# PCM WIKIPEDIA

PROJET DE DEVELOPPEMENT LOGICIEL

### **DOCUMENTATION**

Alexandre Berteaux Baptiste Quéré Hugo Vallée Yoann Le Taillanter Thibault Le Roux M1 Miage ISTIC – Université Rennes 1 2014 - 2015

## INSTALLATION PROCESS AND IMPLEMENTATION OF THE PCM PROJECT

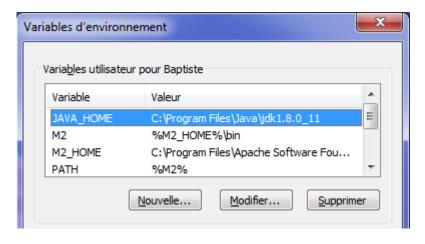
#### > Tools and techniques used:

- IntelliJ IDEA
- Github
- Maven

#### > First step: Installation of Maven

- 1. Download and extract the last version of Maven. Move it to the directory you wish to install Maven. For example: C:\Program Files\Apache Software Foundation.
- 2. Go to the "Environment Variable" in System Properties/Advanced and add the M2\_HOME environment variable with the value of the path where you have installed Maven. For example: C:\Program Files\Apache Software Foundation\apache-maven-3.2.3
- 3. Then, add a new environment variable called M2 with the value %M2\_HOME%\bin
- 4. Make sure that the JAVA\_HOME environment variable is present and contains the value of the path of your JDK. For example: C:\Program Files\Java\jdk1.8.0\_11. If not, create it.
- 5. Update or Create the PATH environment variable with the value *%M2%*
- 6. Verify that Maven is well configured by opening your command terminal and enter *mvn* –*version*
- 7. To complete the installation, you must have git installed on your computer

Your environment variables should look like this:

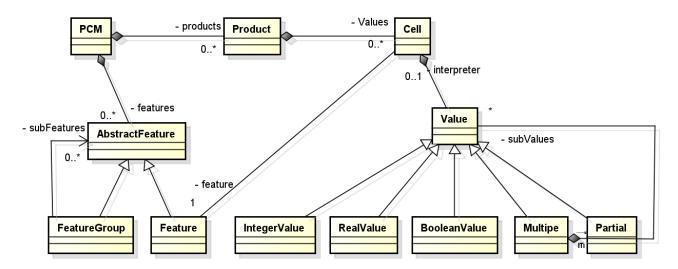


#### > Second step: Implementation of the project

- 1. Open your IntelliJ Idea
- 2. clone https://github.com/gbecan/PCM.git
- 3. clone http://github.com/dukeboard/Kevoree-modeling-framework.git
- 4. cd Kevoree-modeling-framework
- 5. git checkout a758cf07ded04e3dfc44b26bca7fcff8426d5495
- 6. mvn clean install
- 7. open your project PCM
- 8. mvn clean install You're going to have a build failure but it is a normal issue if the fail concerns the org.diverse.PCM.play-app.
- 9. If you want to install "play", download it.

#### **CONCEPTION**

#### Class Diagram



#### STEPS AND PROCESSING

In this part, we will explain how the PCM project works with Wikipedia.

To begin, let's do a quick view of Wikipedia PCM system.

I give your attention on the red box, we can see a small picture of the California flag next to California name. If we click on « edit », we will be able to see how this part is represented.



Still in the red box, you can see { {Flag | California} }. This is a template, a useful but not friendly function of Wikipedia to replace an instance by a smallest code. This is what will be transformed in his real value when our PCM project will preprocess the file. Let's explain this.

#### Editing Comparison between U.S. states and countries by GDP (nominal) (section)

```
B I № № № Advanced > Special characters > Help > Cite
 =State and country data=
{{col-begin}}
{{col-break|width=30%}}
{| class="wikitable sortable"
|+ State GDPs, 2012<ref name="bea">{{cite press release | url=http://www.bea.gov/newsreleases/regional/gdp state/2012/pdf/gsp0612.pdf |
title=WIDESPREAD ECONOMIC GROWTH ACROSS STATES IN 2011 | publisher=Bureau of Economic Analysis | date=June 5, 2012 | accessdate=December
27, 2012 | format=PDF | page=7}}</ref>
! Rank !! State !! GDP (In Millions) ([[USD|M$USD]])
      {{flag|California}} || 1,958,904
| 2|| {{flag|Texas}} || 1,308,132
| 3|| {{flag|New York}} || 1,157,969
| 4|| {{flag|Florida}} || 843,091
| 5|| {{flag|Illinois}} || 751,094
| 6|| {{flag|Pennsylvania}} || 647,443
| 7|| {{flag|Ohio}} || 551,881
```

The program starts with getting a page from Wikipedia, with the getPageCodeFromWikipedia function. It takes the article title as an argument and create the URL link. Then it goes in the edit function, to parse the code as an XML file.

The second step is the preprocess function, it will parse the code to replace all of the templates by its real value.

The third step is the parse function, where the Product Comparison Matrice will be normalized.

That means every multiple column or multiple line will be divided in few simples.

To finish, we can export the new PCM in HTML, CSV or PCM.

The diagram below summarizes what we have just described.

