
Business Process Management in the Annotation Factory

Milan Agatonovic

University of Sheffield

D-Spin – CLARIN Workshop

Munich 11.11.2008



Content

- About GATE Teamware?
- Why and how we use JBPM?
- Challenges and our solution



GATE Teamware - overview

What?

GATE Teamware is a software suite and a methodology for the implementation and support of annotation factories.

Why?

It is intended to provide a framework for annotation services, supplied either as in-house units or as outsourced specialist activities



Or technically speaking ...

GATE Teamware:

- uses the following libraries:

Spring + hibernate + struts + gate + jbpm + CXF

- lives under:

Tomcat + Apache

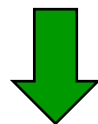
- and has:

mysql as the DB

GATE Teamware – Features

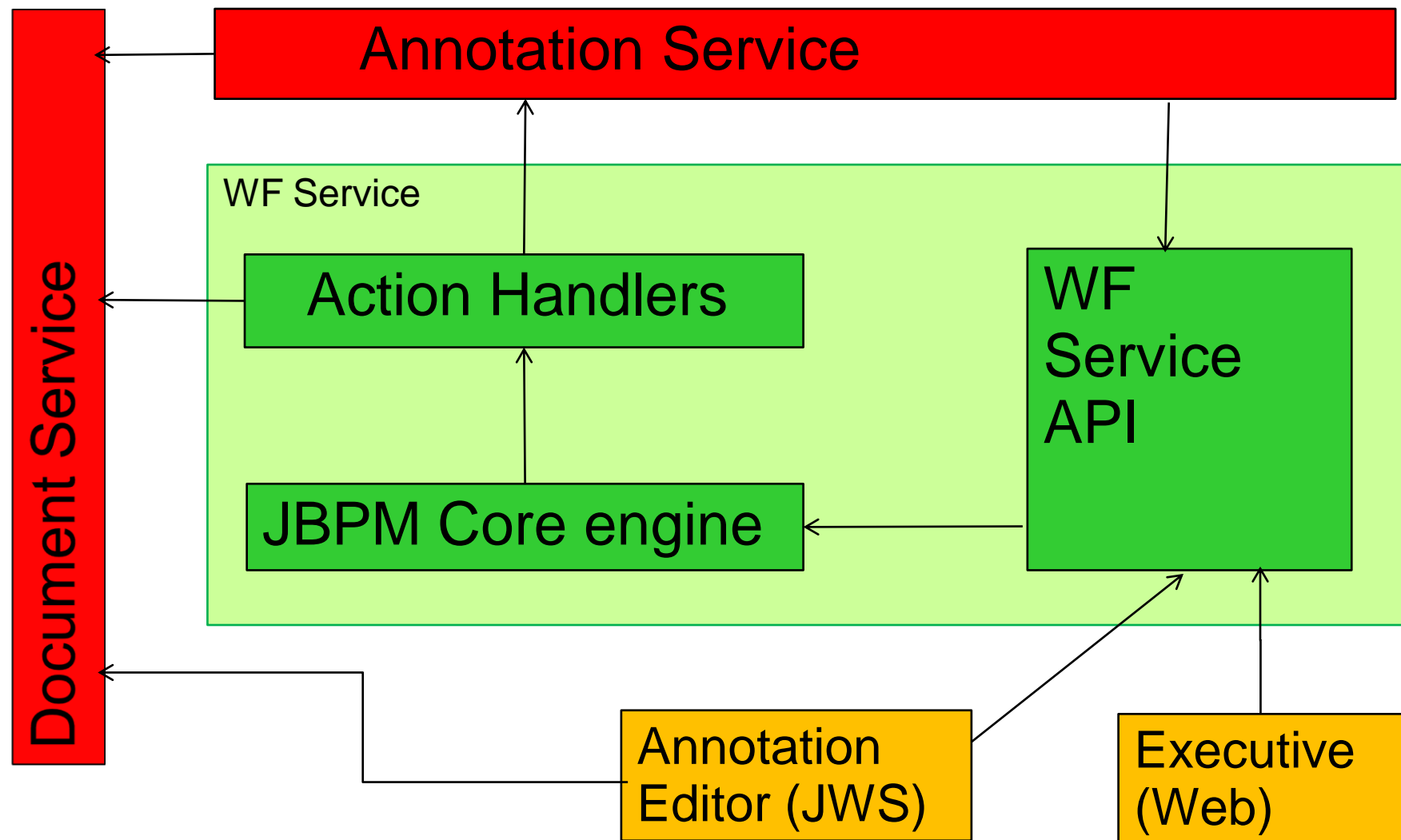
GATE

- Multiple roles (administrators, curators, engineers, annotators)
- Sync & Async service invocation
- Human tasks (annotate, review, etc)
- Complete annotation lifecycle support



Yes, we need a BPM engine!

How it looks like in real life?





Content

- About GATE Teamware?
- Why and how we use JBPM?
- Challenges and our solution

WF Engine requirements

- The support for human tasks ©
- The support for long wait states ©
- Graphical modelling tool ©
- User interface forms generation ©
- Calling services in async & sync way ©
- Open Source

* © denotes CLARIN requirement



Why JBPM and JPDLL?

- Flexible identity component
- Graph based language
- Well known vendor (good experience)
- Convenient for Java developers
- Perfect marriage with Spring

and

Embeddable & Extensible !



Embeddable WF engine

The crucial aspects:

- Persistence
- Services
- Transactions
- Library
- Testability

Extensible?

We have:

- our identity component
- custom task form generation
- JBPM task -> annotation task
- our additional DB queries
- **custom nodes and structures**

What is a “Custom Node”?

Take one of the existing nodes:

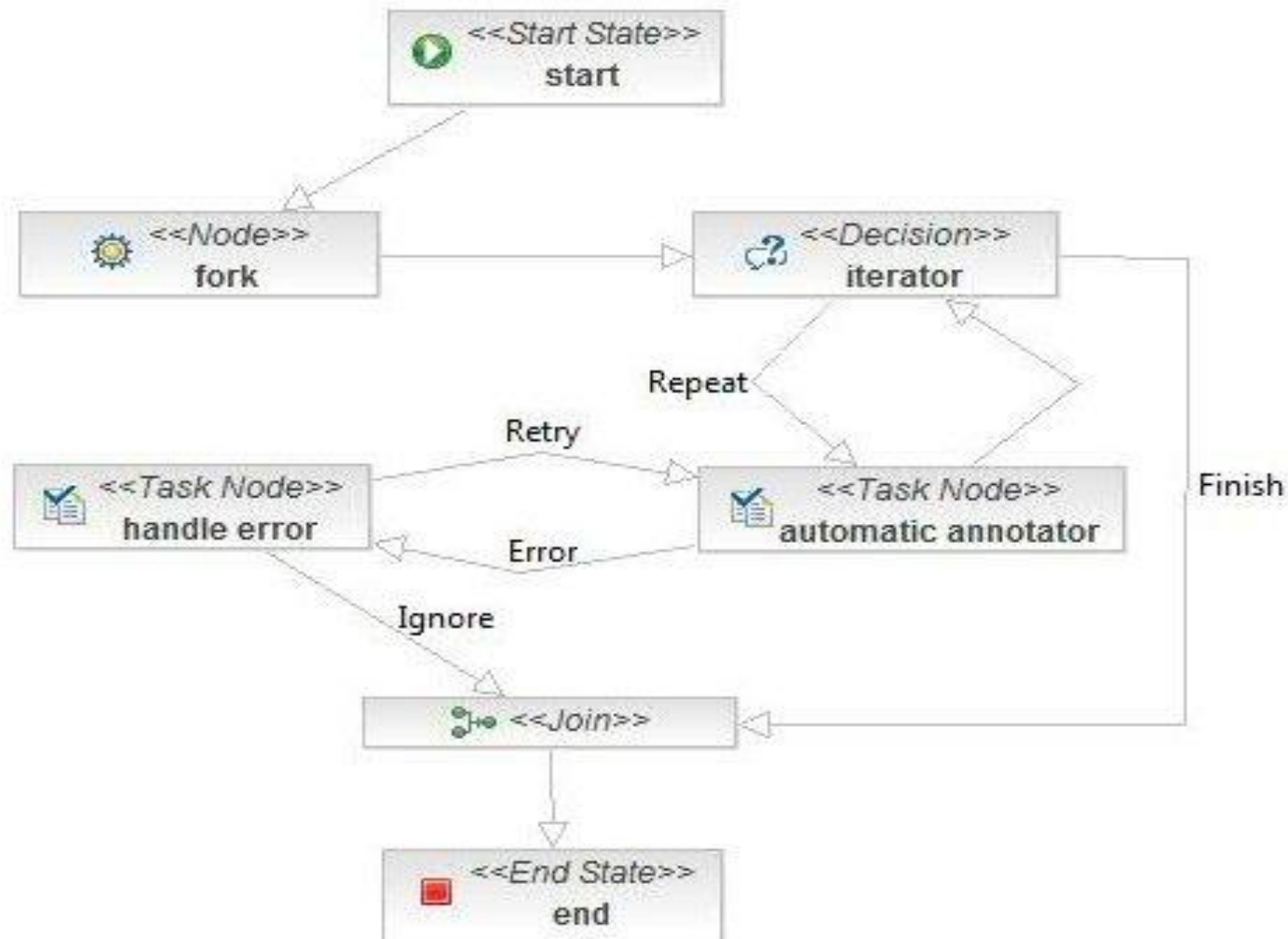
Fork, Join, Task Node, Node, State and:

- Write Action Handler
- Specify its input and output variables
- Put it on some node event
- **That's it. You have a new node.**

Our Custom JBPM Nodes

- Dynamic Fork Node
- Manual Annotation Task Node
- Automatic Annotation Task Node
- Error Handling with Retry Structure
- Loop Structure

Automatic Annotation





Content

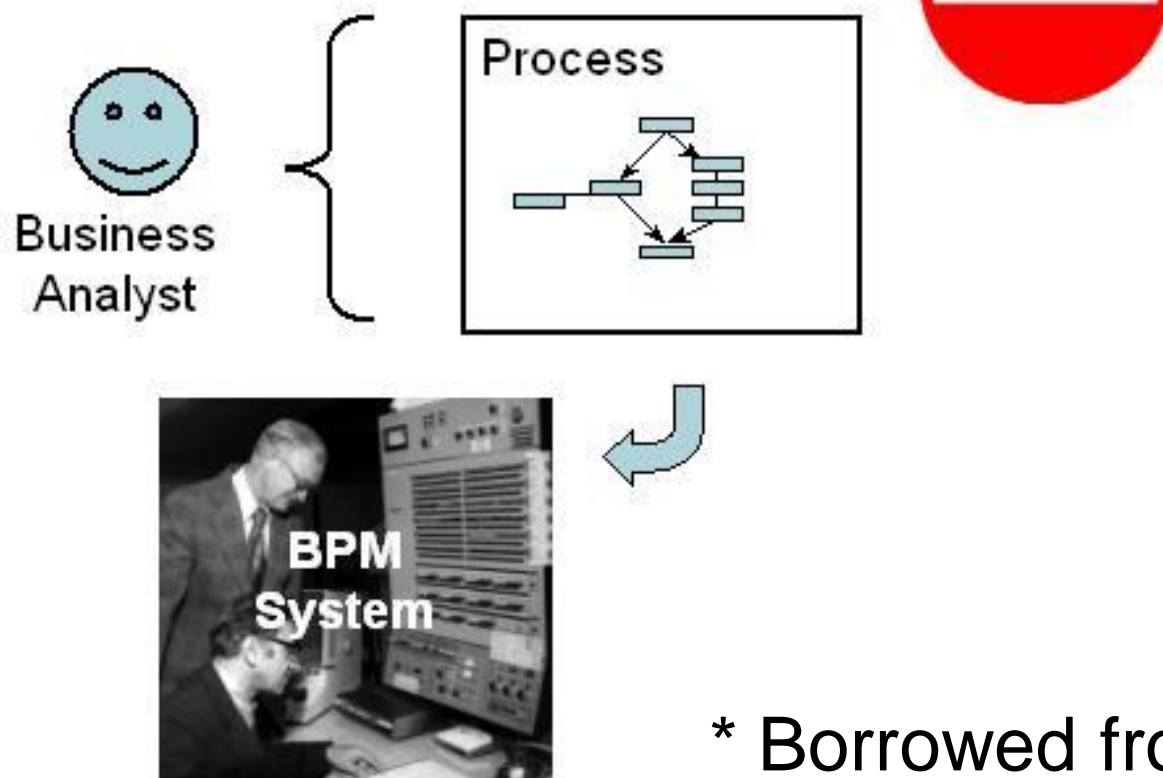
- About GATE Teamware?
- Why and how we use JBPM?
- Challenges and our solution

Usability, Usability ...

The biggest challenge:

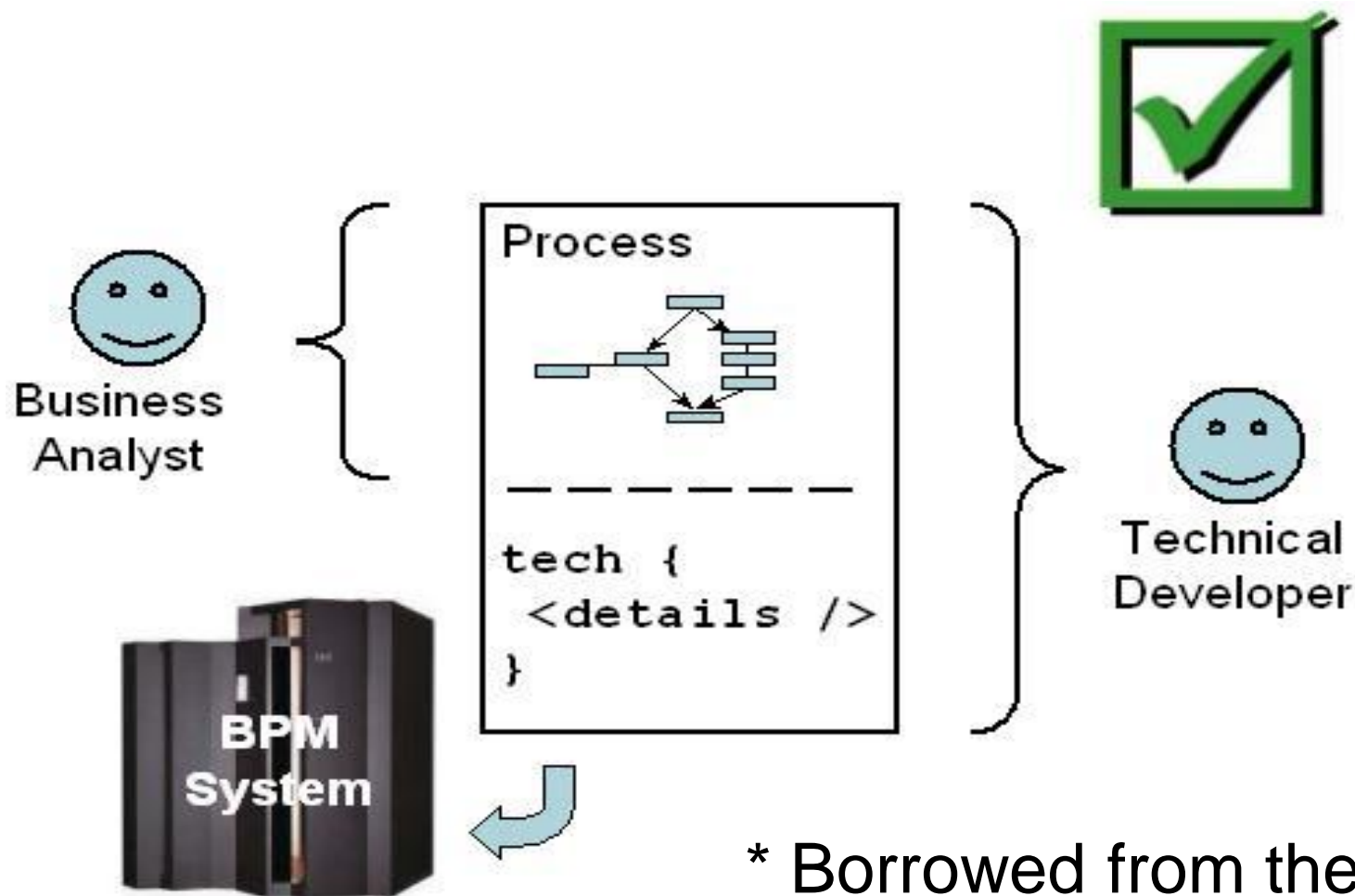
- How to provide workflow modelling tools to a non technical person?
- User wants: “I want to annotate “corpusX”, with “serviceY”, followed by manual annotation with “rulesZ” and review”.

Traditional BPM Approach



* Borrowed from the JBPM developer guide.

JBPM Approach



* Borrowed from the JBPM developer guide.

That's nice, but ...

- We are making tool for researchers.
- They may not have developers around.
- We cannot allow a researcher to execute process directly from the diagram.



Specify Domain Specific Language
(DSL) for annotation factory!



DSL sounds good, but how?

-
- Identify all possible phases (building blocks) in the semantic annotation lifecycle.
 - Define input and output interface for each building block.
 - Examine their deterministic nature (the order of their appearance)
 - Create the master process with all possible blocks and transitions.
 - Generate process composer wizard which (in/ex)cludes blocks and configures them.

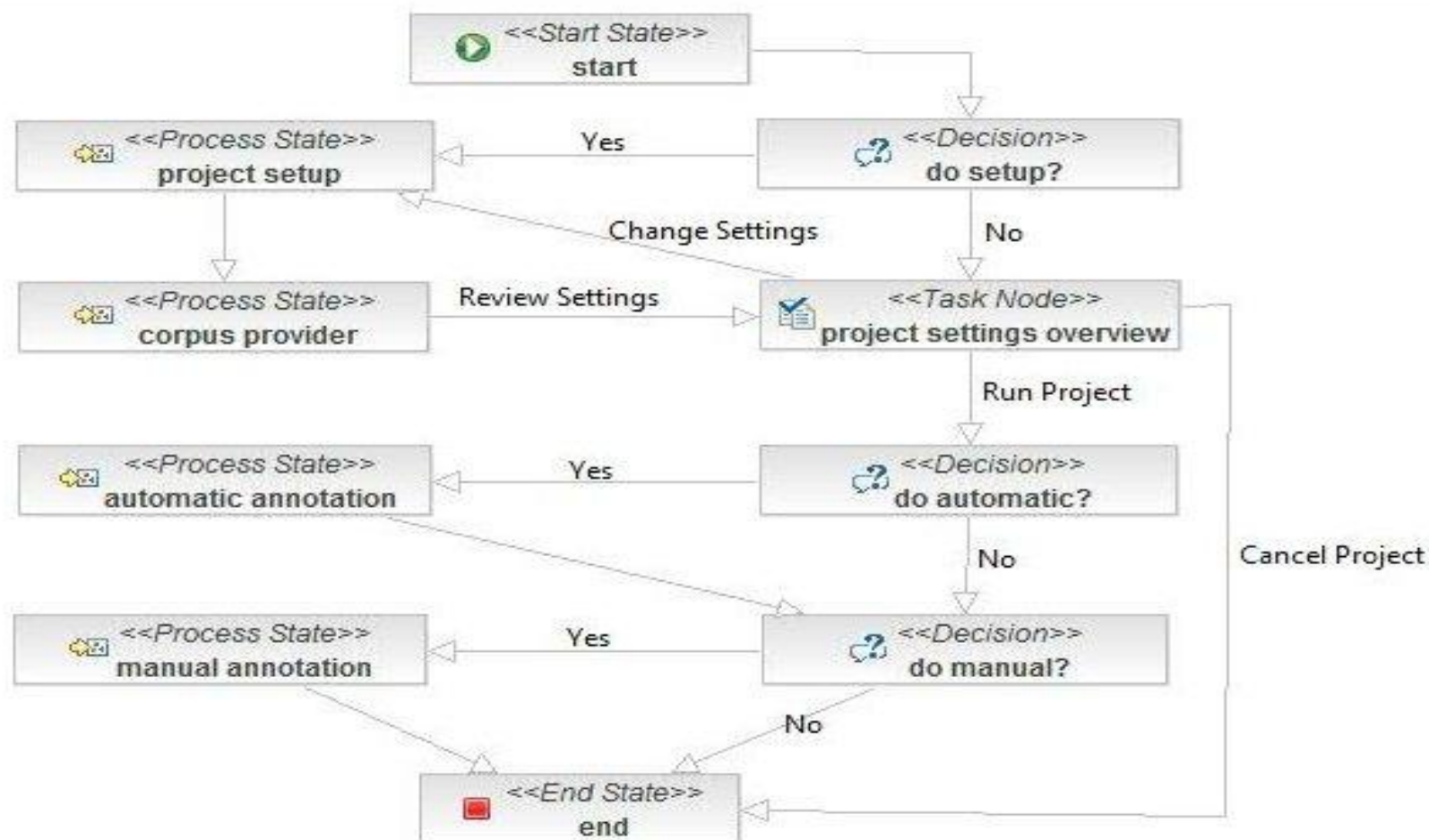
And, we gain...

Modularity: The one Main process (for the domain), but several sub-processes (building blocks), which are included optionally in runtime.

Reusability: All process settings are stored as “project”, which can be later reused


Maintainability: Each sub-process can be separately edited and then deployed. The main process will take its latest version by default.

And in picture





Annotation projects



GATE Teamware

HomeDocumentsSchemasOntologiesWorkflow → Support →

Projects

Create NewBack

One project found.

Name	Description	Updated				
my-anotation-project	automatic + manual annotation	2008-11-05 11:53:09.0				

Export options: CSV | Excel | XML | PDF

Create NewBack

© 2008 GATE | CSS Valid | HELP | Logged in as: Milan Agatonovic

Launch project setup wizard

Starts already configured annotation project

Project Setup Wizard – GAS configuration

GATE

i You have completed 'select annotation modules'!

Configuration Step

gas configuration

Pipelining services

Back Save And Add Another Finish Help

Basic GAS Configuration:

Service Name: *	preprocessing
Service URL: *	http://localhost:8080/matrixware/pr

Advanced GAS Configuration:

Annotation Set Key:	
Annotation Set Value:	
Annotation Set Extra Mappings:	
Parameter Key:	
Parameter Value:	

Service friendly name and URL

Project Setup Wizard – Annotation rules

GATE

i You have completed 'gas configuration'!

Configuration Step

select annotation rules

Necessary
parameters for
manual annotation

Back

Submit

Help

Curators:*

curator
kalina
agaton

Annotators:*

annotator1
annotator2
annotator3

Annotators per Document:

1

Cancel Task Allowed:



Anonymous Annotation:



Annotation Schemas:*

Measurement.xml
Reference.xml
Section.xml

Back

Submit

Help



Project Settings Overview

Project Settings Overview

Back

Run Project

Cancel Project

Change Settings

Help

Automatic Annotation:	<input checked="" type="checkbox"/>
Manual Annotation:	<input checked="" type="checkbox"/>
Corpus:*	ccc ▾
Annotation Schemas:	<div>Measurement.xml</div> <div>Reference.xml</div> <div>Section.xml</div>
Annotators per Document:	2 ▾
Cancel Task Allowed:	<input checked="" type="checkbox"/>
Anonymous Annotation:	<input checked="" type="checkbox"/>
GAS services to be executed:	<div>gas1</div> <div>gas2</div>
Curators:	<div>curator</div> <div>kalina</div> <div>agaton</div>
Annotators:	<div>annotator1</div> <div>annotator2</div> <div>annotator3</div>
Optional Saving Project:	
Project Name:	<div>my-annotation-project</div>
Project Description:	<div>automatic + manual annotation</div>

Shows all options configurable inside wizard with ability to change them, before running the project



References

- GATE (SAFE) website.

<http://gate.ac.uk/safe>

- JBPM Developer Guide

<http://jbpm.org>

- The blog of Tom Baeyens, JBPM lead

<http://processdevelopments.blogspot.com>



Questions?

Demo is available.

Thank you!