

### Introduction

Open Innovation paradigm [1]:

- Innovation projects results are strongly affected by intra-organization and inter-organizations collaborations
- Improving collaborative skills is a hot topic for innovation managers

Innovation as a collaborative process:

- Uncertain and unstructured
- Expert-driven

Need of systematic methodologies able to investigate and support collaborative tasks.

# **Application of Process Mining Techniques for Innovation Analysis and support**

Laura Genga Supervisor: Claudia Diamantini

Department of Information Engineering Marche Polytechnic University, Ancona, Italy

## **Objective**

Discovering knowledge models on **Common Collaboration Practices** with the aim of supporting and improving collaborative projects.

### **Main issues in target scenarios**

- Data retrieval: identification of suitable data sources
- Integration and synchronization of distributed and heterogeneous data
- Analysis of "spaghetti process"[3], with little or no structure

# **Research Approach**

### • Process Driven strategy

 Information extracted from data usually collected by enterprise informative systems

• ERP logs, emails, file versioning...

### Pattern Discovery Techniques

- Hierarchical clustering of subprocesses.
- Identifying the most frequent and important collaboration patterns





# Methodology

	Else         Estir         Format         Yew         Help           2011-03-06         20:40:43         LogTn         10.0.174.200           2011-03-06         20:47:45         LogTn         10.0.174.200           2011-03-07         27:35:46         LogTn         10.0.174.200           2011-03-07         27:35:46         LogTn         10.0.174.200           2011-03-07         27:41:33         LogOut         10.0.174.200           2011-03-07         12:32:26         LogTn         10.0.174.200           2011-03-07         12:37:41         LogOut         10.0.174.200           2011-03-07         12:37:41         LogTn         10.0.174.200           2011-03-07         12:37:41         LogTn         10.0.174.200           2011-03-07         12:37:43         LogTn         10.0.174.200           2011-03-07         12:37:43         LogTn         10.0.174.200           2011-03-07         12:51:31         LogTn         10.0.174.200           2011-03-07         12:51:31         LogTn         10.0.174.200           2011-03-07         12:51:35         LogTn         10.0.174.200           2011-03-07         12:51:34         LogTn         10.0.174.200					
	2011-05-08 10:12:44 LogOut 10.0.174 2011-05-08 10:12:44 LogOut 10.0.174 2011-05-08 10:13:72 LogOut 10.0.174 2011-05-08 10:13:72 LogOut 10.0.174 2011-05-08 10:14:23 LogOut 10.0.174 2011-05-08 10:14:23 LogOut 10.0.174 2011-05-08 10:14:25 LogOut 10.0.174	Case id	Timestamp	Activity	Resource	
<u>Team members</u>			30-12-2010:11.02	a	Pete	\
			31-12-2010:10.06	с	Sue	
	<u>User PC</u>		05-01-2011:15.12	а	Mike	
			06-01-2011:11.18	a	Sara	
			06-01-2011:11.23	с	Pete	
			07-01-2011:14.24	d	Mike	
		2	30-12-2010:11.32	b	Mike	
			30-12-2010:12.12	а	Mike	
			30-12-2010:14.16	C	Pete	
K			05-01-2011:11.22	а	Sara	
	1		05-01-2011:11.40	С	Ellen	
			08-01-2011:12.05	b	Pete	
$\frown$		3	30-12-2010:14.32	а	Pete	
	Y		30-12-2010:15.06	C	Mike	
(2)			30-12-2010:16.34	z	Ellen	
			06-01-2011:09.18	d	Sara	
_			06-01-2011:12.18	h	Sara	
			06-01-2011:13.06	g	Sean	
			08-01-2011:11.43	z	Pete	/
		$\backslash$	09-01-2011:09.55	d	Sara	
			15-01-2011:10.45	f	Ellen	
		3) <u>Pa</u>	<u>ttern Dis</u>	<u>scovery</u>	<u>7</u>	
		A seq_A_C C	s_ C j_	B_C B C_A A	s_B_A A j_A_A	
	<b>→</b>	seq_C_B		↓s	eq_A_C	

s\_B\_C

j\_C\_D

С

В

s\_B\_A

#### 1) Data collection

Collection of data regarding daily activities of a research team :

- files usually stored in enterprise IS and team members PCs
- human knowledge about informal activities not automatically stored,
   e.g. physical meeting

### 2) Data Integration

Obtaining a single dataset containing all collaborative processes logs
Main Steps:

- Data ExtractionData Cleaning
- •Homogenization









С

В

seq\_A\_C





a) Trace preprocessing to obtain corresponding process schema

b) Schema conversion into graphs structures c) SUBDUE [2] launching to derive a lattice structure, by arranging organization practices in a hierarchy with different level of abstractions. Higher-level clusters contain the most important patterns of the domain.

### Conclusion

А

- Identifying the most common organization collaborative practices with their relationships
- Supporting organization collaborative work management
- Exploiting information stored about actual organization collaborative tasks

#### References

- [1] Chesbrough, H.W., Open innovation: the new imperative for creating and profiting from technology, Harvard Business Press, 2003.
- [2] Jonyer, I., Cook, D.J., Holder, L.B., "Graphbased hierarchical conceptual clustering", J. Mach. Learn. Res., vol. 2, pp 19-43, March 2002.
- [3]Van der Aalst, W.M.P., Process mining: discovery, conformance and enhancement of business processes, Springer, 2011.

### **Further Information**

l.genga@univpm.it

Third European Business Intelligence Summer School, July 7-12, 2013 – Schloss Dagstuhl, Wadern, Germany