Verification of Data-Driven Business Processes

dr.ir. Rik Eshuis

Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

Verifying business processes

Real world

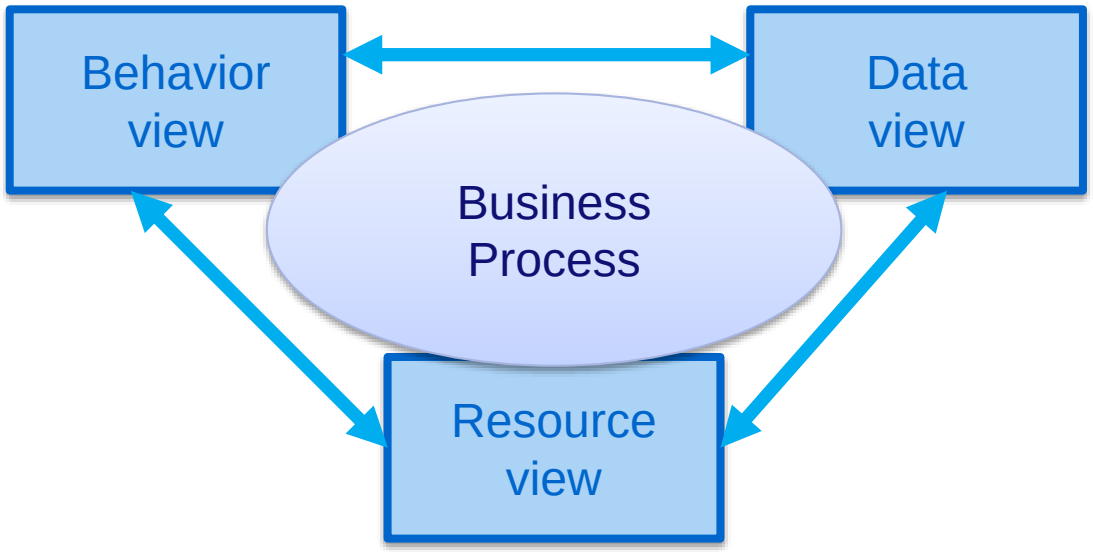
Model world

Process model

verification

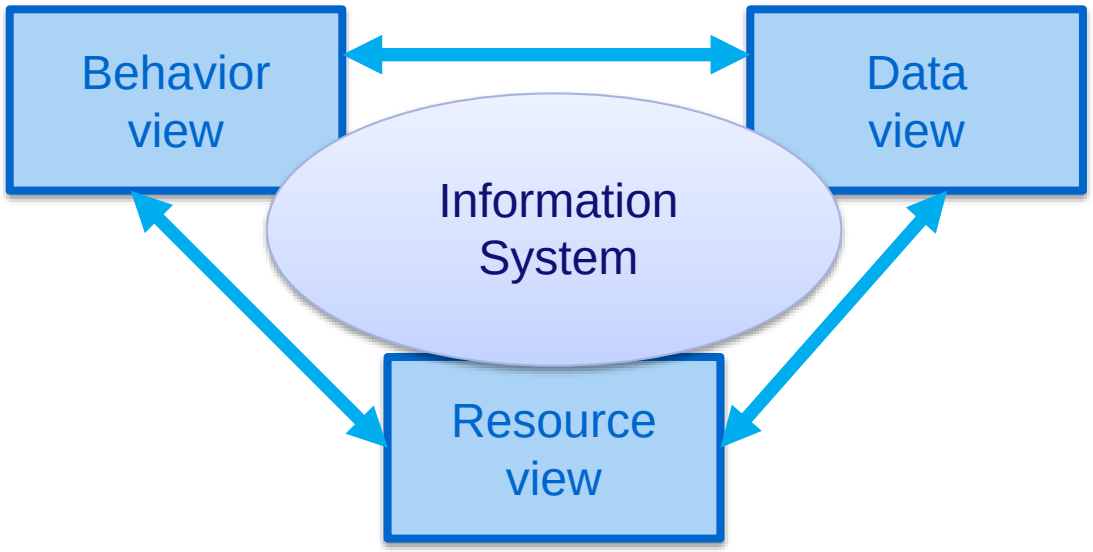
Property

Process modeling perspectives for information systems



Separation of concerns: engineer models separately

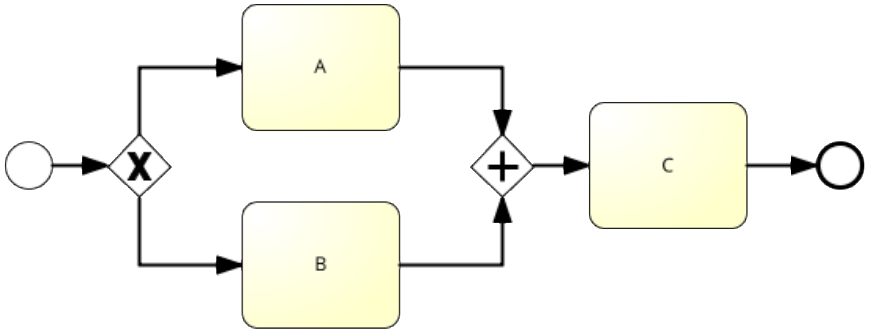
Process modeling perspectives for information systems

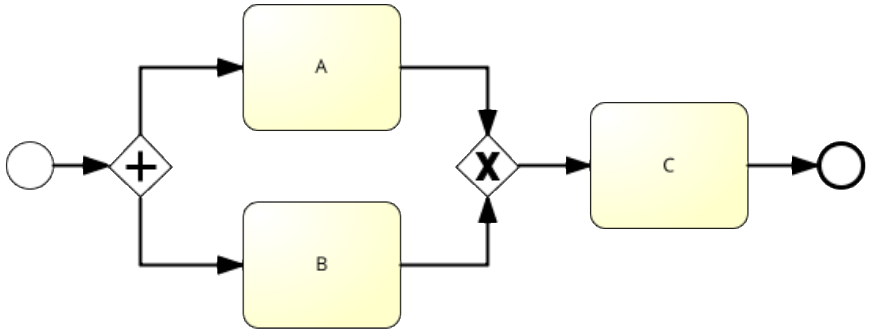


Separation of concerns: engineer models separately

Early process verification approaches: dedicated tools

- Applying algorithms from Petri nets / graph theory
- Correctness formalized in **fixed properties**:
  no deadlock + no lack of synchronization





~ soundness

- Sample tools: Woflan, FlowMake

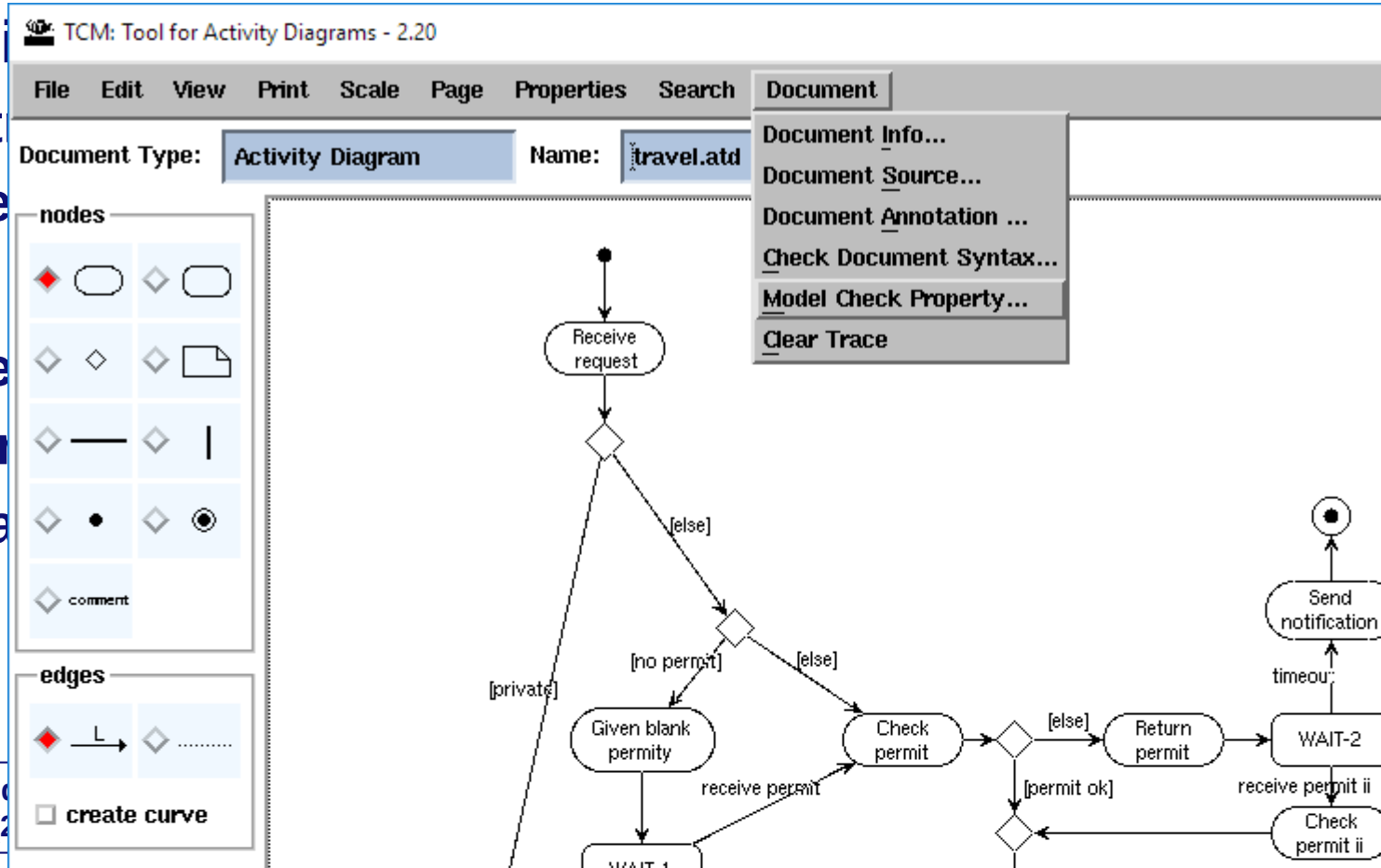
Early process verification approaches: standard tools

- Applying verification tools like SMV and Spin
  - Translate process model into state machine
- Declarative properties (in temporal logic)
  - correctness
  - user defined properties (compliance)
- Sample: integrating NuSMV with activity diagram editor
  - "Safe" activity diagrams
  - Automated translation
  - Visualize counterexamples in diagram of editor

Rik Eshuis: Symbolic model checking of UML activity diagrams. ACM TOSEM 15(1): 1-38 (2006)

Early process verification approaches: standard tools

- Applying verification tools like SMV and Spin
  - Translate process model into state machine
- Declarati
  - correct
  - user de
- Sample:
  - "Safe
  - Auto
  - Visua



Rik Eshuis: Symbolic mo
ACM TOSEM 15(1): 1-38 (2

Verifying business processes: early days

Real world

Model world

Process model

dedicated or standard

abstraction

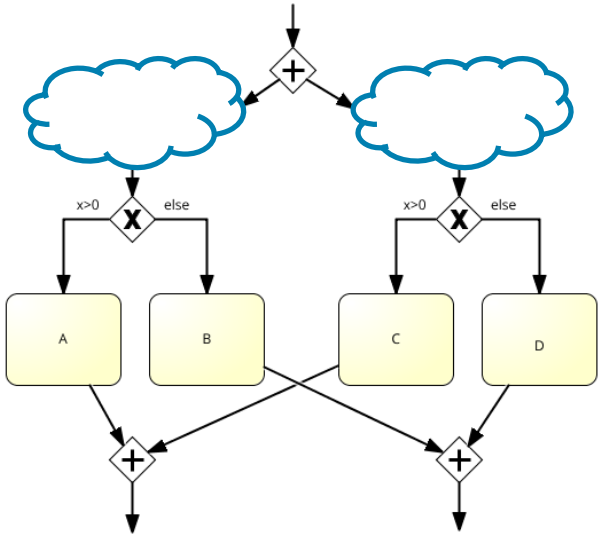
Behavior model

verification

Property

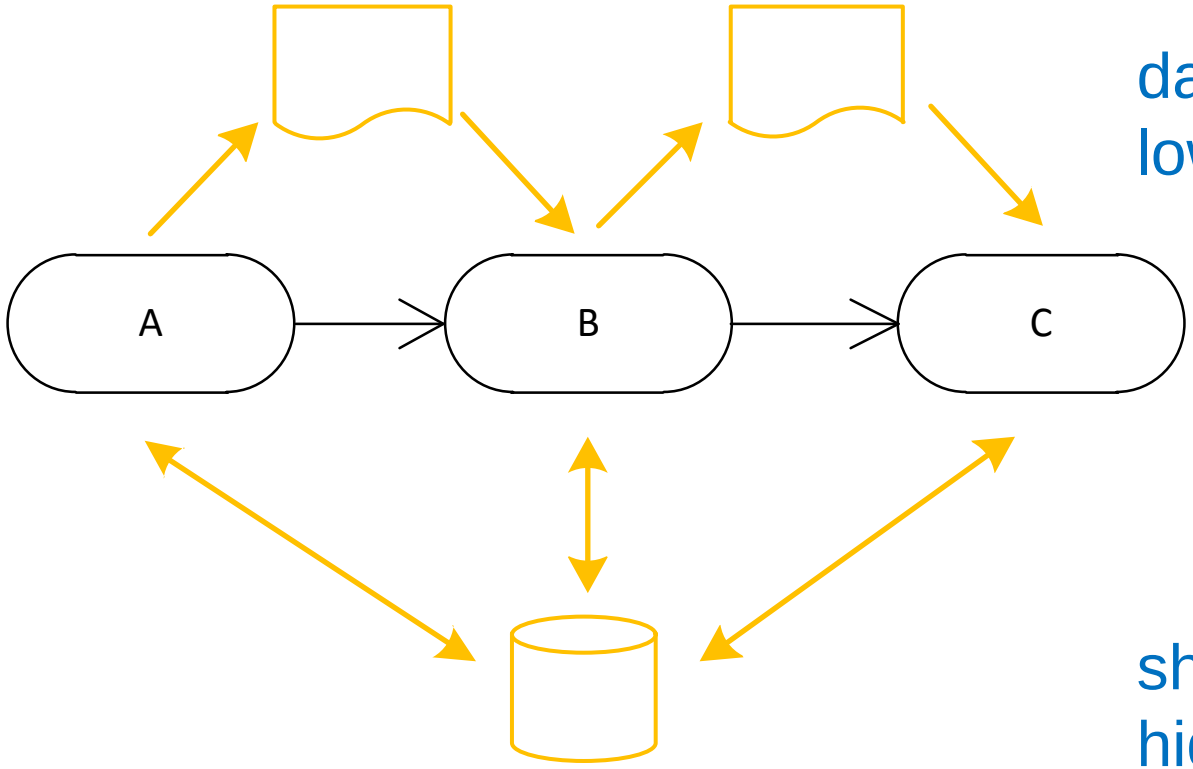
Abstraction

- Early verification tools focus on behavior
- Implies abstraction from other views is required:
  - Example: abstraction from data in classical Petri nets
- At the expense of precision



Gain precision by incorporating data flow aspects in verification problem

Data-driven processes: Two types of data flow

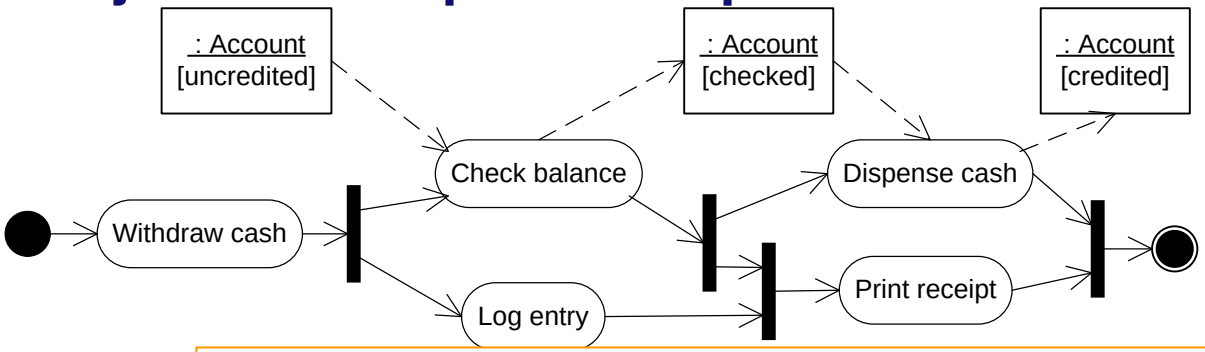


data objects:
low granularity

shared variables:
high granularity

Verifying process models with data objects

- Different object nodes with same name refer to same object
- Integrity constraints based on CRUD lifecycle, e.g.
  - A created object is not used before creation
  - An object is not updated in parallel



Check bal ***unt***

$G\,(\textit{Check balance} \rightarrow \neg\, \textit{Dispense cash})$

$G\,(\textit{Dispense cash} \rightarrow \neg\, \textit{Check balance})$

Rik Eshuis: Symbolic model checking of UML activity diagrams. ACM TOSEM 15(1): 1-38 (2006)

Verifying process models with shared variables

- Activities can read, write variables
- Properties: absence of
  - Missing data: access variable before it is initialized
  - Redundant data: written variable is not used
  - Conflicting data: variable is written by two activities
- Dedicated verification (flow analysis, Petri nets)

Sherry X. Sun, J. Leon Zhao, Jay F. Nunamaker Jr., Olivia R. Liu Sheng: Formulating the Data-Flow Perspective for Business Process Management. Information Systems Research 17(4): 374-391 (2006)

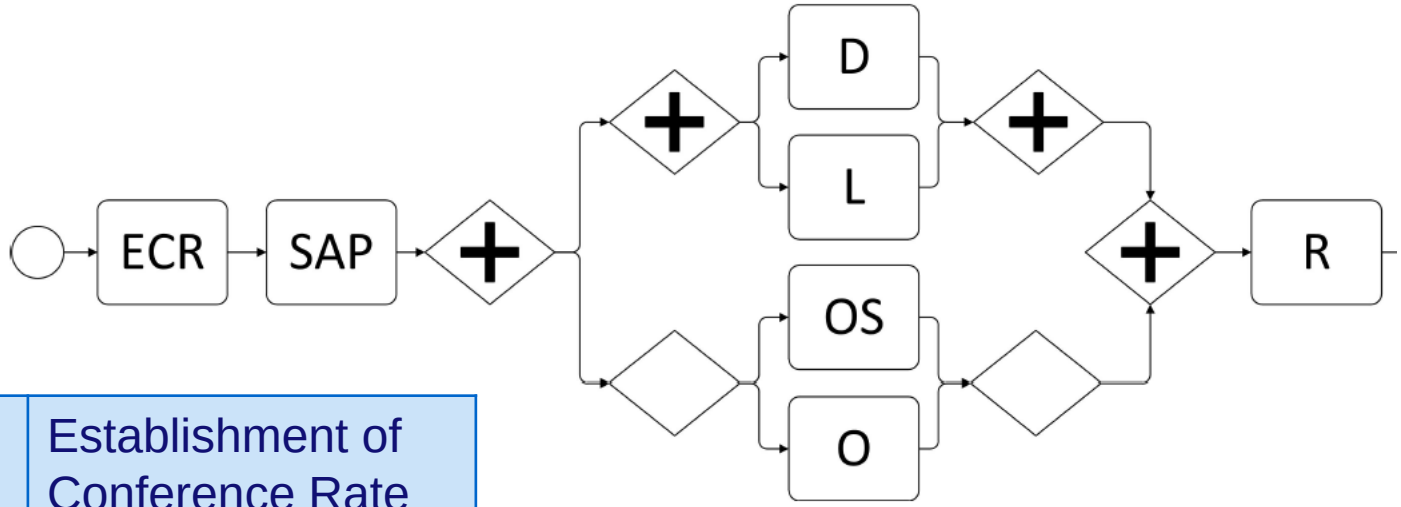
Natalia Sidorova, Christian Stahl, Nikola Trcka: Soundness verification for conceptual workflow nets with data: Early detection of errors with the most precision possible. Inf. Syst. 36(7): 1026-1043 (2011)

Verifying semantic process models

- **Process models with shared variables and pre/postconditions for activities**
- **Process gets stuck at activity if precondition fails**
- ***How can we diagnose correctness of process models with pre/postconditions?***

Diana Borrego, Rik Eshuis, María Teresa Gómez López, Rafael M. Gasca: Diagnosing correctness of semantic workflow models. Data Knowl. Eng. 87: 167-184 (2013)

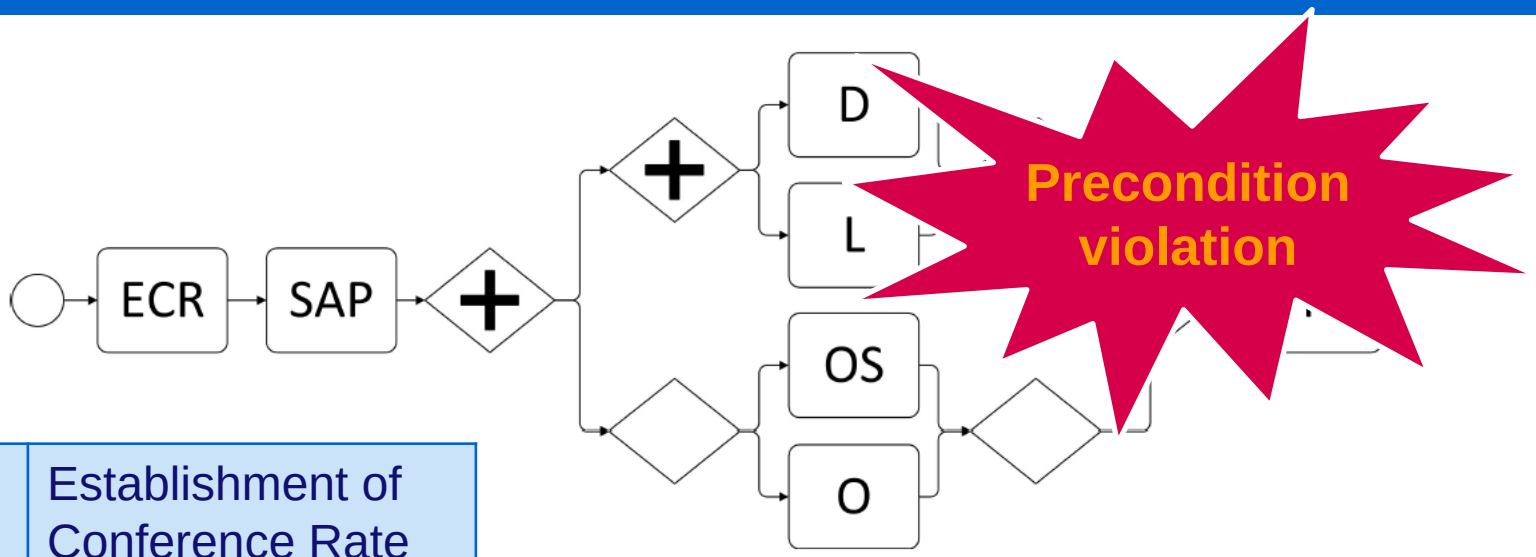
Example



ECR	Establishment of Conference Rate
SAP	Selection of Accepted Papers
D	Dinner
L	Lunch
OS	Other expenses + Social event
O	Other expenses
R	Registration

Activity	Precondition and Postcondition
ECR	**pre:** $true$ **post:** $true$
SAP	**pre:** $true$ **post:** $true$
D	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $regFee * 0.1 \leq dinner \wedge dinner \leq regFee * 0.35$
L	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $regFee * 0.1 \leq 3 * lunch \wedge 3 * lunch \leq regFee * 0.35$
OS	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $others \leq 0.2 * regFee + 0.05 * sponsorship \wedge$ $others \geq 0.05 * regFee + 0.05 * sponsorship$
O	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $others \leq 0.25 * regFee \wedge others \geq 0.05 * regFee$
R	**pre:** $3 * lunch + dinner + others < regFee$ **post:** $numPapers * 1.8 \geq confAtt \wedge numPapers * 0.5 \leq confAtt$
NGS	**pre:** $confAtt * (3 * lunch + dinner + others) <$

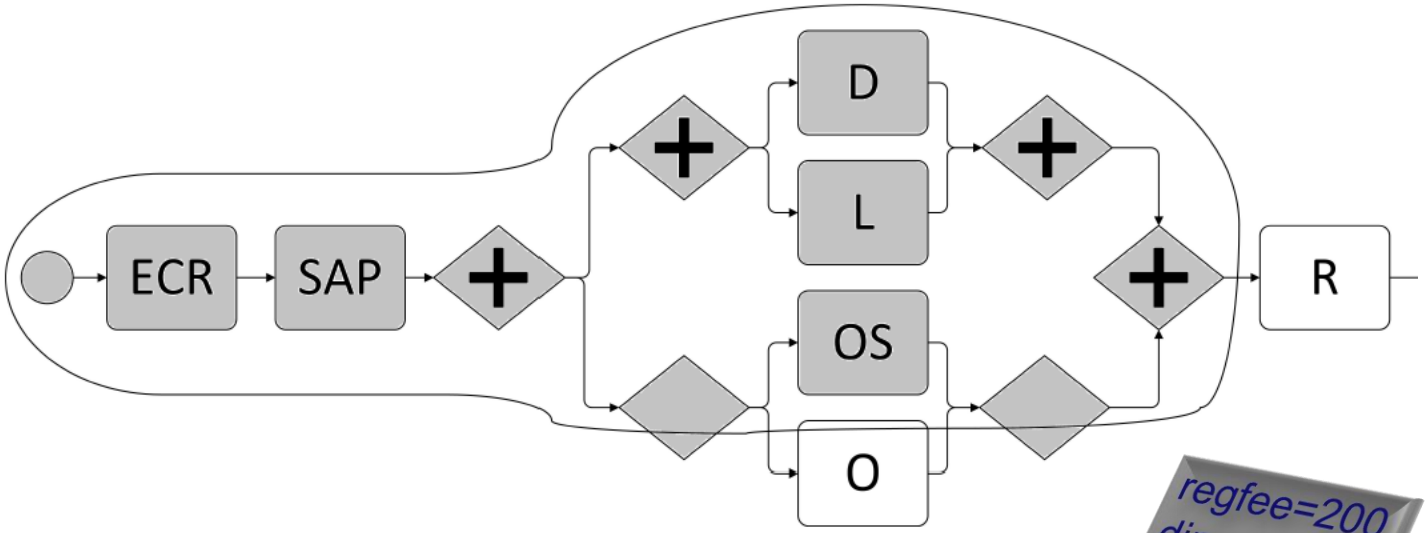
Example



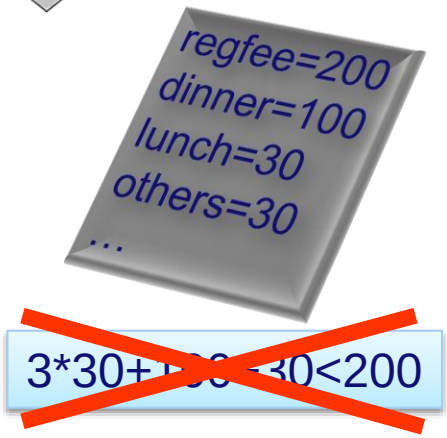
ECR	Establishment of Conference Rate
SAP	Selection of Accepted Papers
D	Dinner
L	Lunch
OS	Other expenses + Social event
O	Other expenses
R	Registration

Activity	Precondition and Postcondition
ECR	**pre:** $true$ **post:** $true$
SAP	**pre:** $true$ **post:** $true$
D	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $regFee * 0.1 \leq dinner \wedge dinner \leq regFee * 0.35$
L	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $regFee * 0.1 \leq 3 * lunch \wedge 3 * lunch \leq regFee * 0.35$
OS	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $others \leq 0.2 * regFee + 0.05 * sponsorship \wedge$ $others \geq 0.05 * regFee + 0.05 * sponsorship$
O	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $others \leq 0.25 * regFee \wedge others \geq 0.05 * regFee$
R	**pre:** $3 * lunch + dinner + others < regFee$ **post:** $numPapers * 1.8 \geq confAtt \wedge numPapers * 0.5 \leq confAtt$
NGS	**pre:** $confAtt * (3 * lunch + dinner + others) <$

Diagnosis

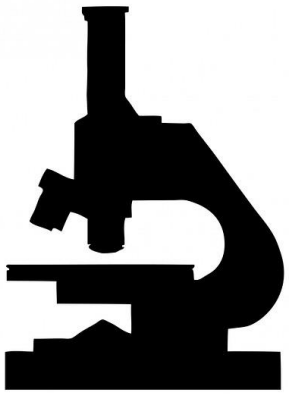


O	**pre:** $sponsorship > 0 \vee numPapers > 60$ **post:** $others \leq 0.25 * regFee \wedge others \geq 0.05 * regFee$
R	**pre:** $3 * lunch + dinner + others < regFee$ **post:** $numPapers * 1.8 \geq confAtt \wedge numPapers * 0.5 \leq confAtt$
NGS	**pre:** $confAtt * (3 * lunch + dinner + others) < confAtt * regFee + sponsorship$ **post:** $guestSpeaker \geq 0.2 * sponsorship \wedge guestSpeaker \leq sponsorship + 0.1 * regFee * confAtt$
IGS	**pre:** $confAtt * (3 * lunch + dinner + others) < confAtt * regFee + sponsorship$ **post:** $guestSpeaker \geq 0.4 * sponsorship \wedge guestSpeaker \leq sponsorship$



Diagnosis approach

- **Workflow graphs + pre/postconditions**
- **Encode correct control flow execution (subgraph) as Integer Programming model**
  - **Rik Eshuis, Akhil Kumar: An integer programming based approach for verification and diagnosis of workflows. *Data Knowl. Eng.* 69(8): 816-835 (2010)**
- **Encode pre/postconditions as Constraint Satisfaction Problem**
- **Combined IP+CSP model processed by CSP solver**

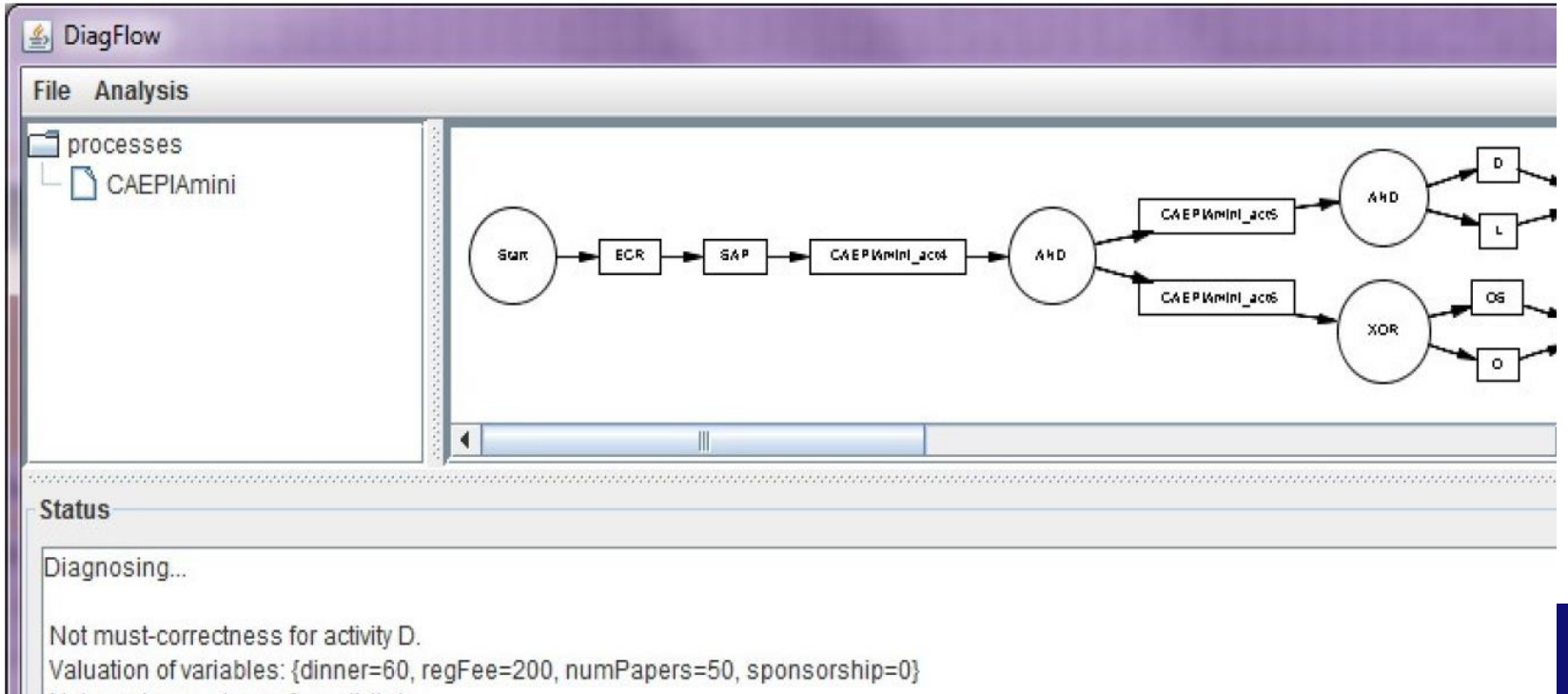


Algorithm

- Convert workflow graph into SSA form
- For each activity a with precondition p
  - Find subgraph that leads to a but does not include a
  - Encode subgraph as CSP+IP model
  - Add not(p) to CSP+IP model
  - If solution exists, then error
- If no error found, then the workflow graph is correct

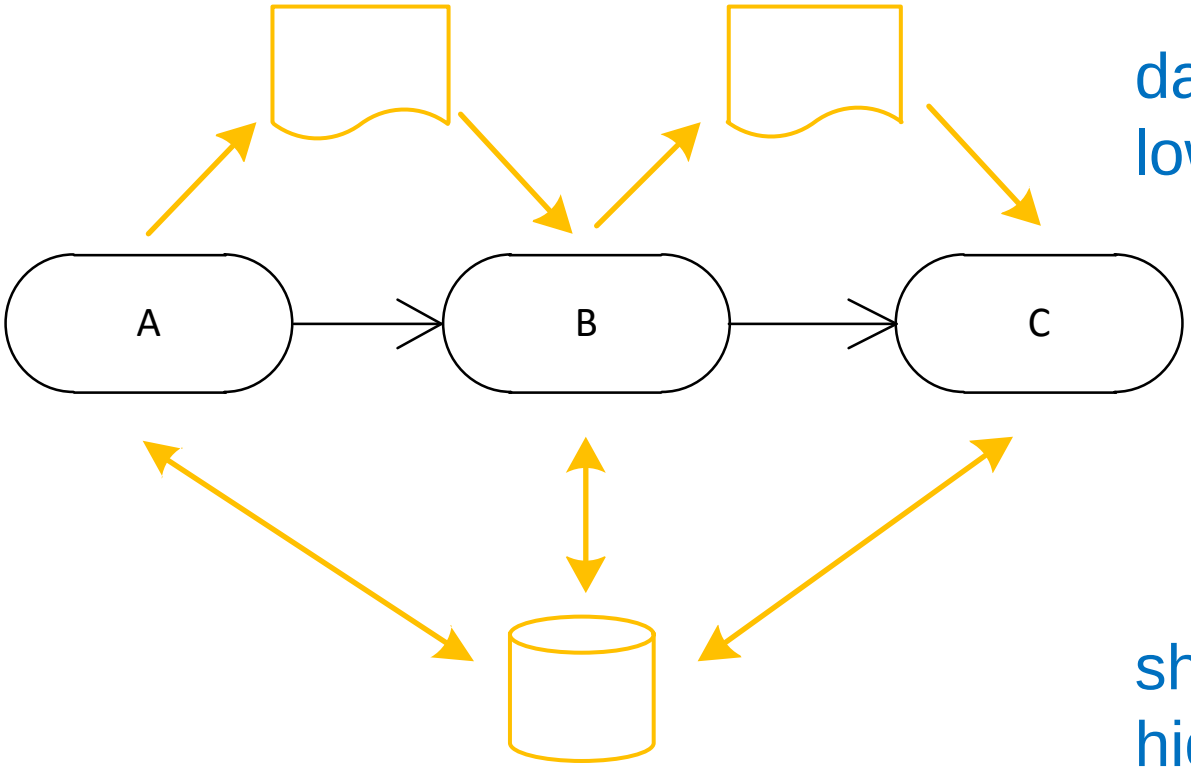
Results

- **Seamless integration of process and data verification**
- **Extensive feedback to fix data flow errors**
- **Fully implemented (using CSP solver)**



Diana Bor
Gasca: Di
Knowl. En

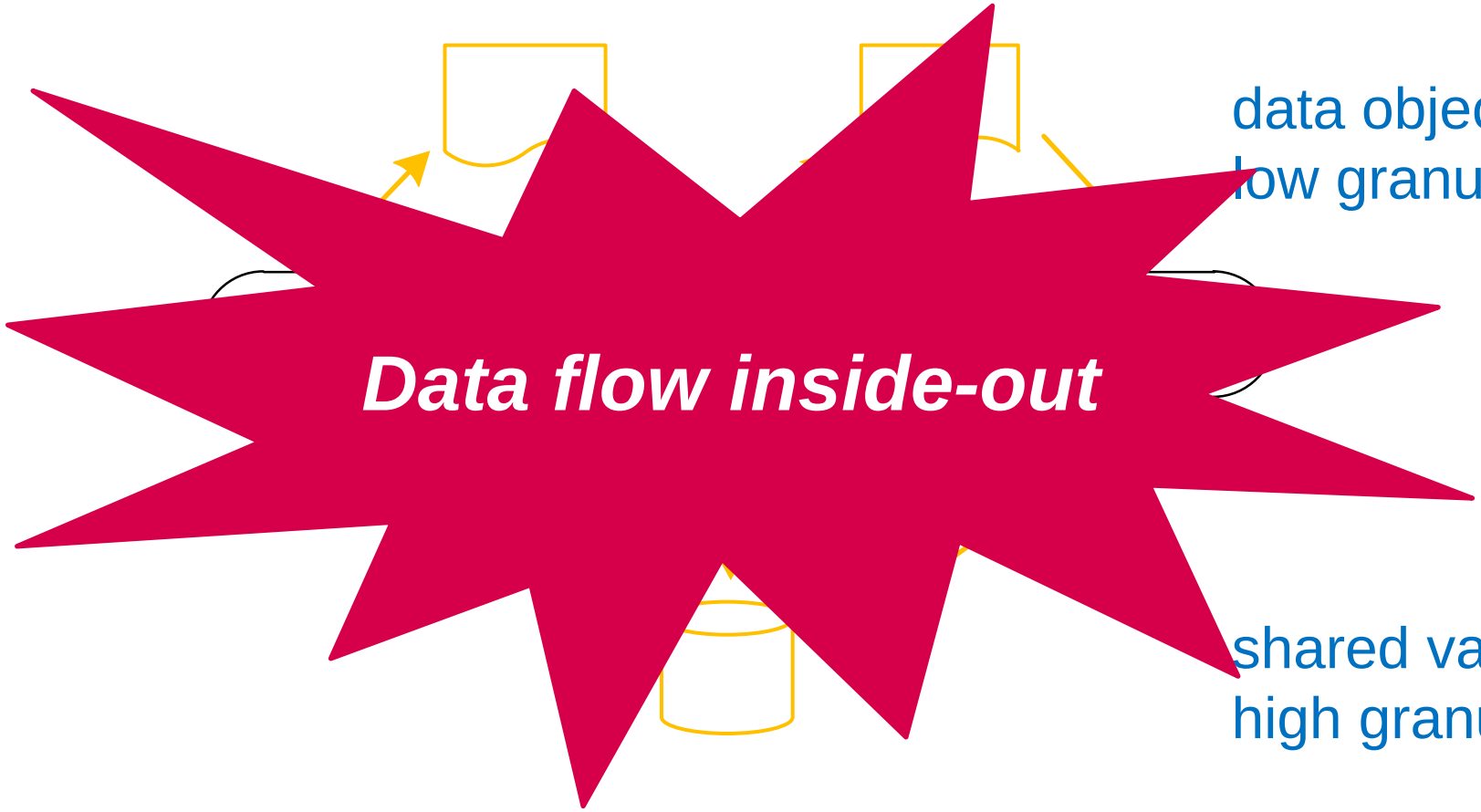
Three types of data flow



data objects:
low granularity

shared variables:
high granularity

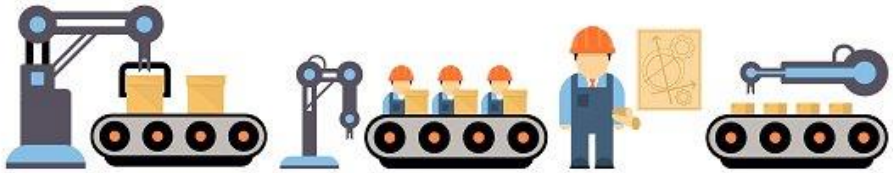
Three types of data flow



data objects:
low granularity

shared variables:
high granularity

From structured to semi-structured processes



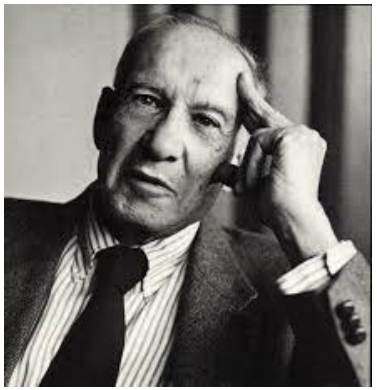
Business processes become …

- **Driven by data**
- **Performed by knowledge workers**
- **Goal-oriented**
- **Adaptive**
- **Unpredictable**

Requires novel verification techniques

Structured vs semi-structured processes: Taylor versus Drucker on work

- Scientific management
- Standardize processes to increase efficiency



- Management by objectives
- Participants choose actions to meet goals

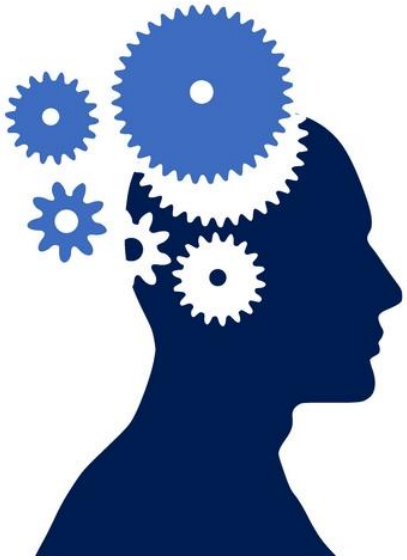
Routine work can be analyzed and a common pattern derived … it can be automated by traditional process automation means.

Knowledge work … does not have the level of repeatability found in routine work. When it comes to work automation, any advantage gained from similarities is overwhelmed by additional costs of having to accommodate the differences.

K. D. Swenson, *Mastering the Unpredictable*. Meghan-Kiffer Press, 2010

Knowledge-intensive Processes (KiPs)

- Performed by knowledge workers (experts)
- Includes knowledge-intensive decision making
- Driven by data and knowledge
- Progressing towards objectives
- Adaptive

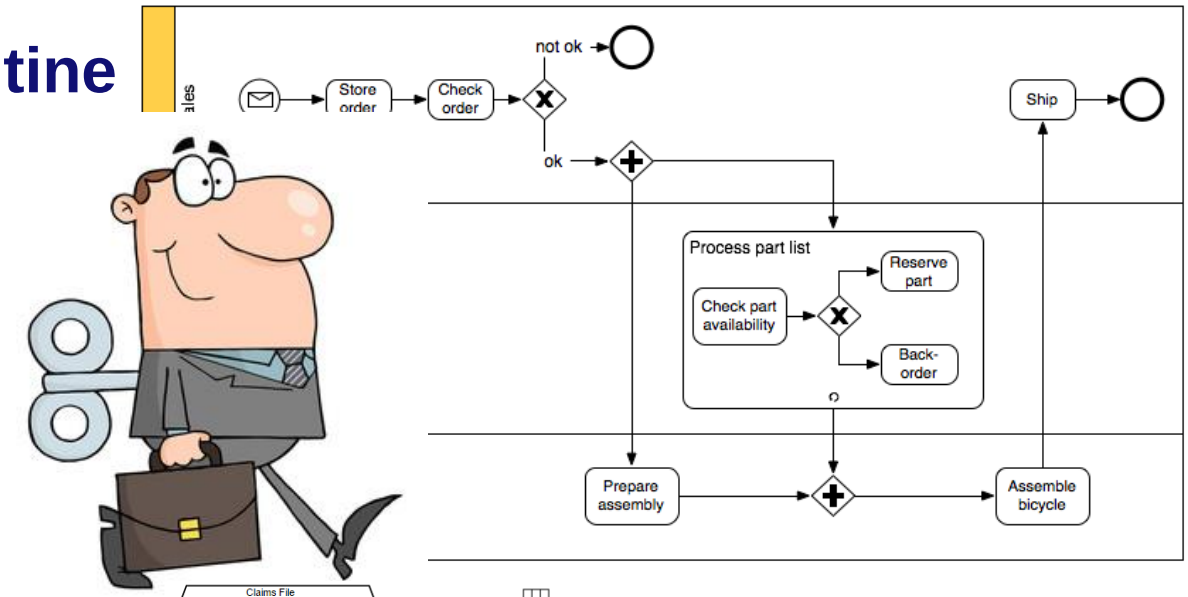


Di Ciccio, Marrella, Russo: *Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches*. J. Data Semantics 4(1): 29-57 (2015)

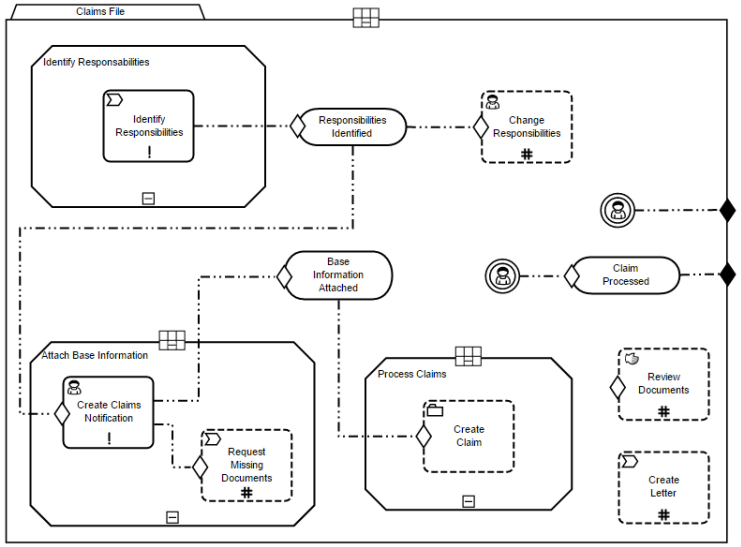
Vaculín, Hull, Heath, Cochran, Nigam, Sukaviriya: *Declarative business artifact centric modeling of decision and knowledge intensive business processes.* EDOC 2011: 151-160

Routine vs knowledge-intensive process

- **Routine**



- **KiP**



Case Management is …

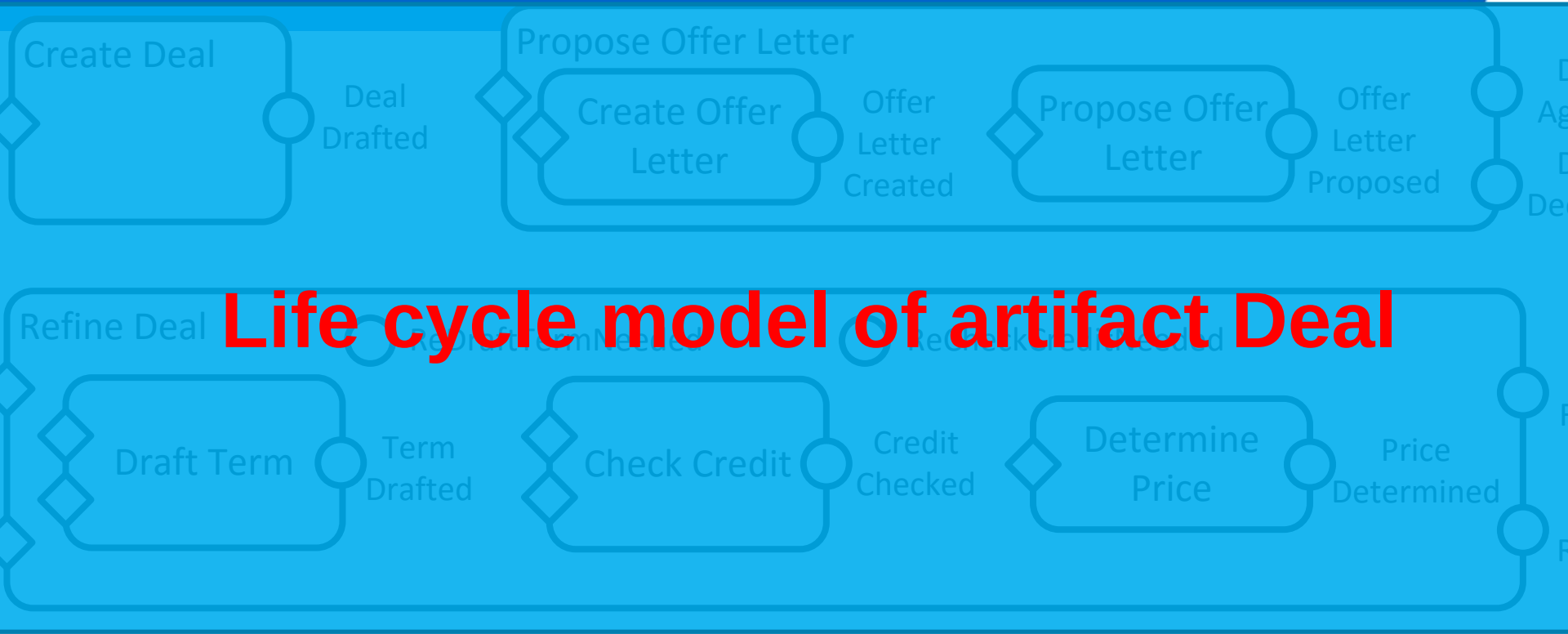
- the management of long-lived collaborative processes that coordinate
  - knowledge,
  - content,
  - correspondence and
  - resources
- to progress a case to achieve a particular goal;
- where the path of execution cannot be predetermined in advance of execution;
- where human judgment is required to determine how the end goal can be achieved;
- and where the state of a case can be altered by external out-of-band events.

Michael White: Case Management: Combining Knowledge With Process, BP Trends 2009

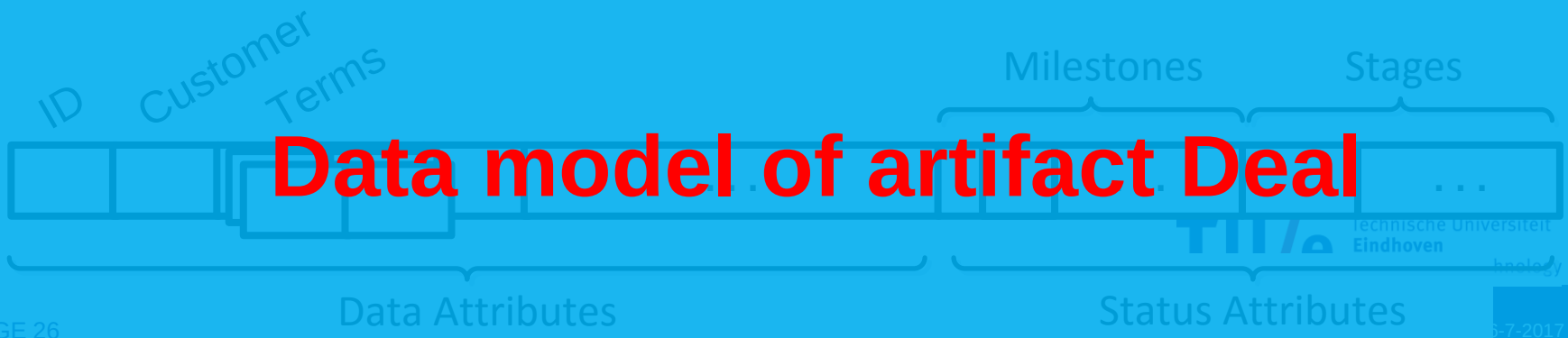
Artifact-centric process models (data-flow inside out)

- **Business artifact: key conceptual business entity used in business operation**
- **Business artifact specified by integrated …**
  - **data model**
  - **life cycle model**
- **Example artifacts: parcel, order, patient, …**
- **Languages**
  - **IBM Guard-Stage-Milestone schemas**
  - **OMG Case Management Model and Notation (CMMN)**

Guard-Stage-Milestone (GSM) schemas

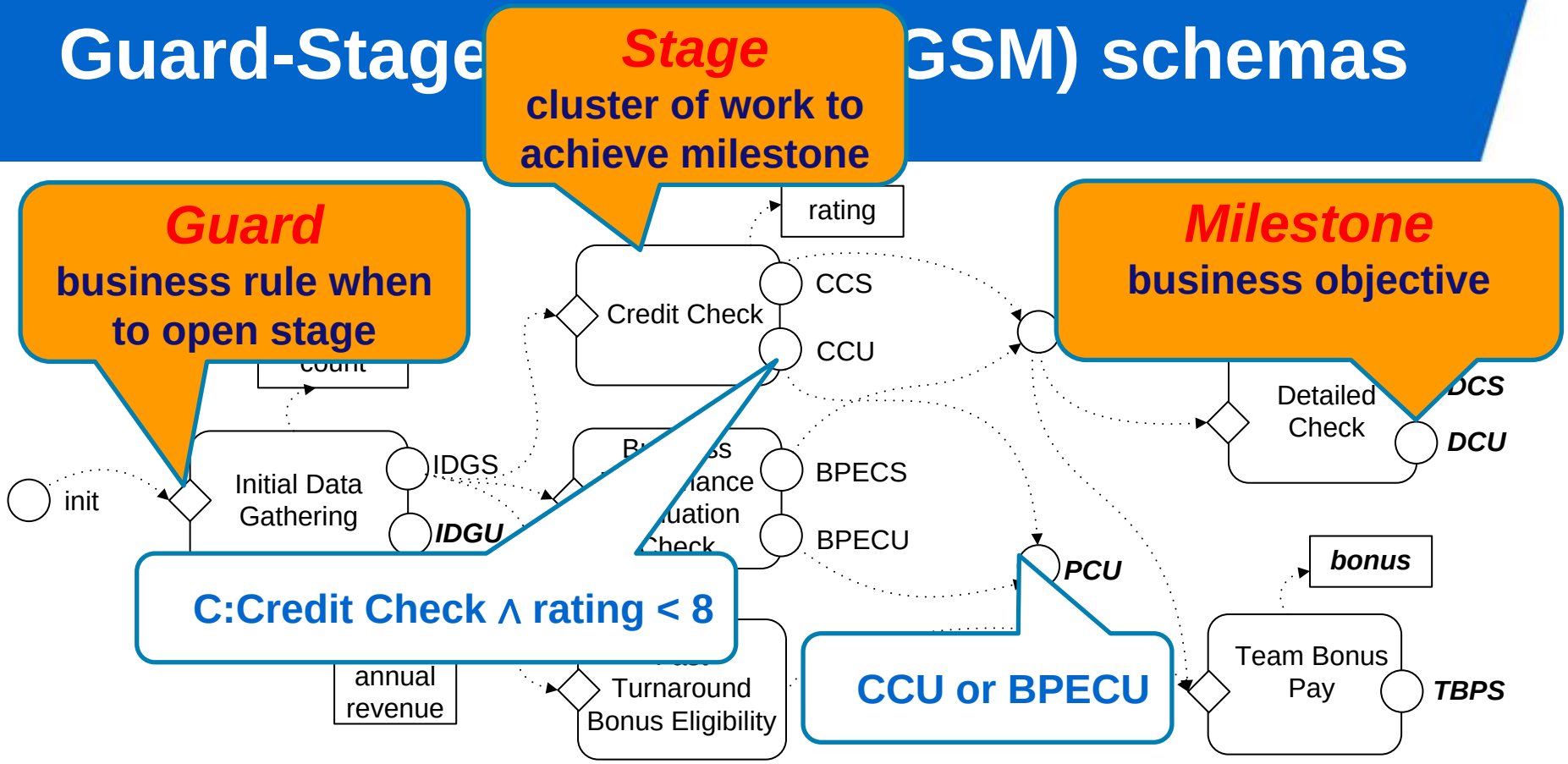


Life cycle model of artifact Deal



Data model of artifact Deal

Guard-Stage (GSM) schemas



- **Business rules specify when stage/milestone changes state**
  - **Might refer to status of other stages/milestones (besides attrib)**
- **Rules need to be evaluated in right order to ensure that all changes have maximal effect**
- **Unit of change triggered by external event is called a B-step**

Two challenges for Case Management (and how they impact verification)

- **How to design Case Management schemas?**
- **How to manage changes in Case Management schemas?**

Challenge 1: How to design Case Management schemas?

- **Blank sheet (lot of work)**
- **Modify template (but who designs template?)**
- **Start from classical process model**

Computing
April 2016, Volume 98, Issue 4, pp 345–373

Synthesizing data-centric models from business process models

Authors | Authors and affiliations

Rik Eshuis, Pieter Van Gorp

Open Access | Article

First Online: 14 February 2015

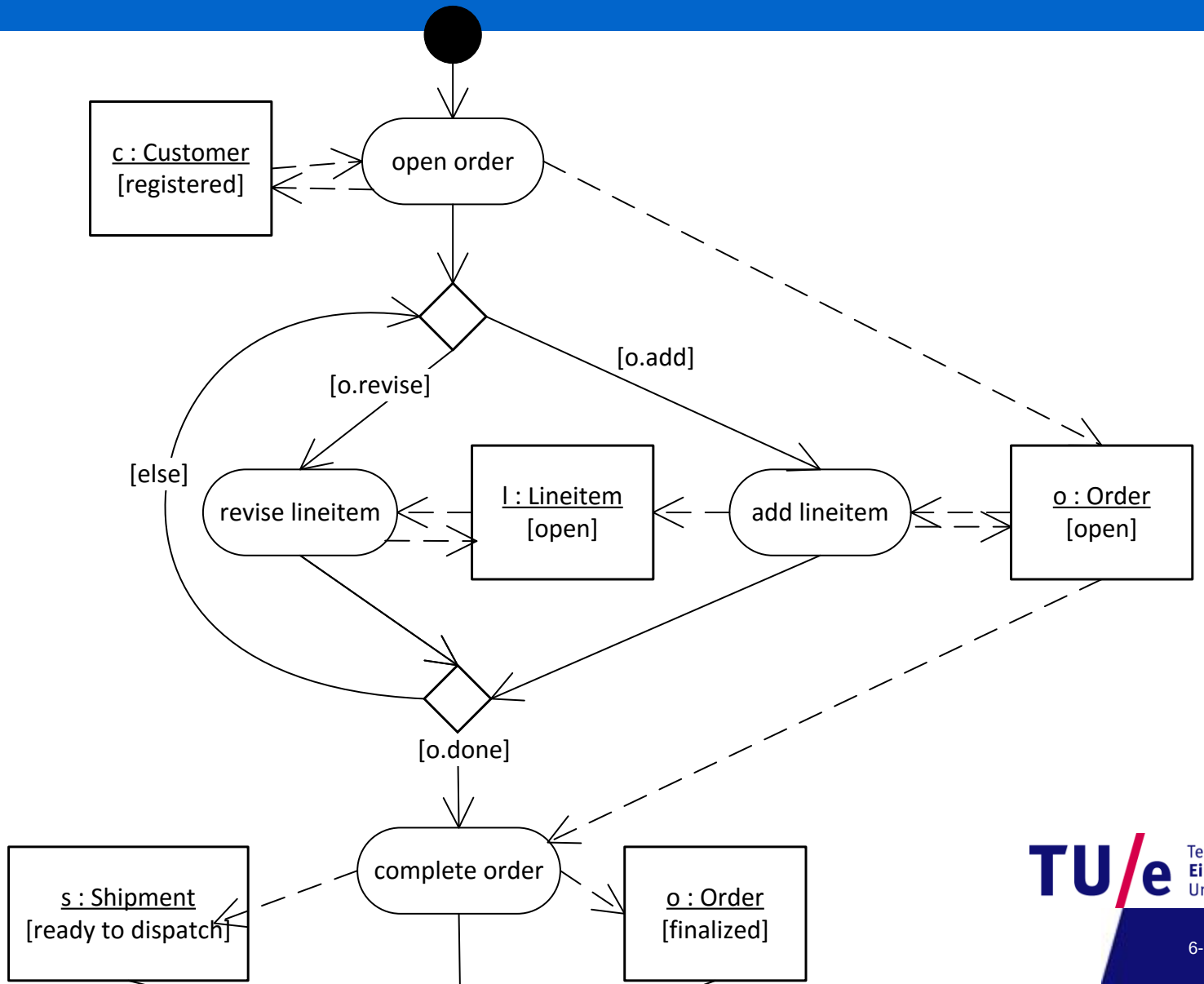
DOI: 10.1007/s00607-015-0442-0

Cite this article as:
Eshuis, R. & Van Gorp, P. Computing (2016) 98: 345. doi:10.1007/s00607-015-0442-0

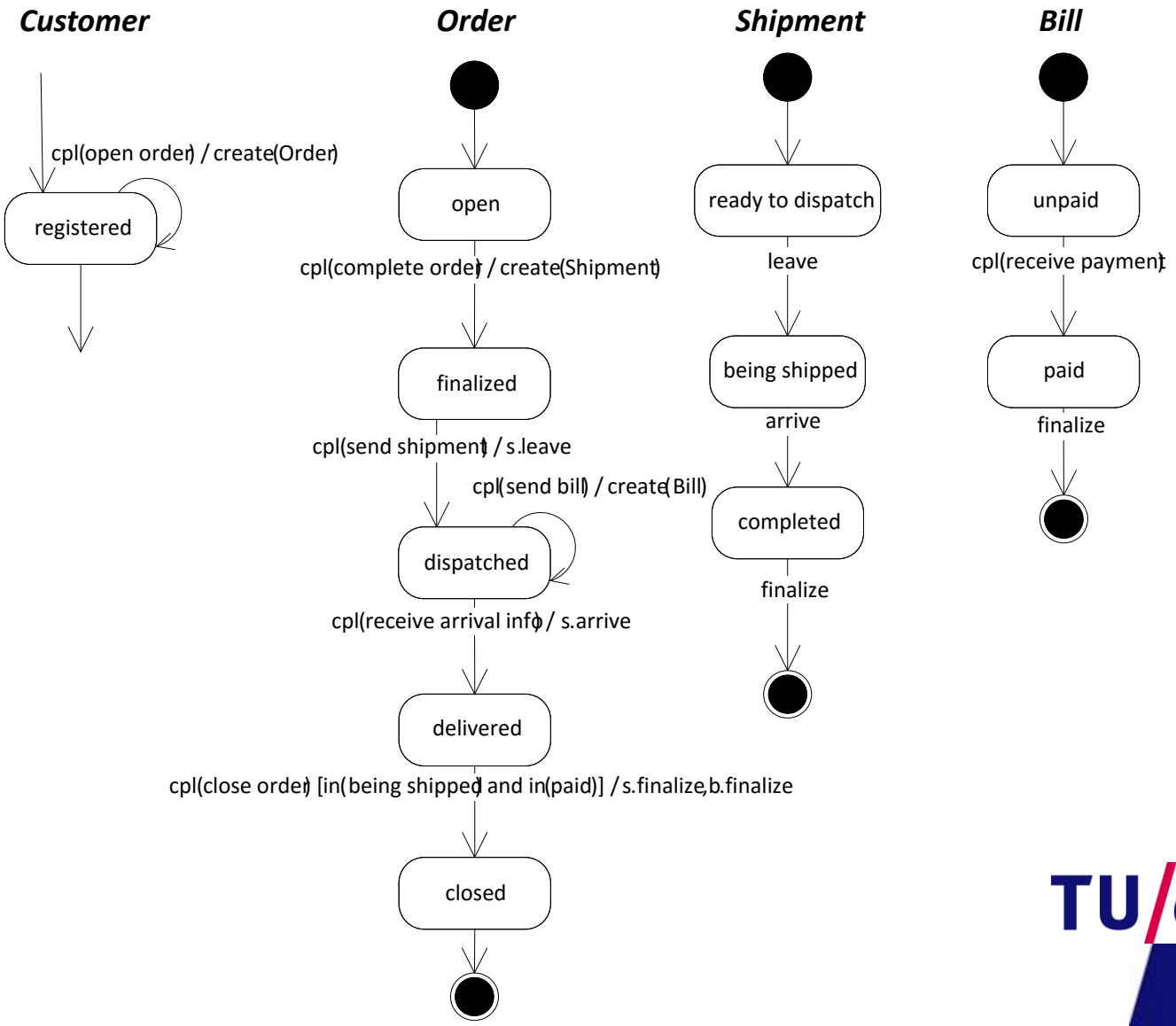
Start from classical process model

1. Create process model with object flows for default behavior
2. Transform into object-centric model (automated)
3. Transform into GSM schema (automated)
4. Extend GSM schema with exceptional behaviour

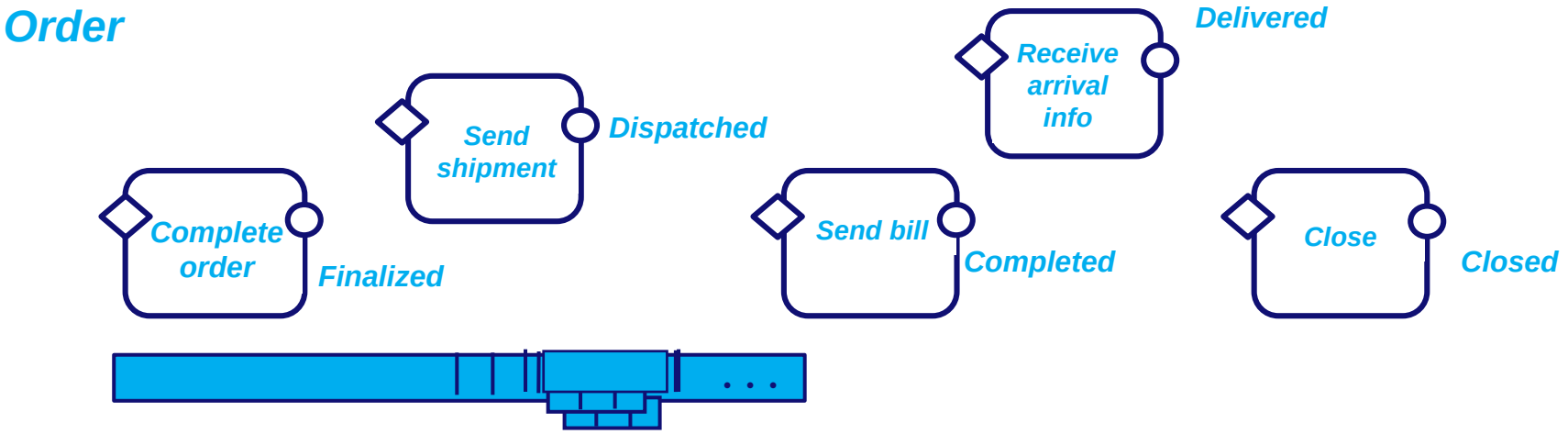
Part of process model



Object-centric model



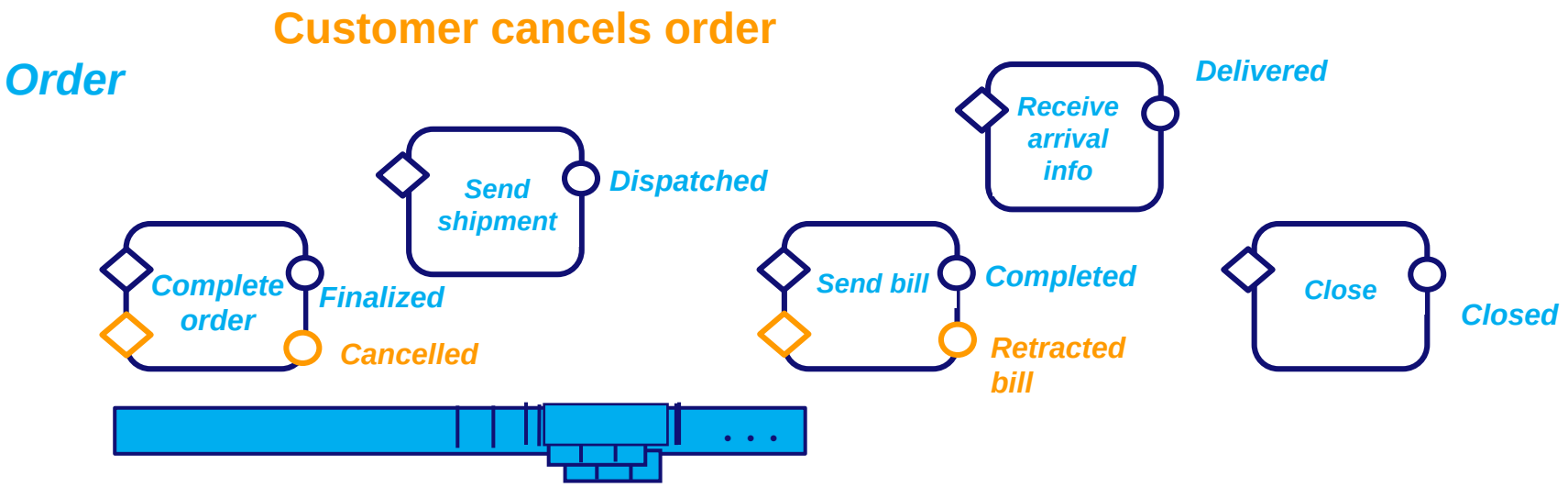
Guard-Stage-Milestone model



Stages	Guard rules	Terminating rules
complete order	ON +open	ON cpl(complete order)
send shipment	ON +finalized	ON cpl(send shipment)
send bill	ON +d	
receive arrival info	ON +se	
close order	ON +d	

Milestones	Achieving rules	Invalidating rules
open	IF true	ON +finalized
finalized	ON -complete order	ON +dispatched
dispatched	ON -send shipment	ON +delivered
send bill completed	ON -send bill	ON -dispatched
delivered	ON -receive arrival info	ON +closed
closed	ON -close order	

Extending Guard-Stage-Milestone model with exceptions



Stages	Guard rules	Terminating rules
complete order	ON +open	ON cpl(complete order)
send shipment	ON +finalized	ON cpl(send shipment)
send bill	ON	
receive arrival info	ON	
close order	ON +d	

plus additional rules …

les	Invalidating rules
open IF true	ON +finalized
ON -complete order	ON +dispatched
ON -send shipment	ON +delivered
ON -send bill	ON -dispatched
ON -receive arrival info	ON +closed
ON -close order	

Verification challenges for Case Management

- **Verify that key properties of default process are preserved in case management schema**

Challenge 2: How to manage changes in Case Management schemas

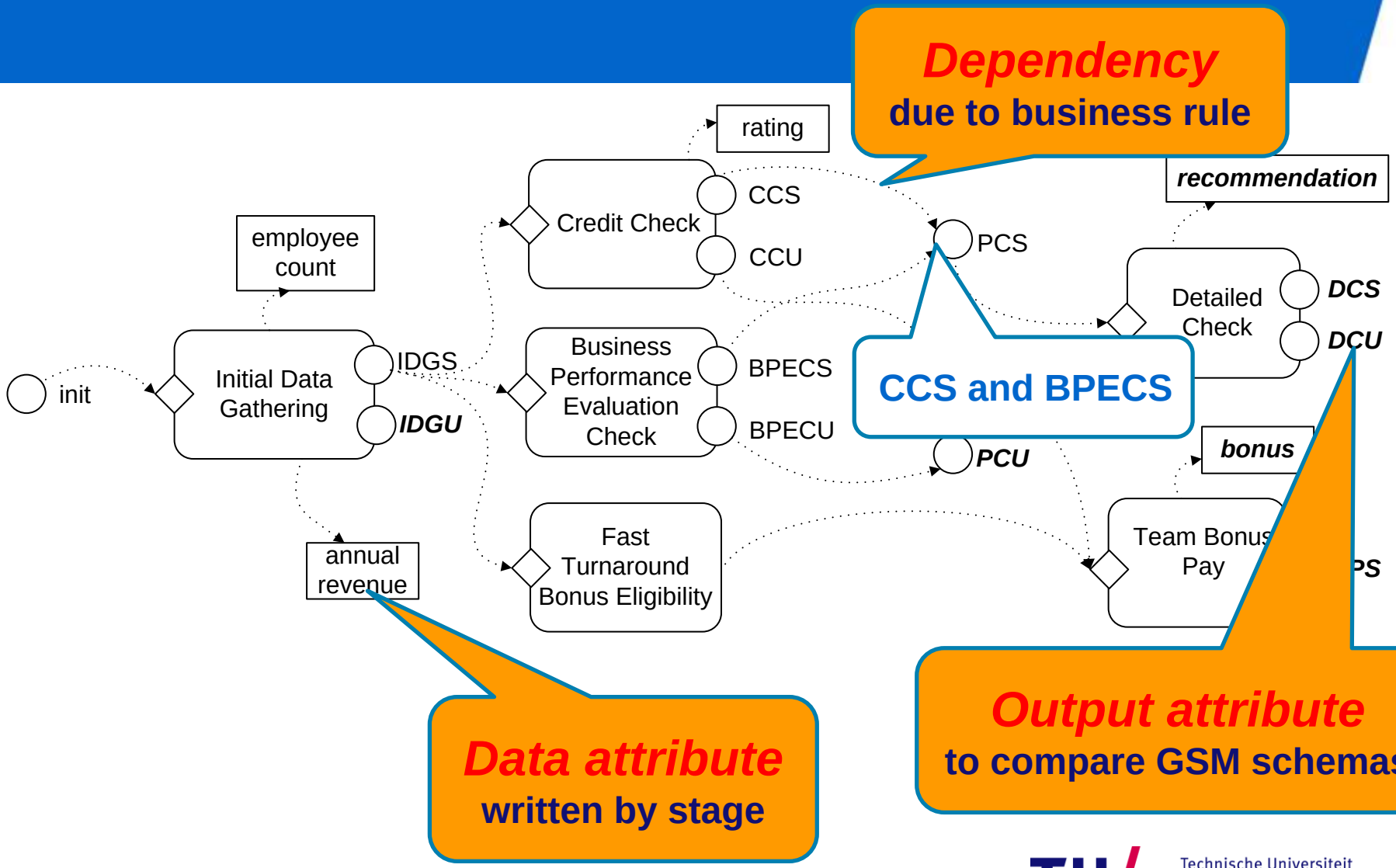
- Case Management schemas are by their very nature constantly changing
  - Adaptive Case Management

How to define:

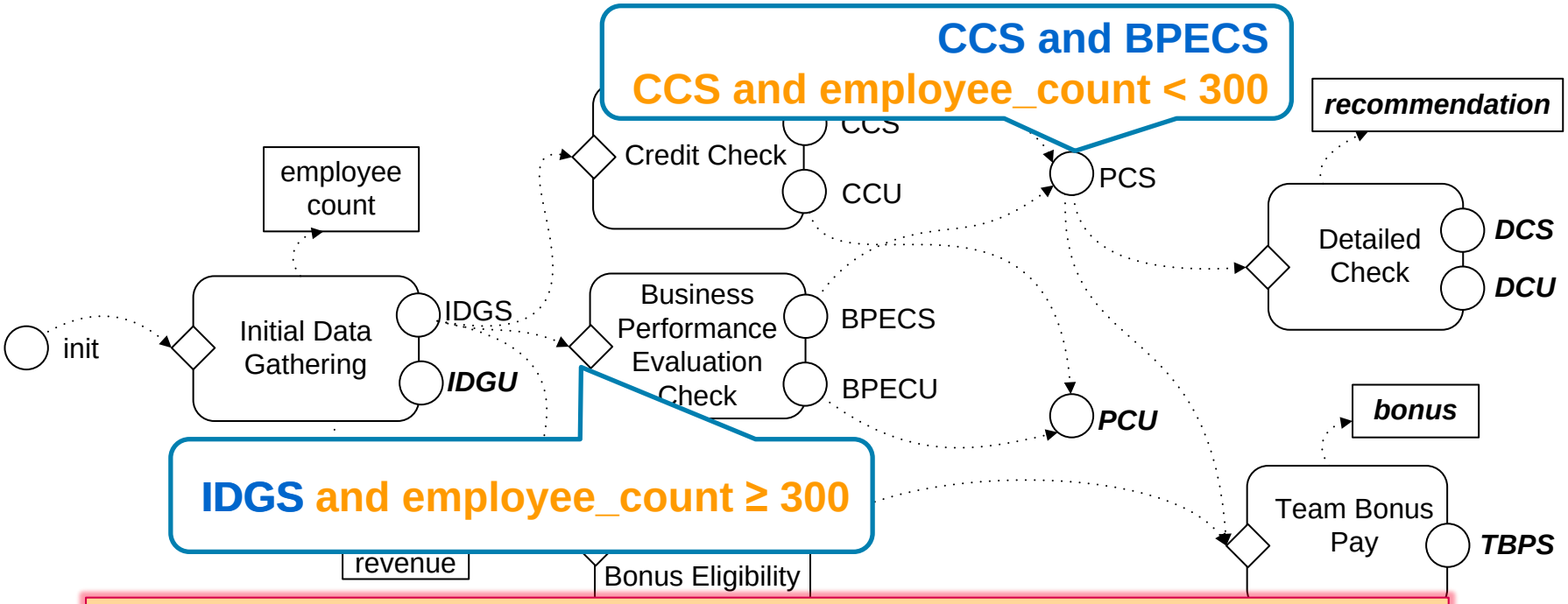
- Conditions under which Case Management schemas can be changed, while preserving properties
- Change operations that guarantee preservation of properties
  - Rule modification, delete/insert stage/milestone

Rik Eshuis, Richard Hull, Mengfei Yi: Property Preservation in Adaptive Case Management. ICSOC 2015: 285-302

Guard-Stage-Milestone (GSM) schemas



Change example 1: modifying rules



"I have only proved it correct, not tried it."

Change assertion:

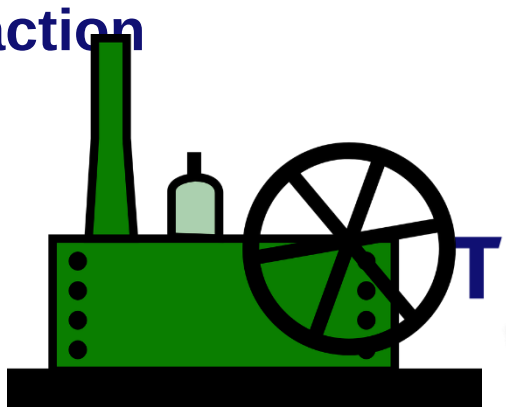
if employee_count ≥ 300, both schemas are equivalent

Verification challenges for Case Management

- Verify that key properties of default process are preserved in case management schema
- Verify that changes in case management schemas preserve user-defined properties

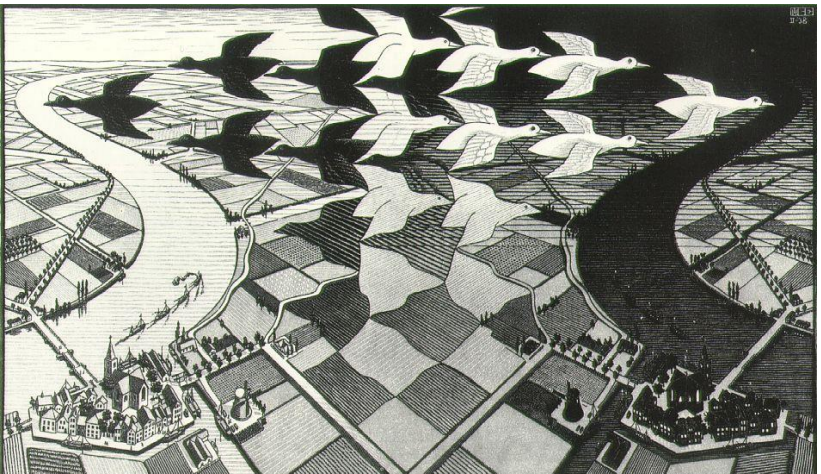
Thinking a bit further

- Instant verification
  - Run-time
  - Zero-time
  - Automated data abstraction
- Resource view
  - Policies
  - Performance analysis



Conclusion

- **Process verification from single to multi-view:**
  - **Incorporating data flow**
  - **Data flow inside-out**
- **Challenges**
  - **Run-time instant verification data-driven processes**
  - **Incorporating resource view**



Thank you!


TU/e
Technische Universiteit
Eindhoven
University of Technology
Where innovation starts
Dr.ir. Rik Eshuis
Assistant Professor Information Systems
School of Industrial Engineering
http://is.ieis.tue.nl/staff/heshuis
h.eshuis@tue.nl


Q
&
A