{R}$. Un *espace analysable* sur $(A, +, ., \preceq)$ est un quadruplet, $(E, \Omega, \mu, \preceq)$, où

1. $(E, +, ., .)$ est une A–algèbre ;
2. Ω est une σ–algèbre de parties de E
3. μ est une mesure invariante par translation sur Ω à valeur dans A ;
4. et $\preceq$ est un ordre complet, compatible avec la multiplication externe.



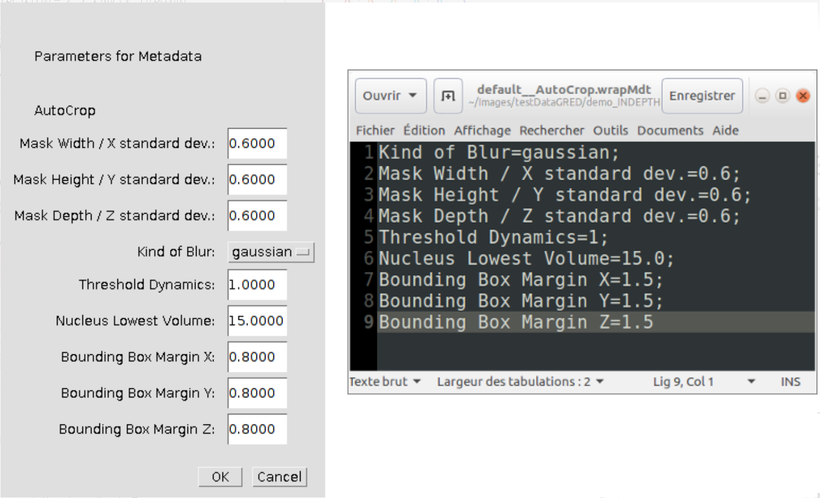
WrapScienceJ Workflow : For the Bioinformatician (1/2)

```java
 */
public class ProcessMetaData extends MetaDataSet {

	/** Kind of blurring mask ("binomial or gaussian) */
	private ChoiceInListSetSingle m metaKindOfBlur;

	/** Smoothing mask Width (for binomial blur) or standard deviation (gaussian blur) */
	private DoubleSetSingle m metaSigmaX; 

	/** Smoothing mask Height (for binomial blur) or standard deviation (gaussian blur) */
	private DoubleSetSingle m metaSigmaY; 

	/** Smoothing mask Depth (for binomial blur) or standard deviation (gaussian blur) */
	private DoubleSetSingle m metaSigmaZ; 

	/** Thresholding is made more dynamic by elevating the values to some exponent
	 * 3 >= typically >= 1
	 */
	private DoubleSetSingle m thresholdDynamics; 

	/** Nucleus Lowest Volume Threshold */
	private DoubleSetSingle m nucleusVolumeThreshold; 

	/** Bounding box's margin on first coordinate for cropping */
	private DoubleSetSingle m xMarginReal;

	/** Bounding box's margin on second coordinate for cropping */
	private DoubleSetSingle m yMarginReal;

	/** Bounding box's margin on third coordinate for cropping */
	private DoubleSetSingle m zMarginReal;
```

WrapScienceJ Workflow : For the Bioinformatician (2/2)

```java
 */
public class Nuclei Auto Crop extends PluginFilterGenericIJ {
    /**
     * @see PluginFilterGenericIJ#getProcess(ImageCore)
     */
    @Override
    public GenericImageProcessConcrete getProcess(ImageCore image) {
        return new AutoCropSegmentProcess(
                                image,
                                "AutoCrop",
                                RetrievalPolicy.ForceDialog,
                                "wrapProcess"+File.separator+"predefined"+File.separator,
                                OutputDataKind.CreatedFromInputCopy,
                                RenderToolFactoryIJ3D.getInstance().getRenderTool(),
                                GlobalOptions.getDefaultGuiFramework());
    }

    /**
     * Allows to test the plugin
     * @param args
     */
    public static void main(String[] args) {

        try {
            Nuclei Auto Crop plugin = new Nuclei Auto Crop();
            GenericImageProcessConcrete process = plugin.getProcess(
                ImageCoreFactoryIJ.getInstance().getImageCore(
                    TestImageThresholding.getSampleImageGray8(5)
                )
            );

            process.testPlugin(process.getRenderTool(), "ConnectedComponents3D", null);
        } catch (IOException e) {
            e.printStackTrace();
        }
    }
```

WrapScienceJ Workflow : For the Biologist (1/2)

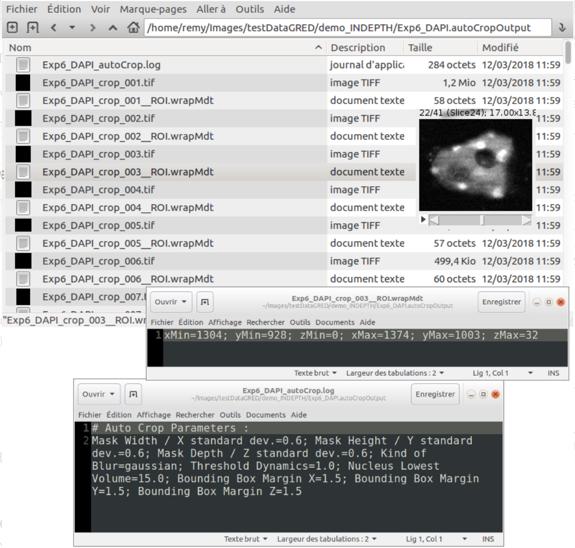


❷ Ω est une σ−algèbre de parties de E

❸ μ est une mesure invariante par translation sur Ω à valeur dans A;

❹ et $\preceq$ est un ordre complet, compatible avec la multiplication externe.

WrapScienceJ Workflow : For the Biologist (2/2)



Questions about *WrapScienceJ*?