Chapter 5 - Process Discovery

The Setting of Process Discovery

- **Process Discovery**: act of gathering information about an existing process and organising it in terms of an as-is process model.
  - **4 tasks of process discovery**
    1. **Define the setting**: assemble a team that will be responsible for the managing process.
    2. **Gather information**: build an understanding of the process by using different discovery methods.
    3. **Conduct the modeling**: do the actual modeling.
    4. **Assure model quality**: guarantee that the resulting model meets different quality criteria.
    - 3 and 4 are done together.
    - Once the setting has been defined (1) the remaining three tasks are performed in an iterative manner.

Process Analyst VS Domain Expert

- **Process Analyst**: responsible for gathering information about business processes and driving the modeling task under the leadership of the process owner. Must be familiar with process modeling languages like BPMN and be skilled at gathering process related information.

> *Example 5.1* Let us consider the following two modeling tasks:
>
> - Modeling the process for ordering books through an online bookstore, from the perspective of the customer.
> - Modeling the same process from the perspective of the bookstore.

- The 1st modeling task is easy, we have done this in the previous chapters.
- The 2nd task is different because we are only able to complete it if we have worked for an online bookstore, which is less common.
  - Process analysts are supposed to model business process which they have experienced neither as process participants or customers.
    - Must gather a lot of information about the process in order to understand how it works from the inside by consulting **domain experts**.
- **Domain Expert**: individual that has intimate knowledge of how a process or specific tasks within that process are performed. Can be a process participant, process owner or an operational manager who coordinates a team of process participants.
  - **Partners, suppliers or customers** can also be seen as domain experts as they provide a complementary view on the process.
  - Domain experts are not experienced with modeling languages.

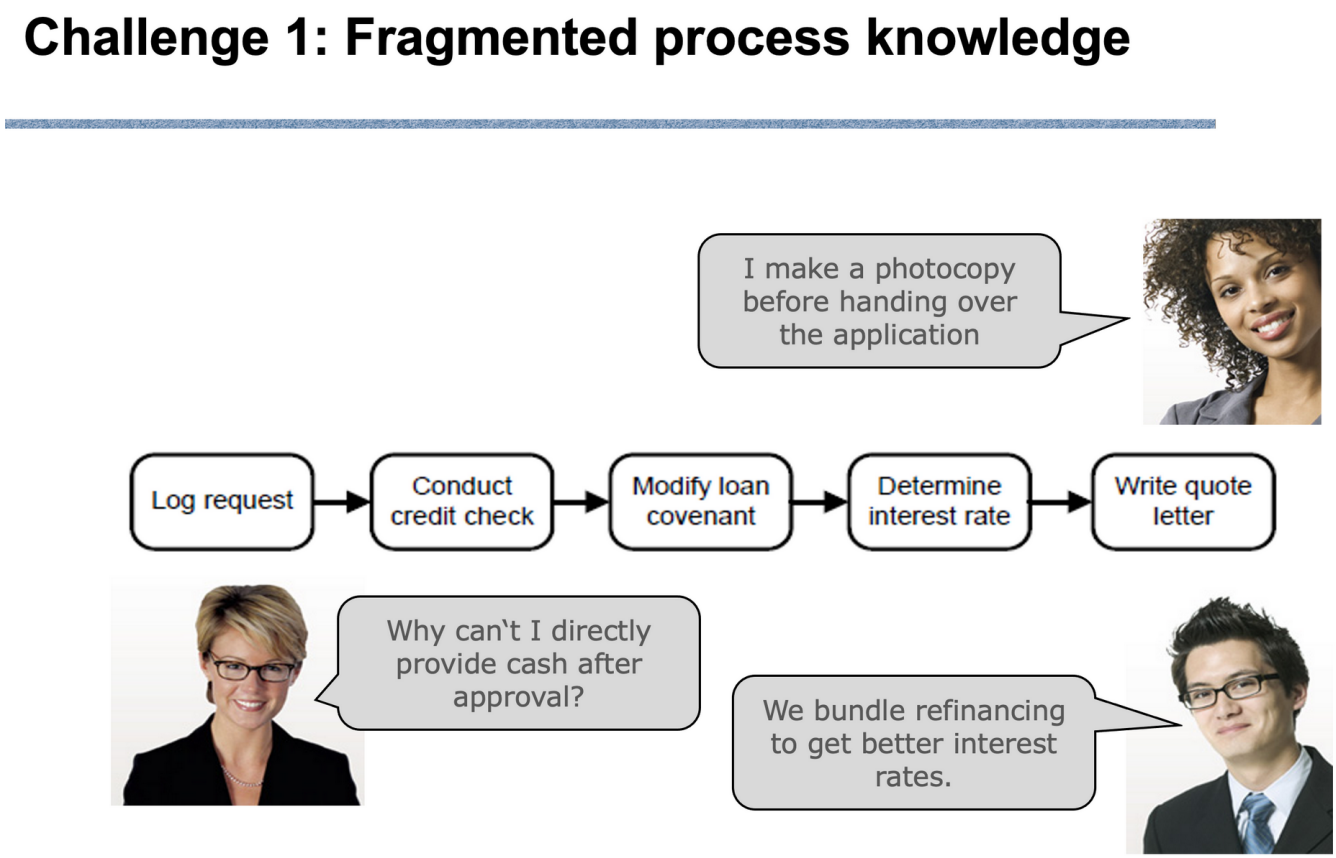
    - It is an advantage when BPM consultants who also have domain expertise can be assigned to process modeling projects.

- It is the **task of the process owner** to secure commitment and involvement of both analysts and domain when defining the setting of process discovery.

3 Challenges of Process Discovery

Fragmented Process Knowledge

Due to **specialisation** and division of labor, the tasks of a process will be performed by different resources. A process analyst must talk with with different domain experts who may have an abstract understanding of the overall process and a very detailed understanding of they're own tasks. **Different domain experts will have different views of the process**, therefore it makes it hard to unite the different views together. This is the reason why process discovery requires many **iterations**.

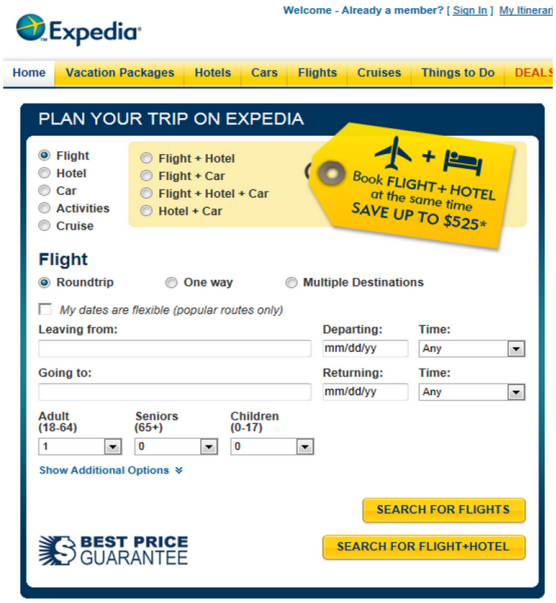


Domain experts think on instance level

Domain experts easily describe the tasks they conducted for a specific process instance but if you ask them general questions about how the process works they will find it difficult. Therefore, process analysts must ask questions on specific aspects of the process for the domain experts.
Ex: What happens if certain conditions do or do not hold, if a certain outcome is achieved or if certain deadlines are not met.

Process analysts then reverse engineer the abstraction from instance level to process level.

Challenge 2: Domain experts think on instance level



"Every trip is different"

"You cannot really compare. Our customers go to different places in different seasons using different modes of transportation"

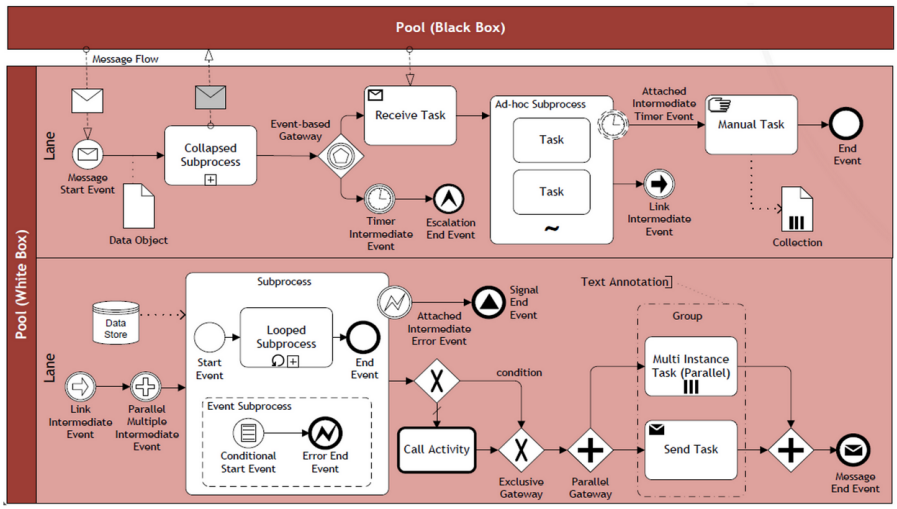
"We can never do anything exactly in the same way. There are so many special conditions"

Knowledge of process modeling is uncommon

Domain experts typically don't know modeling languages so the process analyst must **translate** the model to natural language so the domain expert can point out flaws or ask for clarifications.

Challenge 3: Knowledge of process modeling is uncommon

"Does this diagram correctly show your process?"



What makes a good process analyst? - Profile of an Expert Process Analyst

- Getting the right people on board
- Formulate and test hypotheses of how the process is structured at different levels of detail
- Identify patterns in the information provided by domain experts
- Pay attention to model quality

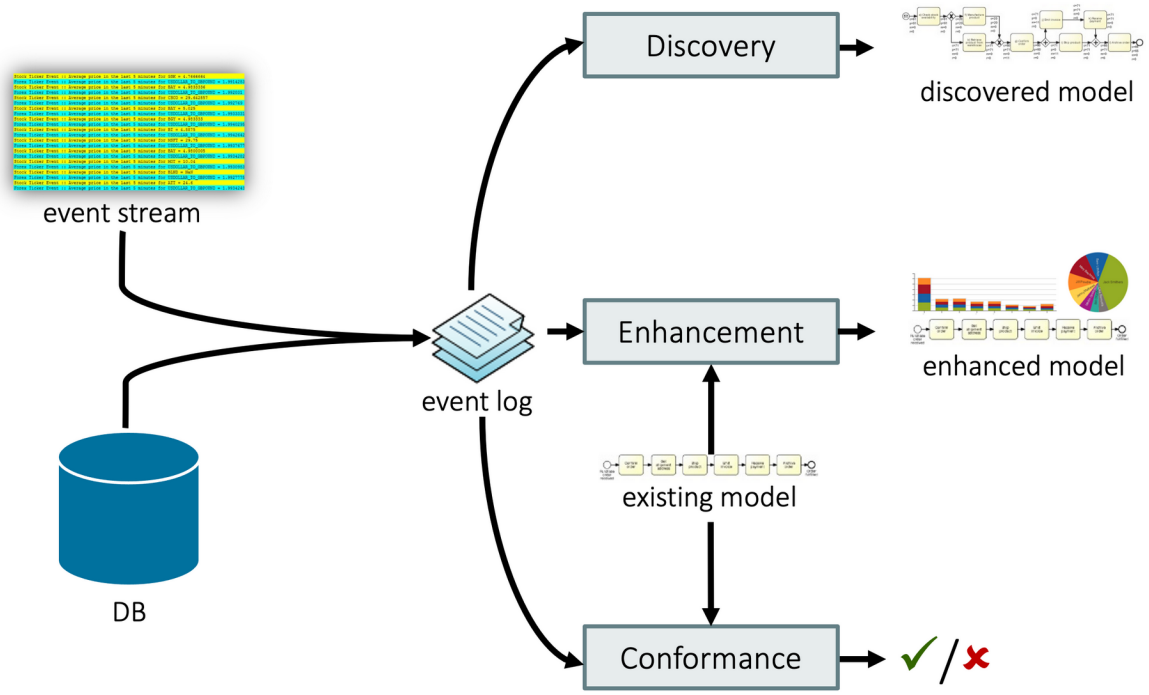
Process Discovery Methods

Evidence Based Discovery (3 methods)

- **Document Analysis**
  - There is usually documentation available relating to an existing business process.
  - **Issues**:
    - Available documentation about the operations of a company is not readily organised in a process-oriented way.
    - The level of granularity of the documentation may not be appropriate.
      - May need to abstract from or refine the information in these documents.
    - Documents are only partially trustworthy as many documents do not necessarily capture reality (some of them may be outdated or describe things idealistically).
  - **Advantages**:
    - Process analyst can use the documentation to get familiar with certain parts of a process and its environment and also to formulate hypotheses.
    - Useful before talking to a domain expert.
- **Observation**
  - Processing of individual cases in order to understand how the process works.
  - Types of roles a process analyst can play:
    - **Active Customer Role**: we trigger the execution of a process and record the steps that are executed and the set of choices that are offered.
      - Ex: create a book order and keep track of which activities are performed on the side of the retailer.
      - Provides good understanding of the process and its milestones, however we only see parts that require customer interaction.
    - **Passive Observer**: appropriate for understanding the end-to-end process but requires access to the people and sites where the process is performed.
      - **Advantage**: reveals how a process is conducted in reality as of today (document analysis captures the past)
      - **Disadvantage**: not all sites are accessible and people may act different because they know they are being observed.

- **Automated Process Discovery**
  - Method that uses event logs.
    - **Advantages**:
      - Event logs capture the actual execution of business processes.
      - Event logs record a rich set of process related info beyond the tasks that have been performed (ex: task timestamps).
      - Method not limited by the confines of an enterprise system can be used to reconstruct end to end processes.
    - **Disadvantage**: event logs are not always available and when they are sometimes only record some tasks.

Automated process discovery and Process mining



Automated discovery: Minimum data requirements

- Activity name and timestamp
- Reference to case id

case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...

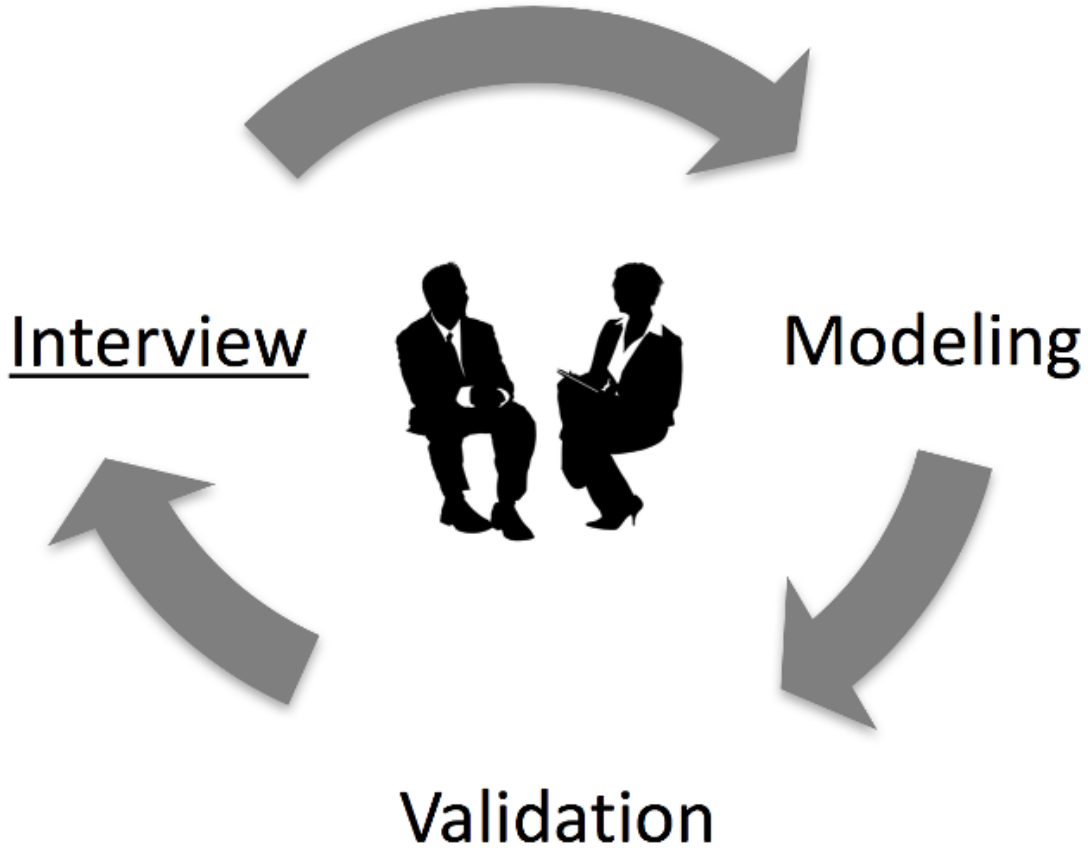
Additional information:

- Activity resource, cost
- Case attributes (e.g. customer reference, type of case…)

Interview-Based discovery

- Interviewing domain experts to inquire about how a process is executed.
- We assume analyst and domain expert share terminology.
- **Approaches**:
  - **Forward**: starting from the triggers we proceed forward until we reach the outcomes.

        - Helpful for understanding which decisions are taken at which stage.
    - **Backward**: starting from the process outcomes (ex: an order being fulfilled) and work our way backwards until we reach process triggers (ex: the receipt of a purchase order).
    - **Structured**: validate our hypotheses but create in the interview a feeling of running through a checklist (if you're not asked about that info, you're gonna hold back on it).
    - **Unstructured/Free-form**: enable domain experts to discuss the process at a level of detail they find appropriate uncovering disregarded aspects of the process
- Overall, we should use questions that aim to identify **exceptional behaviour** (tends to be neglected).



Workshop-based Discovery

- Gather all key stakeholders together.
- Participants interact to create shared understanding.
- Typical one process analyst (facilitator), multiple domain experts, process owner may also attend.
- May be software supported - a model is directly created during the workshop (typically a separate role - tool operator).
- Model is used as a reference point for discussions.
- **Alternative**: use brown-paper workshops.
- Usually 3 to 5 half-day sessions.

Discovery and Culture

Before starting with process discovery, it is important to understand the culture and sentiment of an organisation.

Example: Any difference in discovery?

Consider the following two companies:

- **Company A** is young, founded three years ago, and has grown rapidly to a current toll of one hundred employees.
- **Company B** is owned by the government and operates in a domain with extensive health and security regulations.

How might these different characteristics influence a workshop-based discovery approach?

- There are companies that practice an **open culture** in which employees are encouraged to utter their ideas and their criticism. Such organisations can benefit a lot from workshops as participants are likely to present their ideas freely.
- In **strictly hierarchical organisations**, it is necessary to take special care that every participant gets an equal share of parole in a workshop and that ideas and critique are not held back.

It might be the case that the young dynamic company has a more open culture than the company with extensive health and security regulations. This has to be taken into account when organising a workshop.

Discovery methods: strengths and weaknesses

Aspect	Evidence-based	Interviews	Workshops
Objectivity	High	Medium-high	Medium-high
Richness	Medium	High	High
Time consumption	Low-medium	Medium	Medium
Immediacy of feedback	Low	High	High

Discovery methods: strengths and weaknesses

Method	Strength	Weakness
Document Analysis	• Structured information • Independent from availability of stakeholders	• Outdated material • Wrong level of abstraction
Observation	• Context-rich insight into process	• Potentially intrusive • Stakeholders likely to behave differently • Only few cases
Automatic Discovery	• Extensive set of cases • Objective data	• Potential issue with data quality and level of abstraction
Interview	• Detailed inquiry into process	• Requires sparse time of process stakeholders • Several iterations required before sign-off
Workshop	• Direct resolution of conflicting views	• Requires availability of several stakeholders at the same time

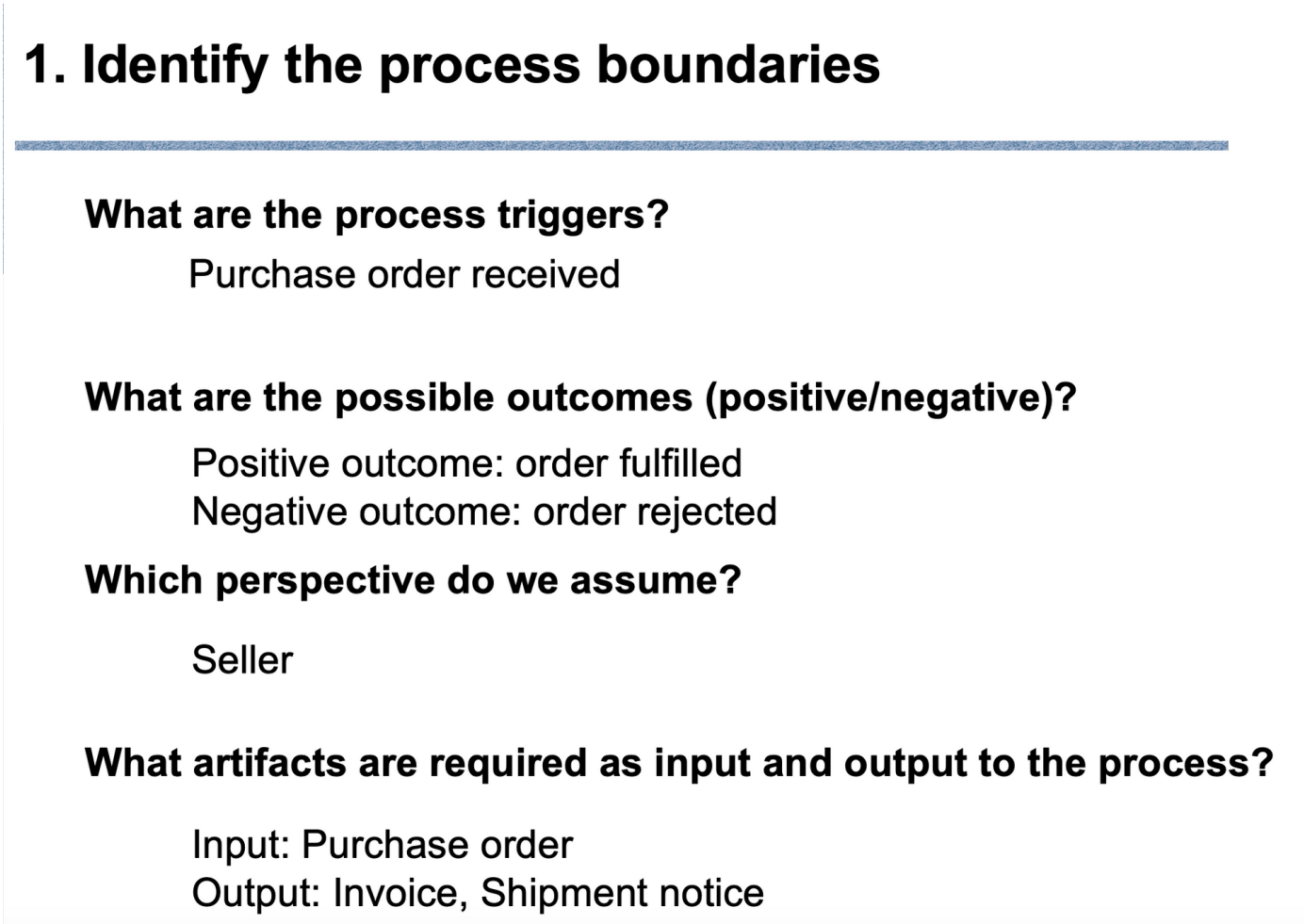
In what situations is it simply not possible to use one or more of the described discovery methods?

- Direct observation not possible if the process partially runs in a remote or dangerous environment.
- Documentation may not exist
- Automated process discovery based on event log data may not be possible due to the lack of input.
- Generally, interviews always possible
- Workshop based discovery critical in strictly hierarchical companies with a non-open culture.

Process Modeling Method

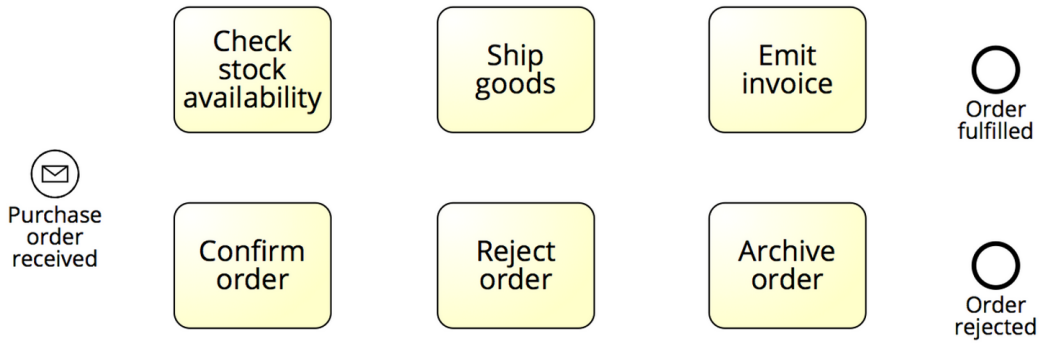
Stepwise method to conduct the modeling

1. Identify the process boundaries



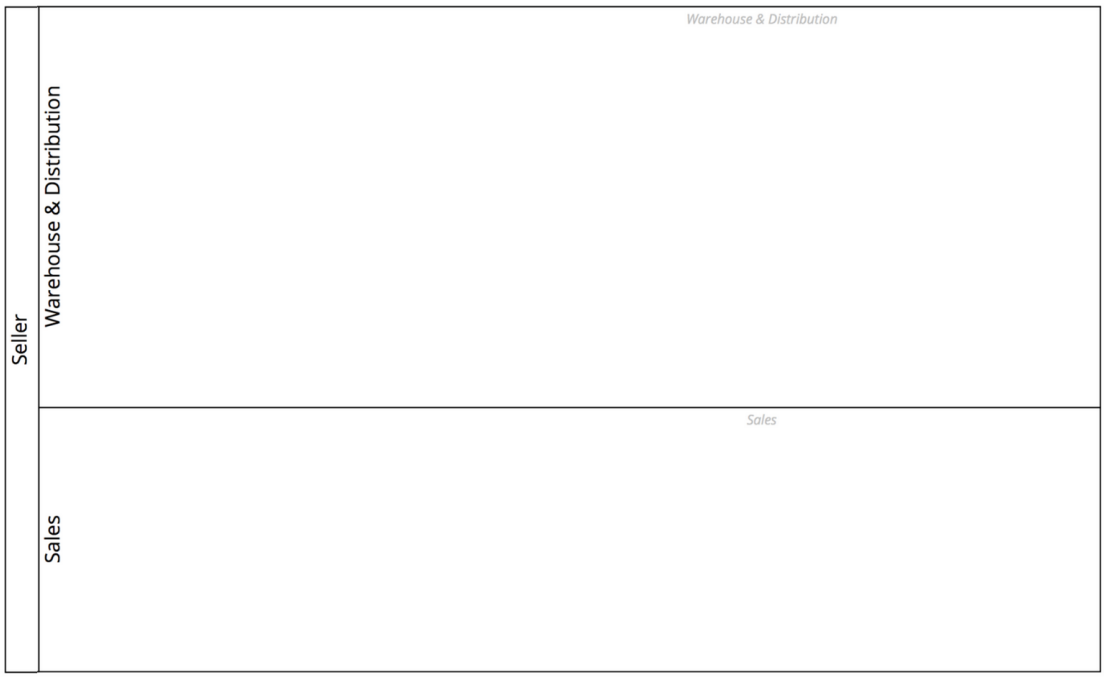
2. Identify activities and events

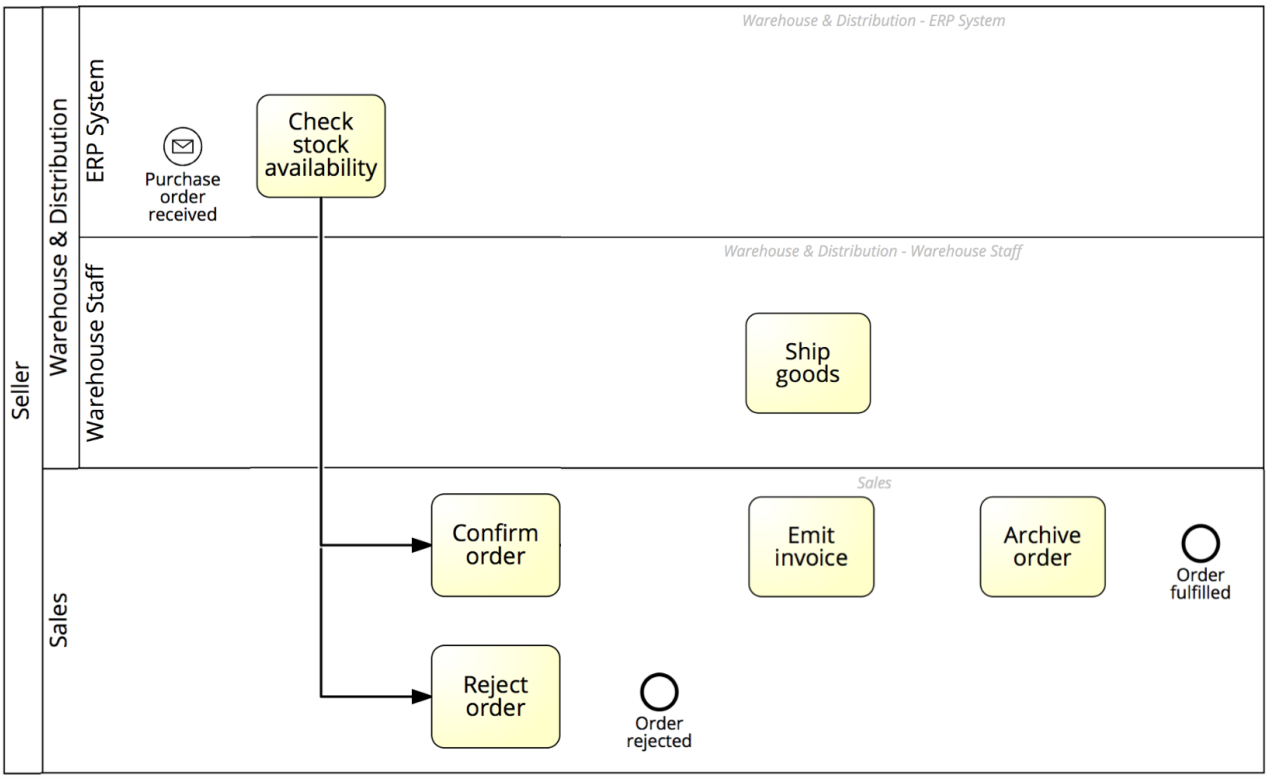
2. Identify activities and events

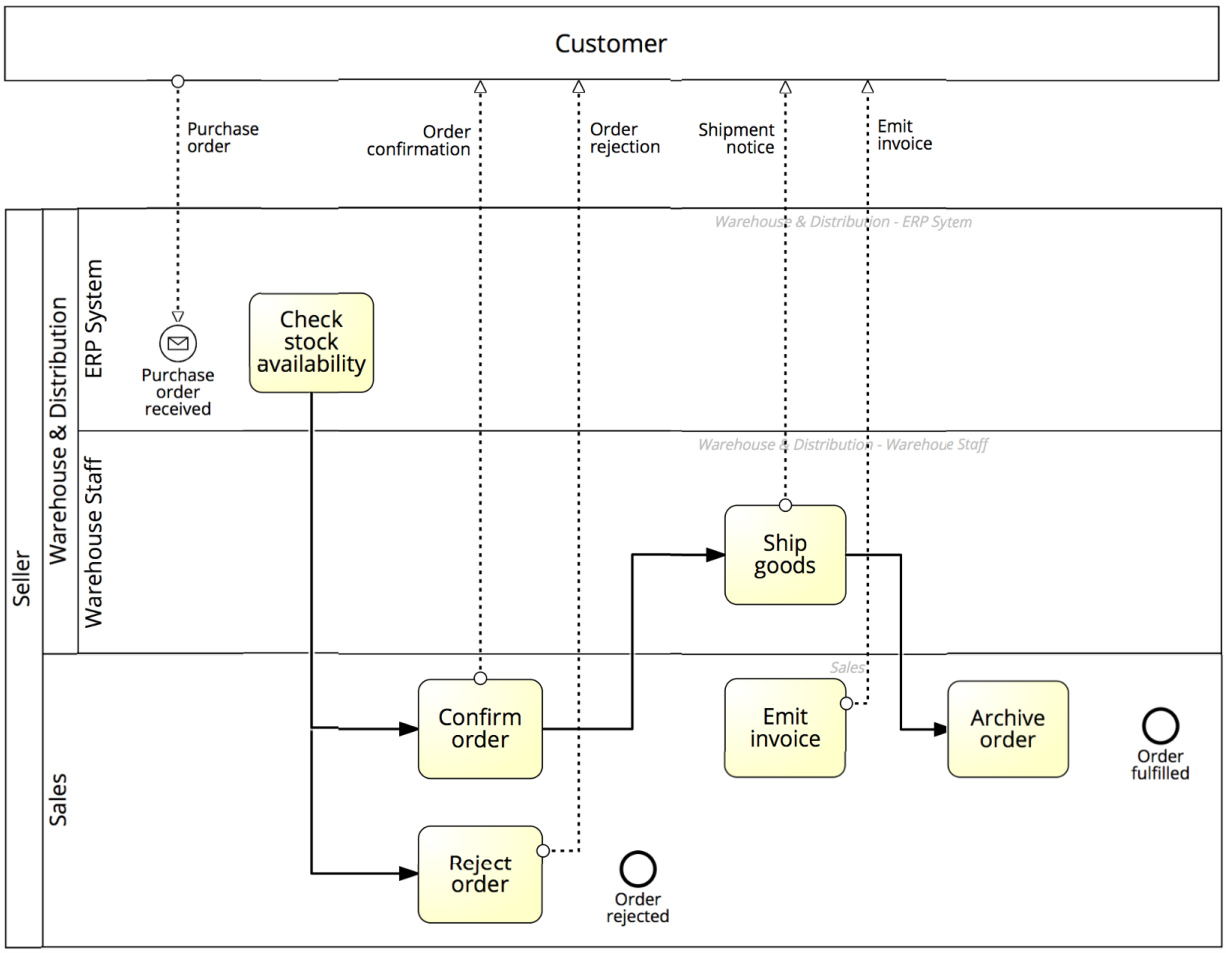


3. Identify resources and their handoffs

3. Identify resources and their handoffs

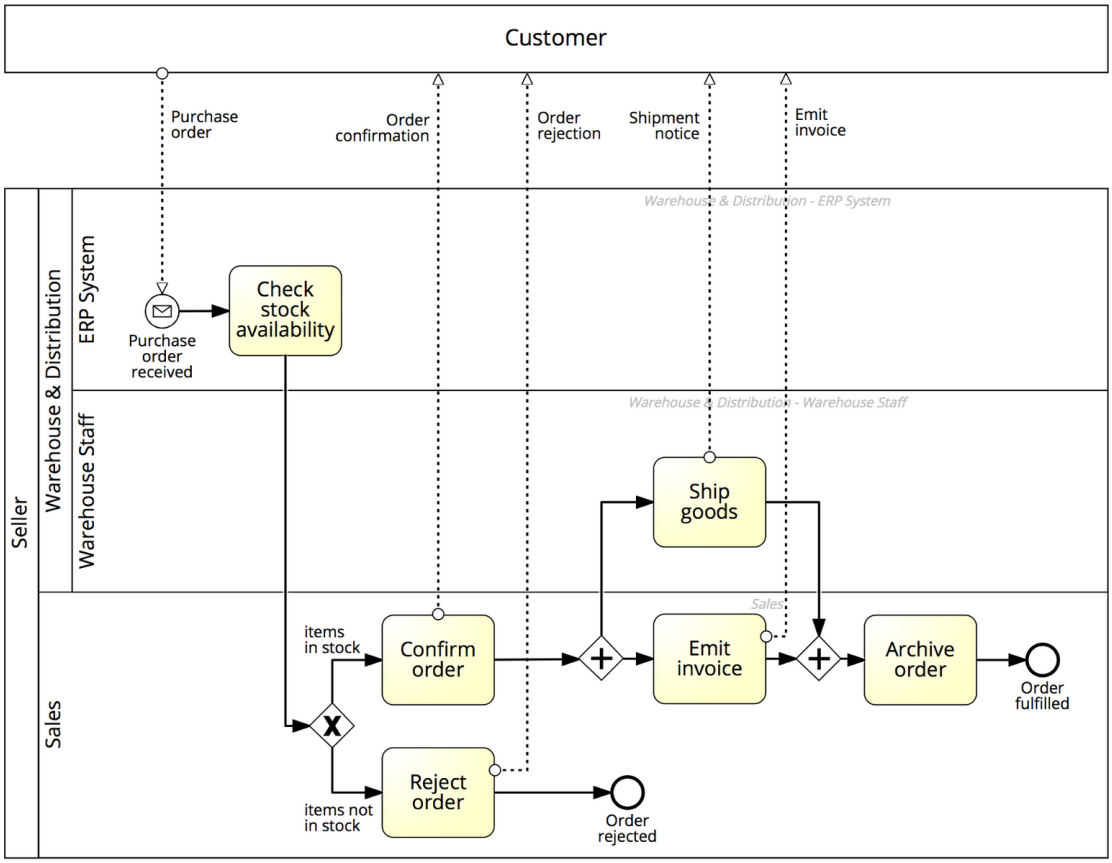






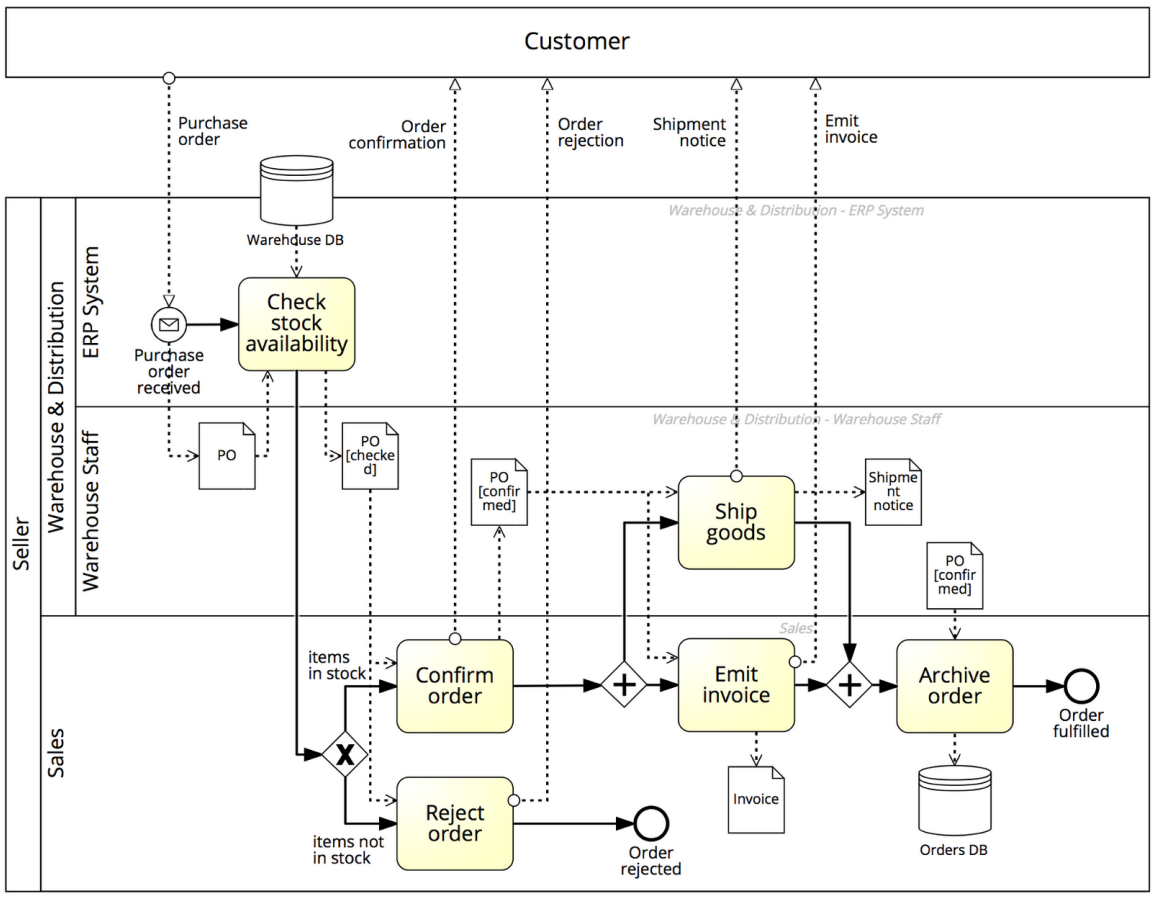
4. Identify the control flow

4. Identify the control flow



5. Identify additional elements (e.g. data objects, different types of events, exception handling...)

5. Identify additional elements



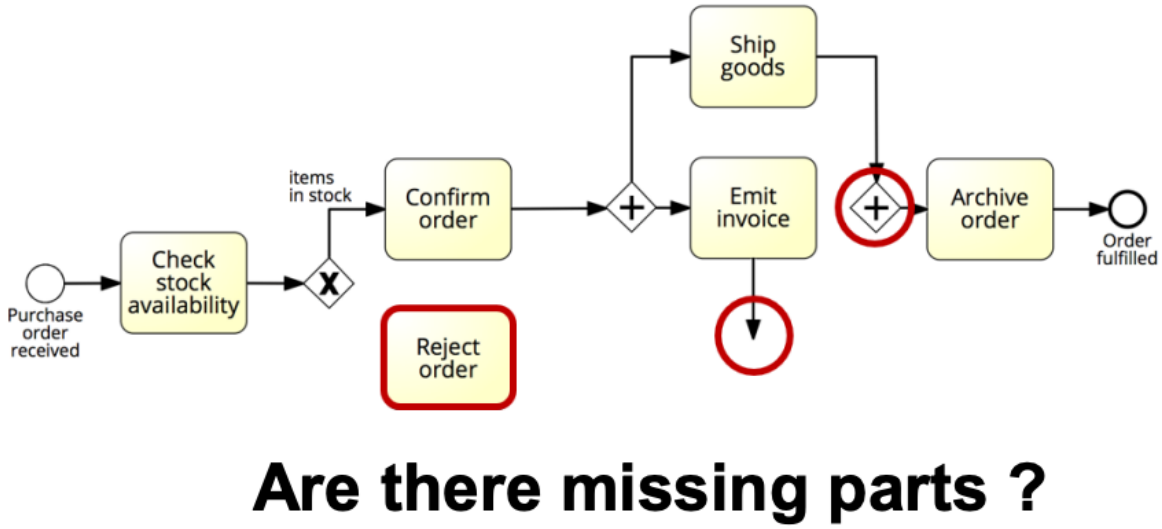
When should we stop modeling a process?

There is no need to go into a level of excruciating detail as it may harm the budget and even slow the improvement process.
The higher the level of detail of a process model, the likelier it is to be wrong.

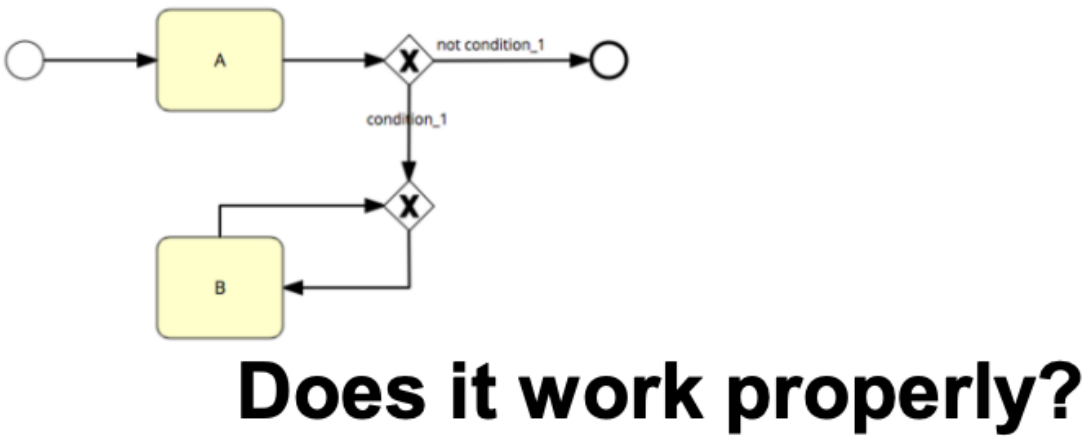
Process Model Quality Assurance

Process Verification

- **Syntactic quality**: refers to the conformance of a process model to syntactic rules of the modeling language used.
    - **Structural Correctness**: Are there missing parts?



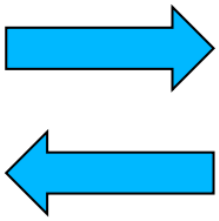
    - **Behavioural Correctness**: Does it work properly?



Process Validation

- **Validation**: is the activity of checking the semantic quality if a model by comparing it with its real-world business process.
- **Semantic Quality**: refers to the adherence of a process model to its real-world process.
- **A model is of high semantic quality if it is semantically correct.**

- Valid (all model instances are correct and relevant) + Complete (all possible process instances are covered)

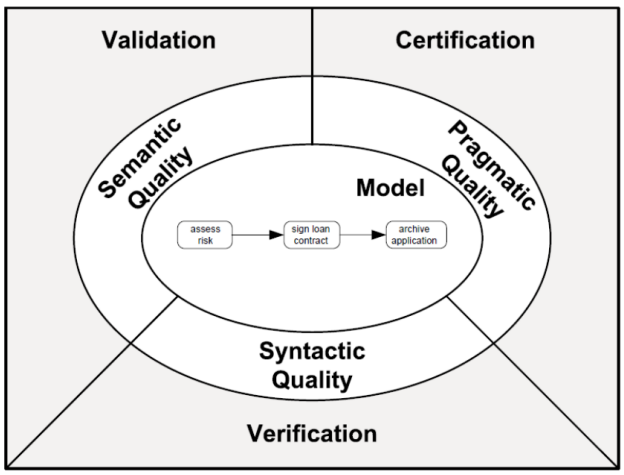


Domain Expert Process Analyst

Certification

- **Pragmatic Quality**: refers to the usability of a process model
- Challenge: anticipate the particular usage of the model.
- **Usability**:
  - **Understandability**: how easy it is to read and comprehend the model.
  - **Maintainability**: how easy is is to apply changes.
  - **Learning**: how good a model reveals how its corresponding process works in reality.
  - Model characteristics that influence usability include size, structural complexity and layout.

- **Validity**
- **Completeness**



- **Understandibility**
- **Mantainability**
- **Learning**

- **Structural correctness**
- **Behavioral correctness**