

Coordination of open data development in Croatia – case study of Environmental Pollution Registry

Neven Vrček, Katarina Tomičić-Pupek

University of Zagreb, Faculty of Organization and
Informatics,

Pavlinska 2, Varaždin, Croatia

neven.vrcek@foi.hr; ktomicic@foi.hr

Delivery

Gathering

Processing

Internaly lifecycle can get quite complex. It can span several organizations (interoperabilty). How about insight? Can we open lifecycle and build open data by design?

Open data lifecycle (II)

- Data collections can consist of **large amount of information** and there can be a need of tracking data changes or lifecycle procedures through a long period of time (in order to **track historical changes**).
- Also, the case study shows that the (pollution) data can be **very sensitive**, so collection and processing should be verified and there should be **a clear data handling methodology** to ensure important quality issues.
- Important quality issues regarding open data systems (disclaimers are not good for development of open data market):
 - **quality of the data** itself, like reliability, consistency and usefulness and
 - **quality of the data handling procedures**.
- Some important quality issues of data handling procedures are **transparency** and the ability to **verify** that the data is processed from its source to the end-user in a way that maintains reliability and correctness.

Open data lifecycle (III)

- In order to achieve **transparency** and establish the ability to **verify** the handling of the data from its source to the end-user in an appropriate way, we suggest the use of **the Business Process paradigm** to represent the lifecycle of the open data.
- The openness and transparency of the consistent process model of handling open data (the lifecycle of open data) enables re-user's trust in data quality and persistence.
- Also **the cost per unit of data** can be calculated and simulated in advance (utilization of graph theory and colored Petri nets)

Process paradigm

- process analysis
- process modelling
- process execution by utilizing information technologies
- important for development, implementation and use of information systems in modern organizations

BPMN 2.0



BPMN 2.0 - Business Process Model and Notation

<http://bpmb.de/poster>

Activities

- Task**: A Task is a unit of work, the job to be performed. When marked with a symbol it indicates a Sub-Process, an activity that can be refined.
- Transaction**: A Transaction is a set of activities that logically belong together; it might follow a specified transaction protocol.
- Event Sub-Process**: An Event Sub-Process is placed into a Process or Sub-Process. It is activated when its start event gets triggered and can interrupt the higher level process context or run in parallel (non-interrupting) depending on the start event.
- Call Activity**: A Call Activity is a wrapper for a globally defined Sub-Process or Task that is reused in the current process.

Activity Markers

- Sub-Process Marker
- Loop Marker
- Parallel All Marker
- Sequential All Marker
- Ad Hoc Marker
- Compensation Marker

Task Types

- Send Task
- Receive Task
- User Task
- Manual Task
- Business Rule Task
- Service Task
- Script Task

- Sequence Flow**: defines the execution order of activities.
- Default Flow**: is the default branch to be chosen if all other conditions evaluate to false.
- Conditional Flow**: has a condition assigned that defines whether or not the flow is used.

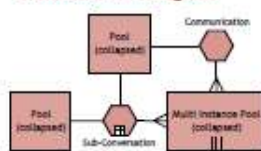
Gateways

- Exclusive Gateway**: When splitting, it routes the sequence flow to exactly one of the outgoing branches. When merging, it awaits one incoming branch to complete before triggering the outgoing flow.
- Event-based Gateway**: It is always followed by catching events or receive tasks. Sequence flow is routed to the subsequent event/task which happens first.
- Parallel Gateway**: When used to split the sequence flow, all outgoing branches are activated simultaneously. When merging parallel branches it waits for all incoming branches to complete before triggering the outgoing flow.
- Inclusive Gateway**: When splitting, one or more branches are activated. All active incoming branches must complete before merging.
- Exclusive Event-based Gateway**: Each occurrence of a subsequent event starts a new process instance.
- Complex Gateway**: Complex merging and branching behavior that is not captured by other gateways.
- Parallel Event-based Gateway (Instantiated)**: The occurrence of all subsequent events starts a new process instance.

Conversations

- A Communication defines a set of logically related message exchanges. When marked with a symbol it indicates a Sub-Conversation, a compound conversation element.
- A Conversation Link connects Communications and Participants.
- A Forked Conversation Link connects Communications and multiple Participants.

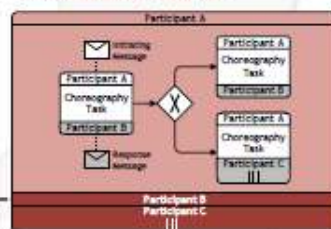
Conversation Diagram



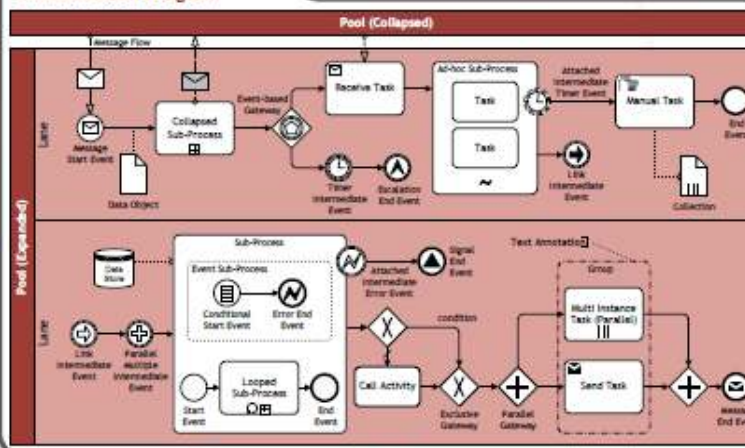
Choreographies

- A Choreography Task represents an Interaction (Message Exchange) between two Participants.
- Multiple Participants Marker denotes a set of Participants of the same kind.
- A Choreography Sub-Process contains a refined choreography with several interactions.

Choreography Diagram



Collaboration Diagram



Swimlanes

- Pools (Participants) and Lanes represent responsibilities for activities in a process. A pool or a lane can be an organization, a role, or a system. Lanes subdivide pools or other lanes hierarchically.
- Message Flow symbolizes information flow across organizational boundaries. Message flow can be attached to pools, activities, or message events.

Events


	Start	Intermediate	End
Typical Level			
Event Sub-Process Interrupting			
Event Sub-Process Non-Interrupting			
Catching			
Boundary Interrupting			
Boundary Non-Interrupting			
Throwing			
None (Unusable)			
Message: Receiving and sending messages			
Timer: Cyclic timer events, points in time, time spans or timeouts			
Escalation: Escalating to an higher level of responsibility			
Conditional: Reacting to changed business conditions or integrating business rules			
Link: Off-page connectors. Two corresponding link events equal a sequence flow.			
Error: Catching or throwing named errors.			
Cancel: Reacting to cancelled transactions or triggering cancellation.			
Compensation: Handling or triggering compensation.			
Signal: Signalling across different processes. A signal thrown can be caught multiple times.			
Multiple: Catching one out of a set of events. Throwing all events defined.			
Parallel Multiple: Catching all out of a set of parallel events.			
Terminate: Triggering the immediate termination of a process.			

Data

- A Data Input is an external input for the entire process. It can be read by an activity.
- A Data Output is a variable available as result of the entire process.
- A Data Object represents information flowing through the process, such as business documents, e-mails, or letters.
- A Collection Data Object represents a collection of information, e.g., a list of order items.
- A Data Store is a place where the process can read or write data, e.g., a database or a filing cabinet. It persists beyond the lifetime of the process instance.
- A Message is used to depict the contents of a communication between two Participants.



Case study

 Preglednik registra onečišćavanja okoliša

AGENCIJA ZA ZAŠTITU OKOLIŠA [Preglednik](#) [Upute za korištenje](#) [Pregled šifranika](#) [Pojmovnik Preglednika ROO](#)

Pretraživanje za godinu

Upit:

Filtar: =

Označi sve kolone Isključi sve

Kolone:

- Podaci o operateru
 - Godina
 - Županija
 - Matični broj subjekta (MBS) ili matični broj obrta (MBO)
 - OIB

Broj zapisa: **1987**

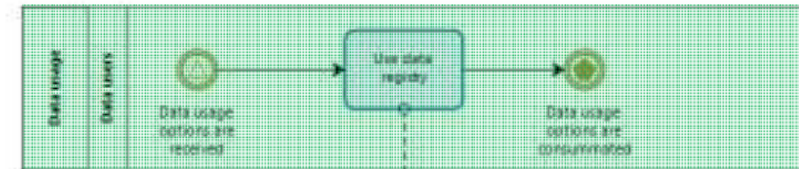
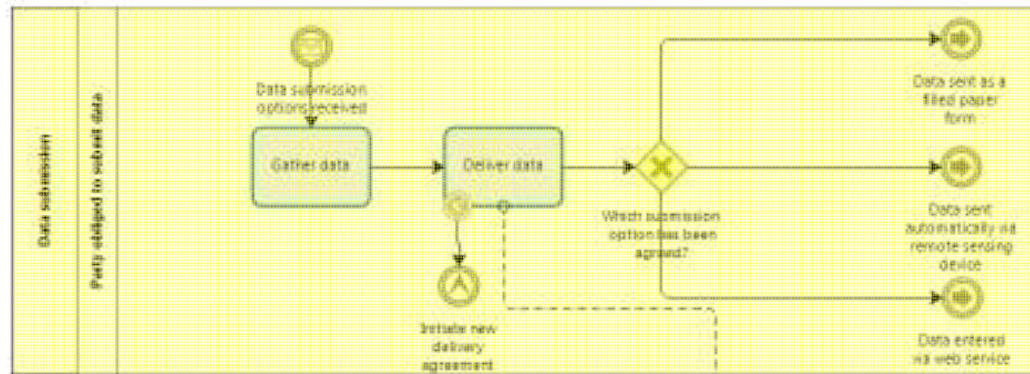
Godina	Županija	Matični broj subjekta (MBS) ili matični broj obrta (MBO)	OIB	Naziv tvrtke ili obrta
2013	Osječko-baranjska	030028386	30605443172	KG Park d.o.o.
2013	Splitsko-dalmatinska	060018277	50405970468	AUTOKUĆA VRDOLJAK d.o.o.
2013	Međimurska	070000473	52347609859	Tegra d.o.o.
2013	Grad Zagreb	080338987	13308543980	Danik doo

Koji je optimalni broj modula IS-a (podsustava, aplikacija)?

BPMN model of Environmental Pollution Register

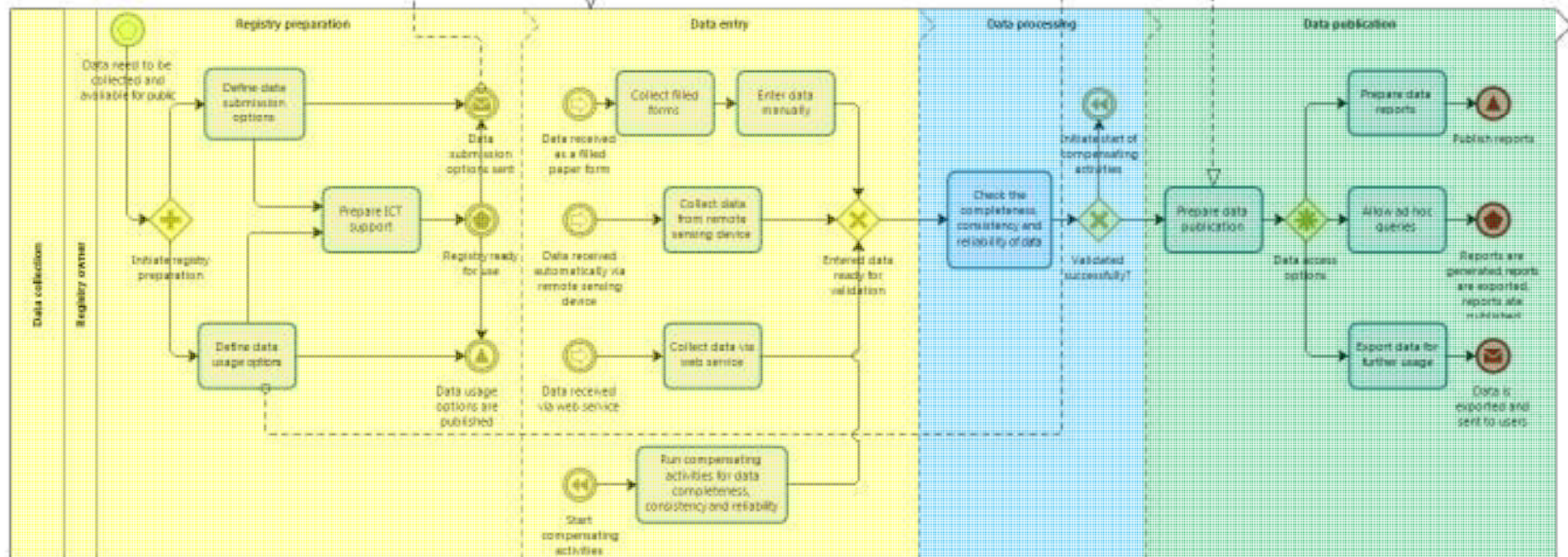


Gathering

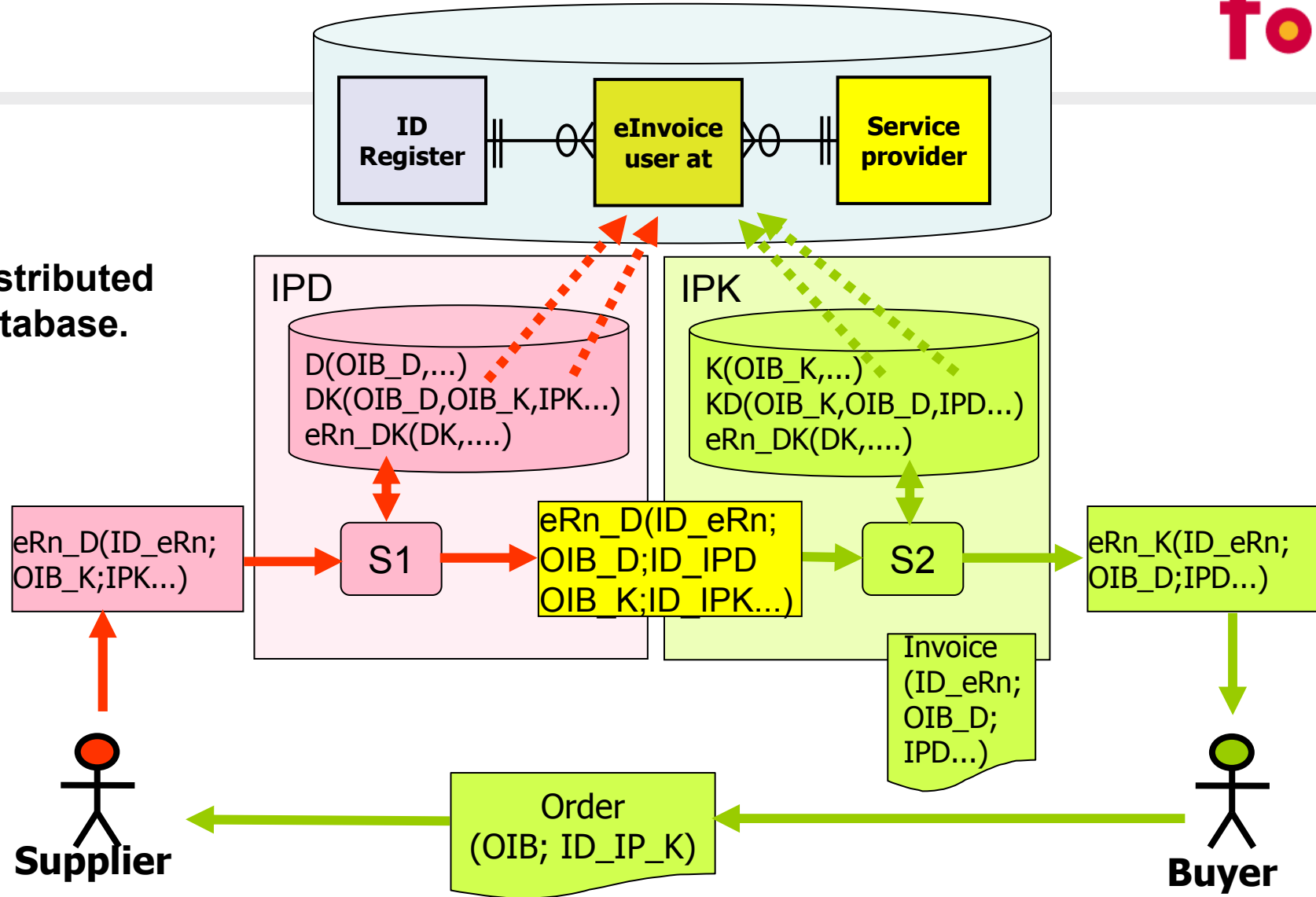


Delivering

Processing



Distributed database.



What we achieved?

- Analysys can be extended to desired level of detail
- Deep insight into business processes for all stakeholders
- Consistency of procedures
- Process simulation (what if analysis)
- Cost evaluation and monitoring

Challenges and benefits regarding data delivering

- It can be useful to mind that users can be external and internal. Internal re-users, like the public bodies themselves, often use the open data only for purposes of redistributing the same information to other external users (e.g. links to web pages)
 - Example of this kind of data consumption by public/governmental bodies themselves - ***Assessment of sea bathing water quality***. Based on this open data collection a National annual report on sea bathing water quality at Croatian beaches is published. Other governmental institutions use this data in their domain (Ministry of tourism, Local departments of public Health,...).

