

Team Process Data Warehouse – Roadmap

Other documents in this collection have elaborated the high level requirements and designs for a TPI* Data Warehouse. These documents described the future “end-state” for the warehouse. Of course, in keeping with best practices for project planning and management, this functionality will be developed using an iterative development approach.

This immediately implies that all of the described functionality will not be present in the first iteration. Instead, dependencies and priorities must be weighted to allocate deliverable requirements to various iterative releases.

The following roadmap describes one proposed development strategy. This strategy is open to continual refinement based on stakeholder priorities, actual progress, and evolving requirements.

Iteration 0: High-Level Requirements and High-Level Design *(Completed Dec 2012)*

The first iteration will focus on the elaboration of requirements. A thorough understanding of current and future requirements will provide the best foundation for a design that avoids backtracking and rework. Requirements should be gathered from as many stakeholders as possible, to include the larger TSP community.

A high-level design will be produced and mapped to these requirements. In keeping with goals for interoperability and extensibility, this design will be evaluated for compatibility with multiple currently existing TSP/TPI tools.

Iteration 1: Core Database Objects *(Completed Feb 2013)*

A database schema will be designed to hold the primary sources of TSP/TPI data: time, size, and defects. This detailed design will be validated against existing TSP/TPI tools to ensure the viability of the ETL process.

SQL DDL scripts will be developed to create objects in various database platforms based on the validated schema.

Iteration 2: Core ETL for Process Dashboard *(Completed Apr 2013)*

Using the open-source Process Dashboard application as a starting point, an ETL process will be developed to extract data from the Process Dashboard and load data into the tables created during the previous iteration.

* TSPSM is used by teams working in a wide variety of problem domains (e.g. software, hardware, services). Since these activities are not limited to software, the name “Team Process Integrated” and the acronym “TPI” are used in this document to describe the full range of TSP-inspired high-maturity processes, and to avoid improper use of Carnegie Mellon service marks. TSP is a service mark of Carnegie Mellon University. Carnegie Mellon University has neither contributed to nor evaluated the contents of this document.

This first ETL iteration will include a significant percentage of effort devoted to the construction of ETL scaffolding, such as:

- The user interface for configuring Process Dashboard data sources
- Logic to extract Process Dashboard data and prepare it for loading into the warehouse
- Data abstraction layers to connect to the warehouse database and load metrics in a platform-independent way

Care will be taken to make certain this this ETL logic is as modular, reusable, and lightweight as possible.

At the completion of this effort, a usable RDBMS will exist that others can begin using for analysis.

Iteration 3: Process Dashboard Core Reporting *(Completed July 2013)*

Once the warehouse contains a core set of data (time / size / defects), it will be possible to generate many of the standard TSP reports based on queries from the data warehouse. As a proof of concept for this possibility, the reporting logic in the Process Dashboard will be retrofitted to work against the data warehouse.

Iteration 4: COTS Reporting *(Enabled July 2014)*

Reports are an extremely important part of the value chain for the data warehouse. At any point in time following the completion of Iteration 2, third-party business intelligence packages (open source and/or commercial) could be incorporated as a means of slicing and dicing the data in the warehouse.

Iteration 5+: Broader Data Scope

Iterations 1 and 2 above can be repeated any number of times to add new types of data to the warehouse – for example, earned value data, workflow process definitions, project metadata, etc. At the completion of each iteration, the RDBMS will contain more and more data that can be used by independent third party analyses.

Iteration 6: Extended Reporting

With a warehouse-enabled calculation engine in place, it would become possible to begin generating very sophisticated reports – for example, reports that:

- Analyze data by workflow steps instead of metrics collection framework phases
- Show the changes that have occurred to a team plan over time
- Quickly analyze data from a large number of historical components, spanning many project iterations
- Slice and filter data in new ways

Iteration 7: Launch Planning Support

The extended reporting capabilities described above could be folded back into tools such as the Team Dashboard and the WBS Editor to provide PROBE-style planning support during a TSP launch.