Introduction

Well defined, organized, implemented, and managed Business Processes are very critical to the success of any organization that wants to operate efficiently. Business Process Management these days is not merely a buzzword but has become a very important field on which companies will want to focus. Typically when the companies are small, they have ad-hoc, non standard, non documented processes.

However, as and when the companies start growing or have ambitious plans for growth, need for business process design and eventually BPM increases. Drivers for BPM implementation are as shown in the diagram below.

There are various other dimensions that companies need to restructure, like IT architecture, operational model, etc., and they also eventually relate to the process architecture of the company.

Companies may consider two options for revamping their processes.

1. **Green Field Processes** – Companies may start afresh to design the processes right from scratch. They can refer to standards like eTOM and ITIL, available in the market and design processes that fit their business objectives, goals, and strategies.

2. **Process Reengineering** – Companies may not want to throw away whatever they have, and can decide to modify the existing processes to build their new process regime. Existing processes in the company are studied, and bottlenecks or areas for improvement are identified. Future processes, designed on top of existing processes, provide solutions for the bottlenecks and problems identified.

   ![Figure 1. Drivers for Business Process Management](image-url)

   - In “Process Reengineering” projects, the following problems may be seen that make analysis of AS-IS process very difficult.
   - Process is not documented. It is known to the people involved in the process and knowledge of the process revolves around them.
All the details are not available in the process.

- Process abstraction is not uniform.
- Process implemented is different from the process designed and documented.
- Process users do not follow the process as demonstrated.

It is very critical, hence, to ensure that the AS-IS process is documented in every detail and as correctly as possible.

**Business Process Discovery (BPD)** – the discipline of BPM is a scientific way to extract the implemented process, validate it with the process owners, and create a baseline that gives a nearly exact AS-IS view of the business processes. It also helps to identify the process hierarchy, process owners, process entities, business rules, process operations, and statuses. Business Process Discovery is not very easy and quick. Due diligence and focus on the problem statement are key to the accuracy of the discovered process. This paper intends to assist in this tricky task of preparing for BPD and validating the output.

**PROCESS DISCOVERY TECHNIQUES**

Business Process Discovery can be done using two techniques.

- Manual – Process Analysts using interviewing techniques to derive the processes.
- Automated – Process Analysts use ABPD tools that can extract the process from databases or logs available with IT.

Process discovery is a complex task, and, hence, either one or both of the techniques should be used. The complexity of process discovery depends on various dimensions. Critical dimensions are

- Organization Structure Complexity – If Organization structure and, therefore, the operating models are complex, process discovery becomes increasingly difficult.
- Process Automation – With automation, organizations can build complex logic and, thus, can create complex processes. Hence, increased automation is directly proportional to increased difficulty.
- Process Maturity – If the organization has the following, then process discovery become easier.
  - Process Architecture with defined level of abstraction
  - Process Owners
  - Documented Processes

![Discovery Complexity Graph](image)
As the complexity of discovery increases, automated discovery plays a critical role. However, manual discovery technique is still required to extract the activities that are done manually as well as to validate the process extracted by automated discovery technique.

The following table provides a relatively simple matrix for selection of appropriate technique.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Process Type</th>
<th>Discovery Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Heavily or Fully Manual</td>
<td>Manual Discovery</td>
</tr>
<tr>
<td>2</td>
<td>Partially Automated Process Flow</td>
<td>Manual + Automated Discovery</td>
</tr>
<tr>
<td>3</td>
<td>Zero Touch, Fully Automated Process Flow</td>
<td>Automated Discovery</td>
</tr>
</tbody>
</table>

Wherever partial or full automation is available, automated discovery should always be done first. Manual discovery should be carried out to identify the manual process and to validate the process discovered automatically.

Scenario 3, mentioned above, is less likely as very few companies have fully automated their processes. Also scenario 1 is applicable to smaller companies where IT infrastructure is non-existent or is limited to business enabling functions. Scenario 2 is the most likely scenario and, therefore, most of the time both the discovery techniques are required to be applied.

**Manual Discovery**

Manual process discovery involves interviewing techniques. Information is gathered from the stakeholders (read sources of information) and normally is in scattered format. Process Analysts have to join the pieces of information together to form the process flow. The right set of questions should be asked in order to get the required details of the process. Knowledge of the domain helps analysts to identify such questions and get relevant answers from the stakeholders.

Manual Discovery typically goes through following phases.

- **Discovery Workshops**
  Process Analysts organize workshops or meetings with the required SMEs or Process Owners. Context is set, and objective is clearly communicated. SMEs provide a view of the process that they follow in their domain. Analysts take note of all the information from different SMEs so that it can be used to define the entire process. Analysts can ask leading questions to understand the process handovers between different units.

  The workshops can be conducted based on the following dimensions:
Approach

- Centralized
Process Analysts can bring all the stakeholders together in one workshop and get information from all the stakeholders at one go. The information will be like islands that the analyst may bridge, based on his understanding of the domain or by asking leading questions. If there are any conflicting views between the stakeholders, they can be cleared up immediately or noted down for tracking. The centralized approach should be considered when the process to be extracted is not very big and complex.

- Distributed
Different meetings can be scheduled to gather information from different stakeholders individually or relating to a particular domain. Once information is gathered from all the meetings, analyst can bridge the gaps. There may be a conflict of views, which will be clear only in later stages. Hence, analysts have to ensure that they contact the stakeholders immediately whenever they sight any such potential conflict.

Process Hierarchy

- Top Down
Analyst can begin at the highest level of abstraction of the processes and drill down to the detailed level. This means discussions with higher-level managers first, and then talking with the process owners or SMEs. Top Down discussions provide a very good business view of the problems. Analysts can relate the business view to the operational view very easily when they discuss with process owners or SMEs.

- Bottom Up
Analysts gather detailed information from SMEs. Based on the information gathered, related pieces are tied together to build a higher level view. This may be little difficult as the information gathered may not be complete.

Technique

- Free Form
Process Owners or SMEs provide the information, and analysts try to build process based on the information given by them. SMEs may miss out on some details and, hence, may leave a gap in the processes that isn’t immediately apparent. Free Form technique can be used
when the domain is very specific to the company or the analyst does not have enough knowledge of the domain.

- **Structured**

Analysts create a detailed questionnaire based on their understanding of the domain. SMEs provide answers to the questions and any additional information that they have. This scientific technique ensures consistency in the information provided by different SMEs. It is particularly helpful when the approach taken is of distributed meetings.

- **Process Modeling**

Once all the pieces of information are put together and process flow defined, it is very important to present in a structured format and as agreed with the customer. If any process modeling tools are used, different process models should be created and linked to provide a comprehensive AS-IS process view. If tools are not used, the processes should be documented as per customer specifications. Generally, Microsoft Visio can be used in the absence of full-fledged process modeling tools.

Different types of process diagrams can be created, based on the requirement. The different process models are

- Hierarchical Process Flow Models
- End to End Process Flow Models
- Process Swim Lanes
- Organization Chart
- Resource Diagrams
- Entity Flow Diagrams

Ensure that correct details are put in the various models or in the process documentation. It is very important not to put too many details in the process, or the process will look too complicated or tedious.

- **Review Workshops**

The AS-IS process discovered from the discussions needs to be signed off by the stakeholders. Process analysts have to organize a review (walkthrough) of the process defined. Any review comments are worked upon and the process modified in response. Review ensures that the process documented or modeled is as per the information given. While conducting process walkthrough, the analyst should ensure that all the conflicting points during the process discovery discussion sessions are reviewed and then agreed upon by the related stakeholders.

**Benefits**

Manual Process Discovery provides the following benefits:

- The process that is practiced is discovered. It may be different from the process that has been documented. People who practice the process may find convenient workarounds for the standard processes, and this is captured during manual process discovery.
- Many practical problems are also discovered that may provide a better view of the problem statement to the analysts. This helps in defining better TO-BE Process.
- Since stakeholders are involved in the discussions and for signoff, they are more than willing to contribute, become more accessible, and share a lot more insights.
• Where automation is very limited or not implemented correctly, manual process discovery is the only way to gather information.

Challenges

Though manual process discovery has certain benefits, it may be time consuming and difficult to gather the required information. It poses the following challenges:

• It may be difficult to identify the right mix of stakeholders / SMEs for discovery.
• The availability of stakeholders for discussion may break the sequence of gathering information.
• Information on the following may not become available in manual discovery.
  o Process Performance
  o Process Scalability
  o Business Exception / Failure statistics
• Resolution of conflicts between stakeholders
• Bridging the gap in the information received and defining the process

Dos and Don’ts

• Understand the problem statement carefully and confirm the scope at the beginning.
• Decide on the approach, hierarchy, and technique to follow. It’s always good to prepare before the workshops, and structured questionnaires should be the preferred way to extract information.
• During discussions, if stakeholders are stuck at any issues, ensure that you note the issue and request resolution later. It helps to save time to discuss all the areas within the given timeframe.
• Understand the handover points and ask questions to ensure that there are no issues in process handovers from one process owner to the other.

Automated Discovery

In order to become more competitive, responsive, and efficient, companies try to maximize automation of their business / processes. IT systems implement complex business logic and complicated business processes. They can store each and every transaction made with all the relevant details. Hence, if the data that is operated and stored by different systems are analyzed, greater details of the process, business rules, volumetric will be available.

However, doing this manually will be an almost impossible task, as the smallest of the systems will have hundreds or thousands of rows of data. There are applications available that do this task for you. These tools can be set up to read the databases, and they can create state models and eventually draw process models.
Figure 3. ABPD Tools

The tools read the key data attributes, such as unique identifiers of the entities, their relationships, various timestamps, and status. Timestamp fields provide information on when the entities were created or modified. Hence, it is critical to know the timestamp. Discovery, with the help of such applications, is called Automated Business Process Discovery (ABPD). However, it may not be entirely automated for following reasons.

- Understanding of the database structure, its entities, and their relationship, and operations performed on the entities is required to be done manually.
- After this analysis of data, the relationship has to be set in the tool manually.
- Some of the tools only generate state models; process models may have to be drawn manually.

**Benefits**

- Provides exact AS-IS visualization of the process that has been implemented / automated and not the one that has been defined and documented.
- Volumetric analysis of all the process steps is readily available. This is critical for quicker and accurate problem analysis.
- Provides view of all the exception paths along with volumes. It helps analysts to identify the bottlenecks in the processes very easily.
- Process view can be changed very easily to exclude exception paths or to reduce complexity, and vice versa.
- Process discovery can be done repeatedly with ease, to reflect any changes done in the processes.
- Process discovery can be done on archived data; therefore, with no impact on any day-to-day business.

**Challenges**

- Security requirements of companies may not allow access to some business critical data of the company.
- Understanding of the entities, relationships, and operations poses a real challenge, as it may be technology specific.
- Database may not be completely relational, and constraints may not be available. It makes identification of process flow for a single transaction very difficult.
- Where data may not be stored in separate attributes, but in some other format like CLOB or XML, establishing relationship between entities may not be possible, making discovery very difficult.
- Connecting to database may not be easy because of different connection technologies.
- Understanding database and initial set up of tools to read the data is time consuming.

**Dos and Don'ts**

- Identify all the systems that are involved in the process or problem under scope. Identify the databases that they update. If you don't have knowledge of the database, it's always good to seek the help of database administrators and system designers to understand the database.
- Identify the entities that are created, modified by the systems in scope. Shortlist the tables that relate to these entities and identify the data fields that contain timestamp. Timestamp is very important as tools use timestamp to identify the process flow.
- If data is stored in complex data type or in XML format, ensure that the tool can read such a format. If not supported, you probably need to find some other tool that supports reading such data types.
- Tools will have a limitation on the number of records that it can read from the database. Hence, ensure that you pick up minimum sufficient data from the database to read.
- If the outcome of the discovery doesn't give you an entire process map or a major functionality is not present in the discovered process, increase the data set. This can be an iterative trial and error process.
- Note down all the exception paths and their volume. This data is critical input for bottleneck analysis.
- Set up review calls with stakeholders and validate the process that you have captured.

**Discovery Tools**

There are quite a few tools available in the market, and some BPM tools are adding capability to discover business process or integrate with the discovery tools to analyze the discovered business process.

Most of the early stage tools have been built on the mathematical / analytical model called “Petri Net.” However, recently developed tools have preferred a modeling language “EPC” to develop discovery algorithm.

Some of the tools have been listed below.

- PRoM
- Fujitsu Interstage
- BMC Discovery
- IDS-SCHEER ARIS PPM5
- Pallas Athena BPM one Process Mining
- Metastorm ProVision
Emerging Trends

Though Business Process Discovery is not a new concept, its importance and use has increased recently. It has seen lot of changes from the 90s when the process discovery was in its initial phases.

Application of Business Process Discovery

ABPD has seen increasing acceptance in different industries. For complex organizations, ABPD is a very big relief. More and more companies will use ABPD for any process re-engineering and process management projects. ABPD, combined with process simulation, will not only provide the AS-IS view but the performance capability of the companies as well, providing much better insight into the problems than what current ABPD implementation can give.

Better Analysis

The analytical capability of the tools is on the increase. Currently, analysis is largely manual; however, future generations of the tools will see more automated analysis. Various analysis reports will be one click away. Tools will have the capability to compare AS-IS process with the standard library processes to identify gaps and will automatically suggest TO-BE processes as well.

Closer Integration with end-to-end BPM

The discovered AS-IS process view provides mainly activity flow or state model. The AS-IS process has to be manually drawn, based on the extracted process, to create a standardized process model. However, efforts are being made to integrate the ABPD into end-to-end BPM process so that the discovered process is automatically drawn as per standard modeling language. Since discovery will be coupled with end-to-end BPM, the discovered process can be simulated, compared against library, and a TO-BE process can be defined automatically.

References

2. www.bmc.com/products/offering/ITBM.html
3. www.bpm-research.com
4. www.prom.win.tue.nl
5. www.computers.us.fujitsu.com/interstage/process_discovery

------

Author

Sandeep Jadhav is a lead consultant with Infosys Technologies Limited with 8 years of experience in business process management, optimization, and transformation programs. He has lead and successfully delivered process transformation programs across various enterprises across the world. His current focus area is end-to-end process management automation to minimize the cycle time right from process discovery to process monitoring and control stages. Sandeep can be contacted at sandeep_jadhav@infosys.com
BPTrends Linkedin Discussion Group
We recently created a BPTrends Discussion Group on Linkedin to allow our members, readers and friends to freely exchange ideas on a wide variety of BPM related topics. We encourage you to initiate a new discussion on this publication or on other BPM related topics of interest to you, or to contribute to existing discussions. Go to Linkedin and join the BPTrends Discussion Group.