vl·e

virtual laboratory for e·science

Dynamic workflow

Adam Belloum
Systems and Networking Engineering research group
Institute of Informatics
University of Amsterdam

UvA

UNIVERSITEIT VAN AMSTERDAM

Outline

- Definition of workflows (**Business**, **Scientific**, **Dynamic**)
- Make the case for workflow and Dynamic workflow
- Described two different ways support Dynamic workflows

Business Workflows

- "The **automation** of a business process, in whole or parts, where **documents**, **information** or **tasks** are passed from one participant to another to be processed, according to a set of **procedural rules** "

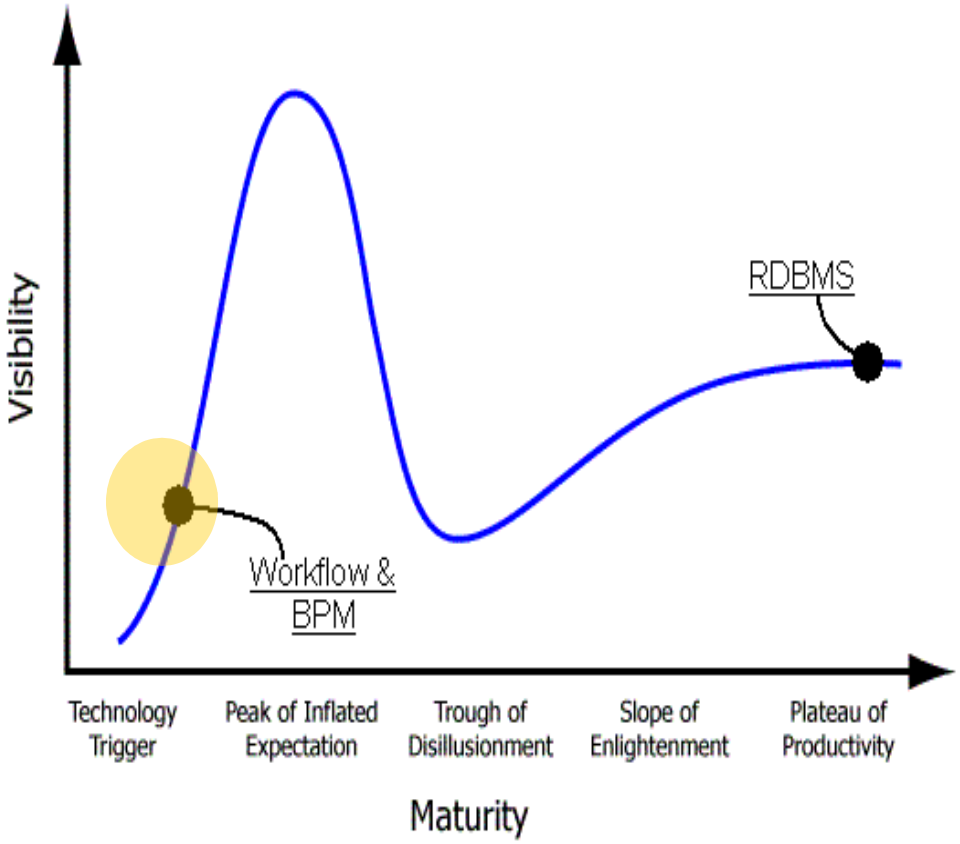
– Workflow Management Coalition

Scientific Workflows

- "These are networks of **analytical steps** that may involve, e.g., database access and querying steps, **data analysis** and **mining** steps, and many other steps including computationally **intensive jobs** on high performance cluster computers."

Bertram Ludäscher et al. (Kepler project)

Workflow Hype Curve



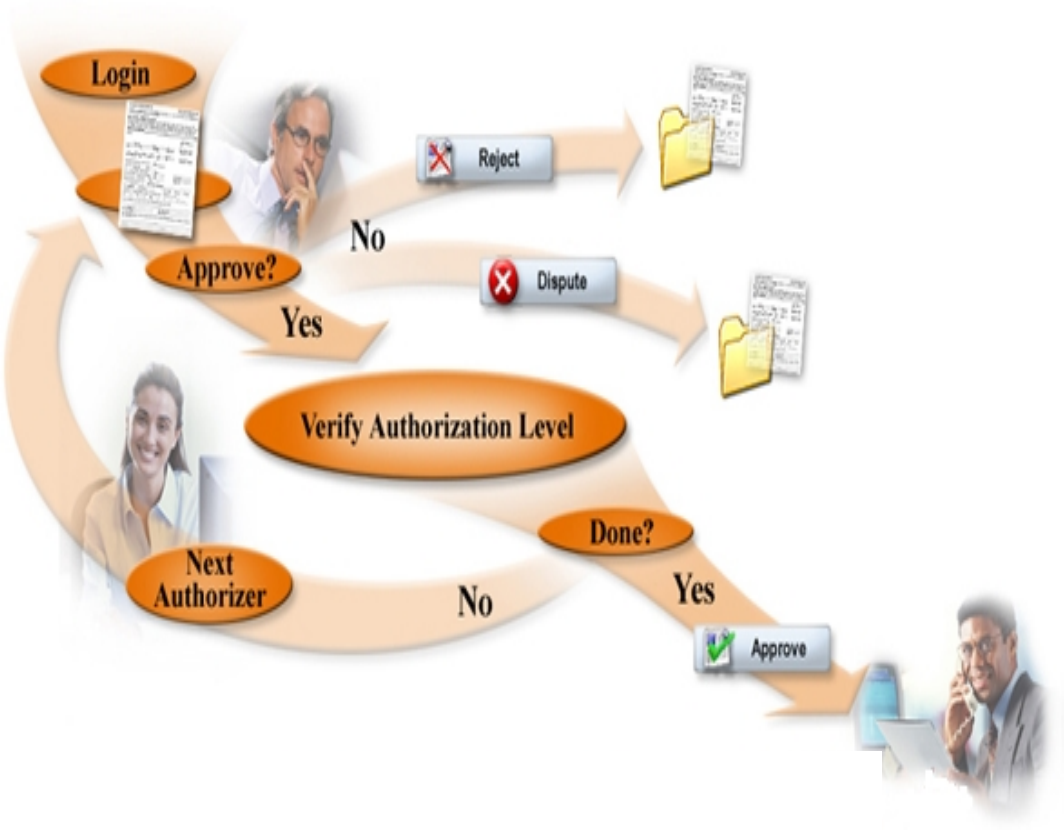
- "When talking about an **RDBMS** in a software development team most people will **get the picture** and shake their heads slightly up and down **confirming they understand what you're saying**. …"

http://www.theserverside.com/articles/article.tss?l=Workflow

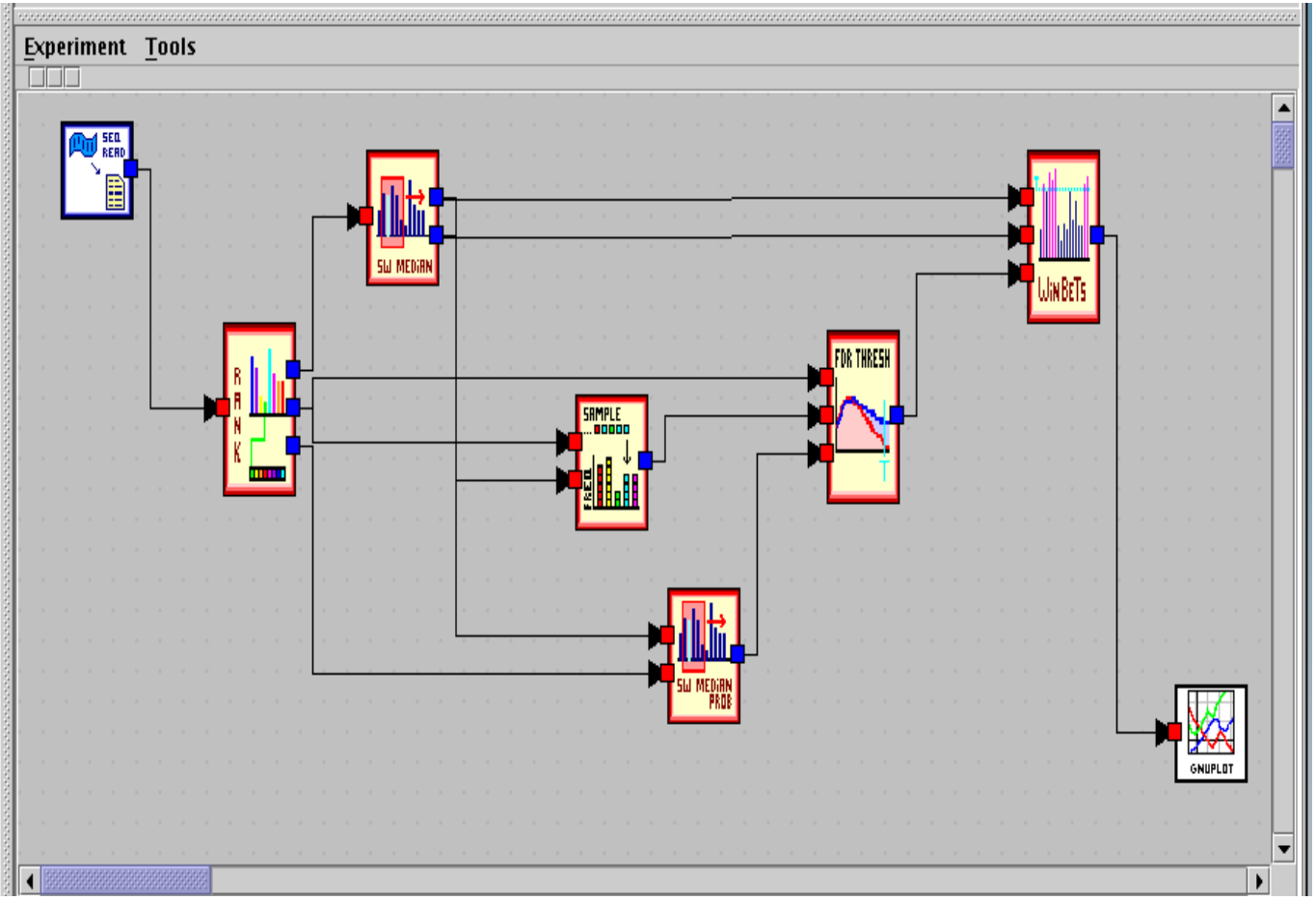
The case for **workflows**

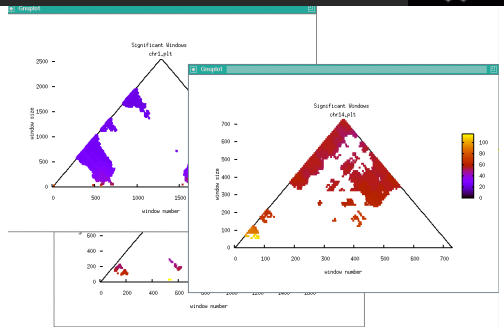
- **Capturing knowledge/enforce best practice**
  - Capture business process based on the company policy
  - Capture best practices of scientist, expert from a specific domain.
- **Incorporate human decision in the process**
  - There are cases that can not be automated both in business and scientific workflow
- **Easy development**
- **Increase the re-usability**

Business workflow …

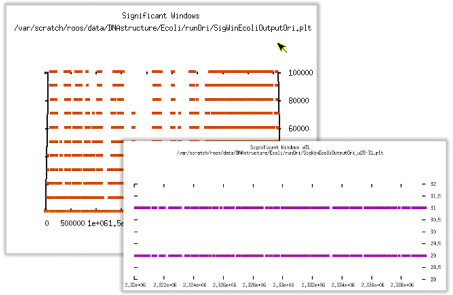


Scientific workflow

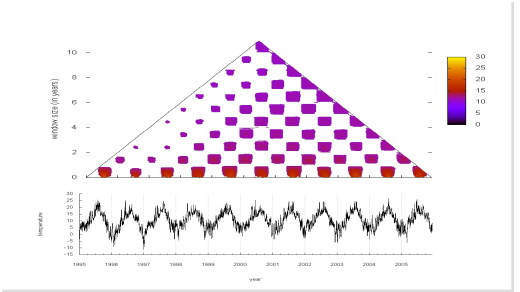




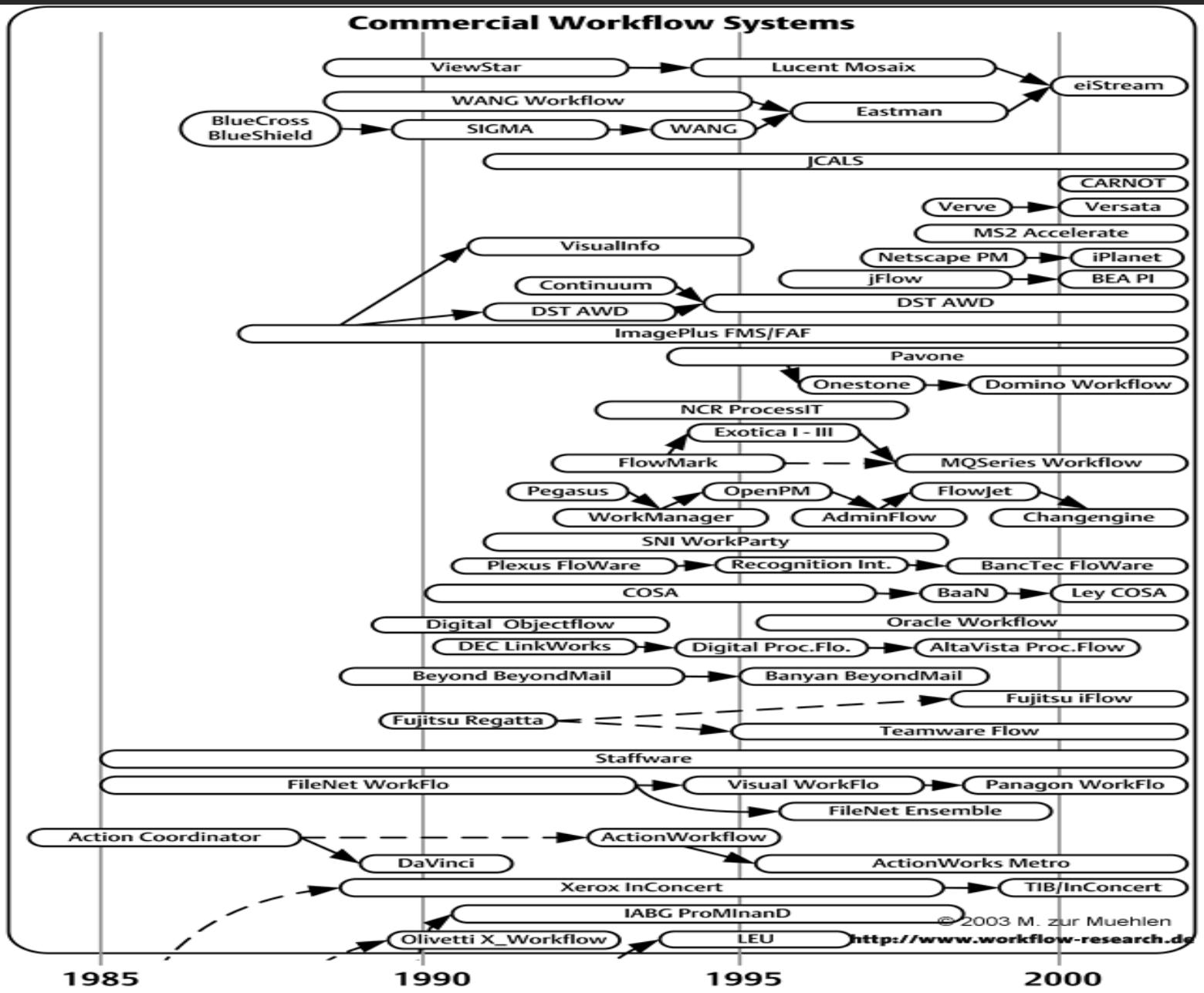
Human transcriptome map



DNA curvature of the *Escherichia coli* chromosome



Temperature in Amsterdam



http://www.workflow-research.de/Research/index.html

Scientific workflow management systems

- Askalon: **http://www.dps.uibk.ac.at/projects/askalon**
- Gridbus: **http://gridbus.csse.unimelb.edu.au/workflow**
- ICENI: **http://www.lesc.ic.ac.uk/iceni**
- Karajan: **http://www-unix.globus.org/cog/java**
- Kepler: **http://kepler-project.org**
- Pegasus: **http://pegasus.isi.edu**
- Taverna: **http://taverna.sourceforge.net**
- Triana: **http://www.trianacode.org**
- WS-VLAM: **http://www.science.uva.nl/~gvlam/ws-vlam**

http://www.science.uva.nl/~gvlam/project/doc/P2/SWMSRecommendationReport.pdf

The case for **Dynamic** workflows

- Different **alternatives** to proceed with the workflow and there is no way to decide which one is best at **design time**
- **Repeat** a sub-workflow until we get a certain condition is reached
- The **basic** structure or **semantics** changes. This may be because of an automated planner changing the workflow or optimized it on the fly based on prior execution knowledge.
- **Failure** in one part of the workflow may cause the whole workflow to fail if there is no alternative branch.

Dynamic workflow in a few words …

- ***WFMC*** or the e-Science definition

+

- The ability to **adapt automatically** or **via user input** at **run time** to changes without **compromising,** business logic, performance, safety etc

Manual changes (User in the loop)

- User stepping in and changing the flow on the fly.
- Can be achieved by providing computational **steering** capabilities

"Computational steering is a valuable mechanism for scientific investigation in **parameters** of a running program can be **altered** and the results visualized immediately"

Automatic changes (Autonomic Behavior)

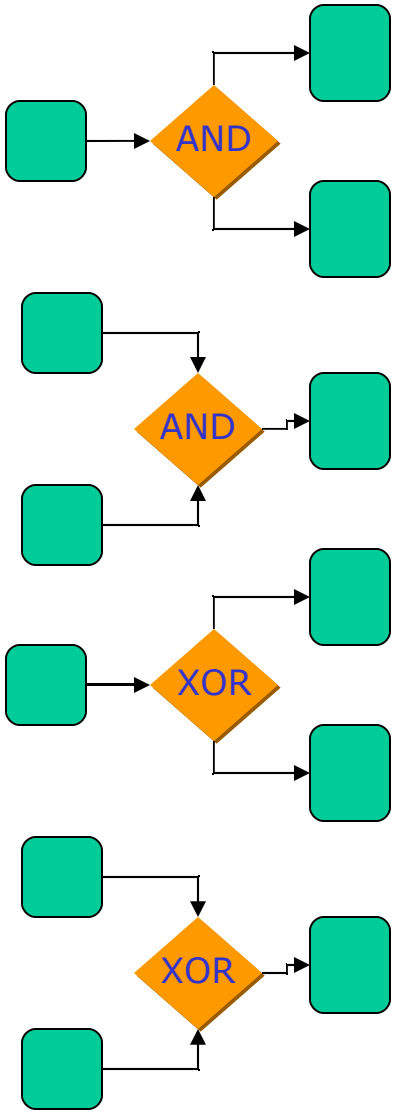
- The change of the flow is done on the fly by the system itself.
- Can be achieved by providing **control flow** capabilities

20 Control-flow patterns, are described in the workflow patterns page

www.workflowpatterns.com

Basic Control Patterns

- Parallel Split
  - execute activities in parallel
- Synchronization
  - synchronize two parallel threads of execution
- Exclusive Choice
  - choose one execution path from many alternatives
- Simple Merge
  - merge two alternative execution paths
- More control patterns can be found on the workflow patterns page



http://workflowpatterns.com

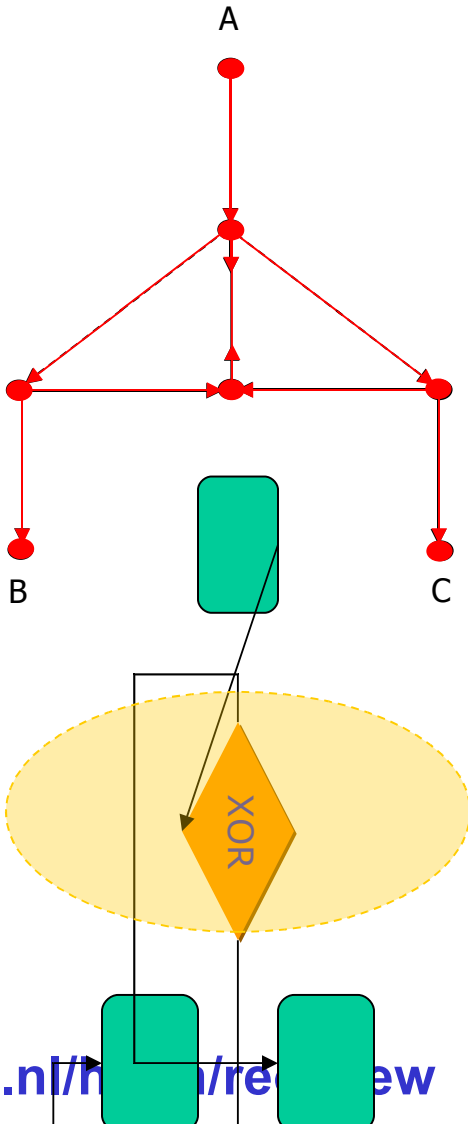
REO approach

- *Reo* is an **exogenous** coordination language based on a calculus of connector composition.
- Atomic connectors are called channels.
- Every channel represents a primitive interaction (protocol), explicitly defined as a binary constraint.
- Channel composition yields more complex interaction protocols, represented as more complex constraints.
- *Reo* connectors are dynamically reconfigurable.

- Synchronous channel
  - write/take 
- Synchronous drain: two sources
  - write/write 
- Synchronous spout: two sinks
  - take/take 
- Lossy synchronous channel 
- Asynchronous FIFO1 channel
  - write/take 

Exclusive Router using Reo connector

- Synchrony and exclusion propagate through synchronous segments of a circuit.



http://www.cwi.nl/h /re ew

Conclusion

- **Semantics** can play an important to support developing more dynamic workflow management systems
- **Interaction** with activities in workflow, steering workflows themselves and changing them during runs
- **How much** can or should be **automated**.

References

1. Workflow management coalition
   http://www.wfmc.org/
2. The *Workflow Patterns Initiative*
   http:www.workflowpatterns.com
3. Workflow Research **www.workflow-research.de/Research/index.html**
4. Composition by Anonymous Third Parties by Farhad Arbad, CWI
   http://www.cwi.nl/htbin/reo/view
5. Dynamic workflows and User Steering
   http://vtcpc.isi.edu/wiki/images/e/e3/Dynamic1.pdf

vl·e

http://www.vl-e.nl/