PANTOPIX

Unlocking the Potential of Legacy Data with AI and Knowledge Graphs

Information Energy 2025

Sofia Darie and Nikhil Acharya

3rd of April 2025

Agenda

PANTOPIX

- Introduction
- The Challenges of Legacy Data
- Where the Value Lie
- How do we create value
- Key Learnings

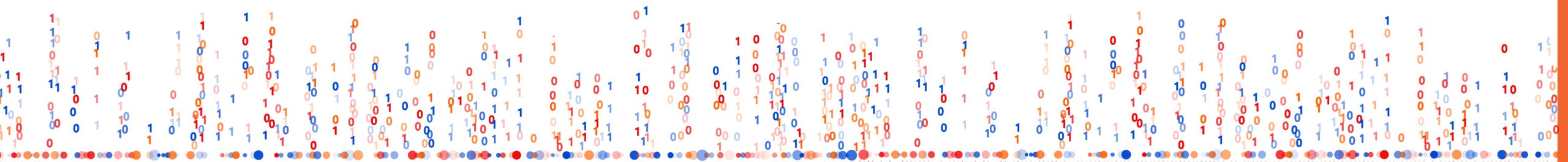
PANTOPIX

Intelligent Solutions for Technical Communication

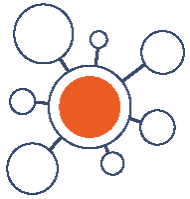
We empower our customers to step confidently into their digital future.

Together, we're crafting the intelligent information landscape of tomorrow.

With exceptional consulting expertise and cutting-edge technological innovation.

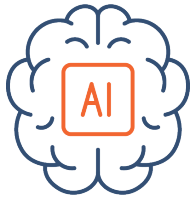


Turning information into value



Knowledge Graphs

Connecting information silos with the help of semantic technologies.

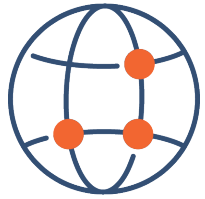


AI-powered Solutions

AI as a driver for intelligent provision of information.

Content Management & Delivery

Conception, configuration, migration and integration.

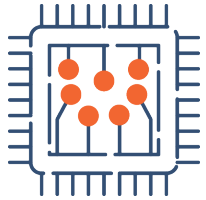


iiRDS

The technical standard enables a content-neutral transfer of information.

Data Pipelines

Transformation, publication and migration of data.
Automated and intelligent.



DITA

Topic-based creation, distribution and use of technical information.

About us

Sofia Darie

Consultant

Nikhil Acharya

Knowledge Engineer

PANTOPIX

The Challenges Legacy Data

A Short Introduction

What is Legacy Data?

- Technical documentation for in and out of support products.
- Technical documentation difficult to operate with in modern digital ecosystems
  - Manuals, PDF´s, Scanned Documents, old XML Formats
  - Siloed information across different departments and systems
  - Structured and unstructured data
  - Disconnected data sources (An old CCMS System)
  - Outdated formats

The Challenges of Legacy Data

Lack of standardization

- Standards for information creation
- Inconsistent metadata concept and assignment
- Inconsistent naming conventions and classifications

Difficult to identify and retrieve

- Outdated indexing methods
- Information stored across multiple systems

Difficult to contextualize the information

- Loss of context through the format
- Generic description of data
- Lack of detail

Resources and time consuming

- Human and financial resources
- Cross department alignment
- Complex migration processes

PANTOPIX

Where the Value Lie

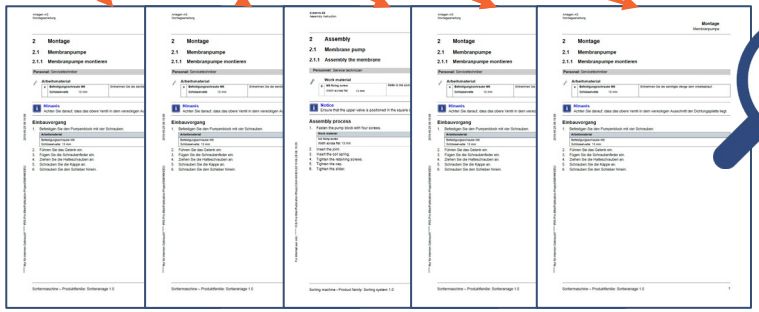
A Closer View

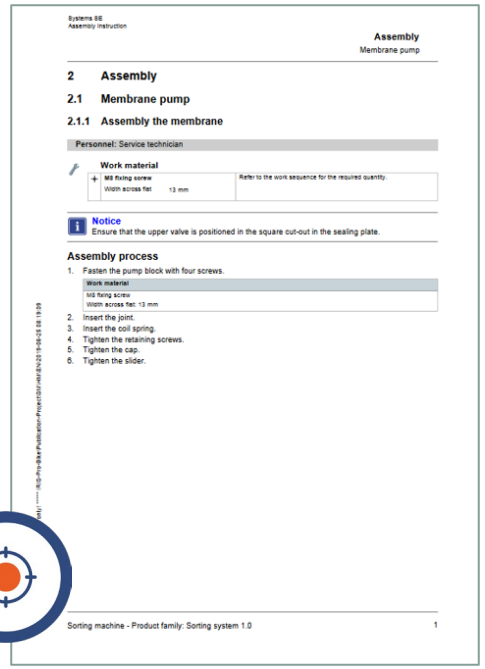
Where does the Value Lie?

- The foundation of modern digital ecosystems
- An integral component of a product or software
- Essential for both pre-sales and after-sales product support
- Equips field and service teams with comprehensive maintenance knowledge
- Holds valuable data that can be interconnected for greater insights

The Value Lies Within

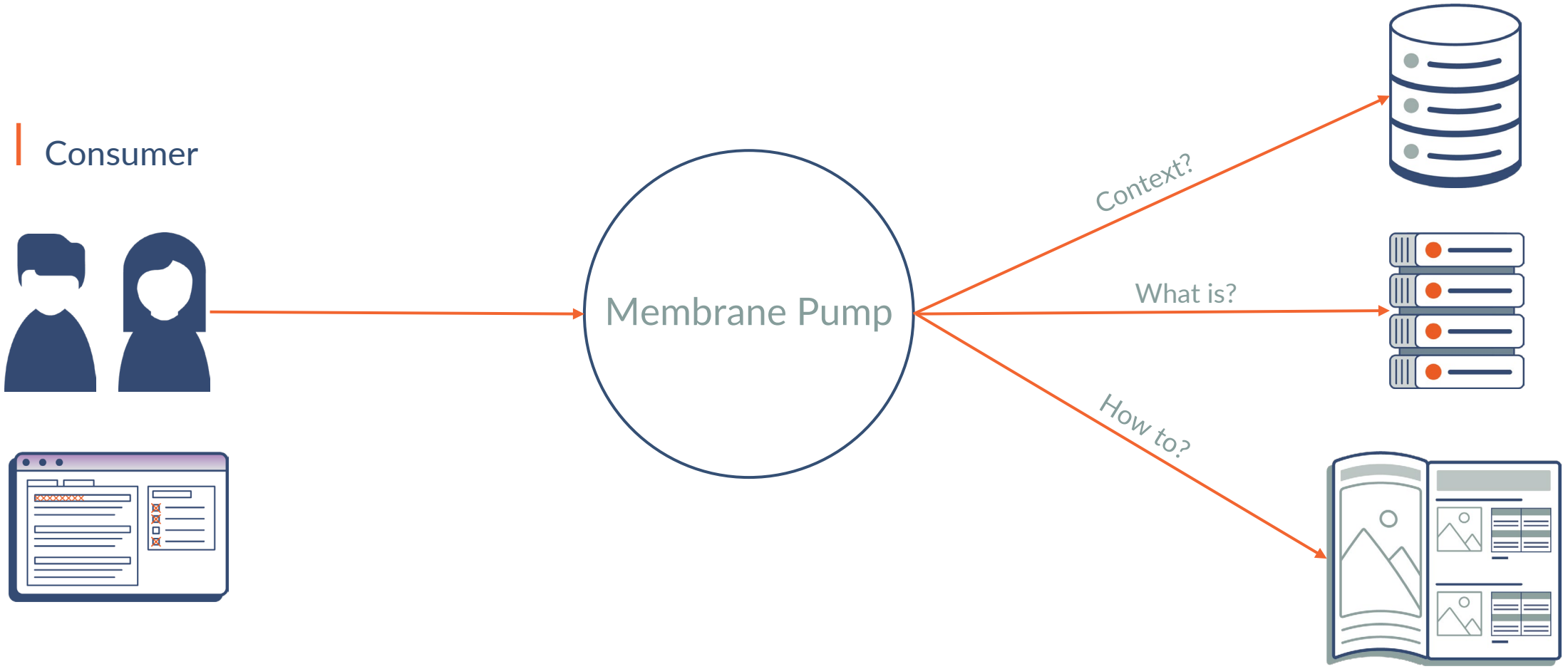
Metadata:
Document type: Assembly Instruction
Product: Sorting System 1.0
Variant: 1.0





Metadata:
Topic type: Task
Product: Sorting System 1.0
Activity: Assembly
PLC Phase: Maintenance
Component: Membrane Pump
Spare Part(s): Membrane
Tools: Fixing screw
Consumable: Kit

A Bird´s Eye View of the Why



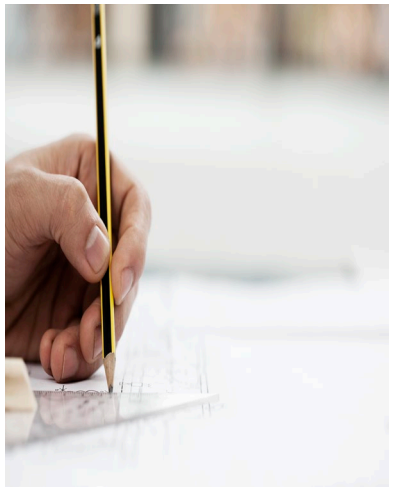
PANTOPIX

How do we Create Value?

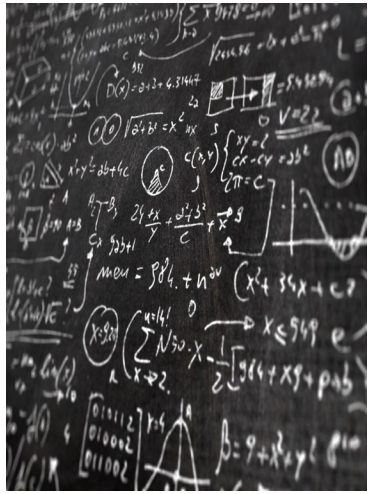
How to Manually Handle Old Legacy Documents

Fetch

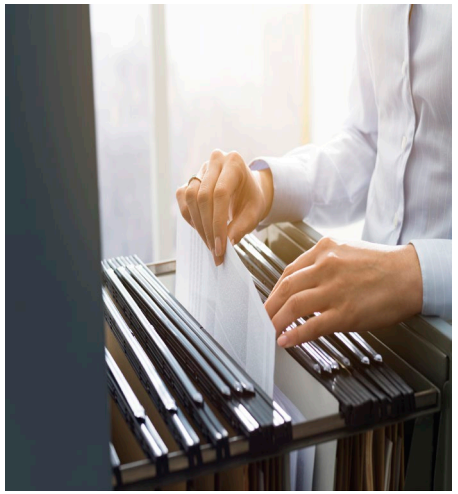
Classify



Extract

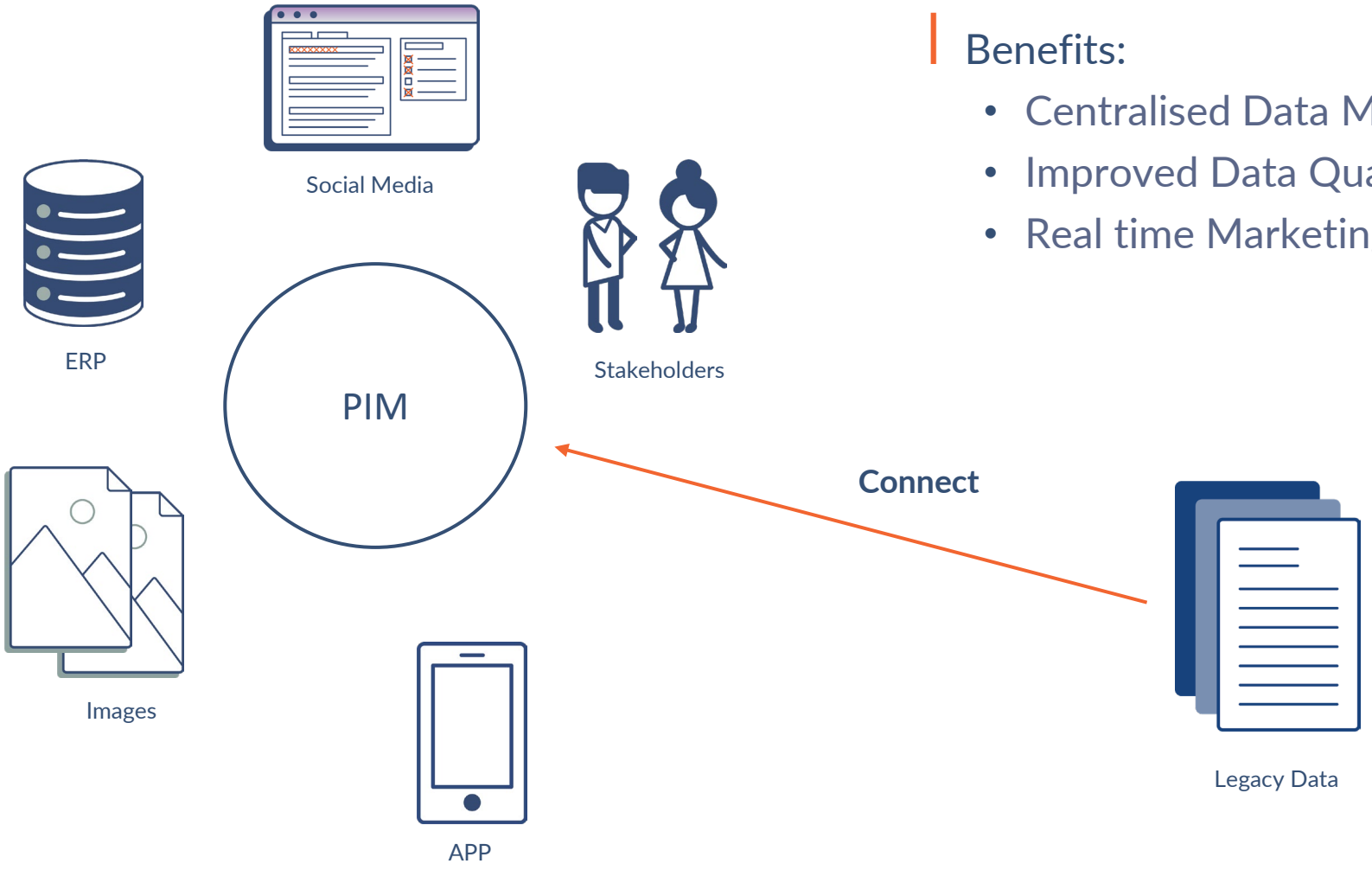


Analyse



Update

PIM Data Ecosystem



Benefits:

- Centralised Data Management
- Improved Data Quality
- Real time Marketing

Adding Legacy Data to PIM

Challenges

- Complexity and diversity
- Lack of Documentation
- Complex integration process
- Maintaining data security and compliance
- Shortage of skilled personnel who understand legacy data

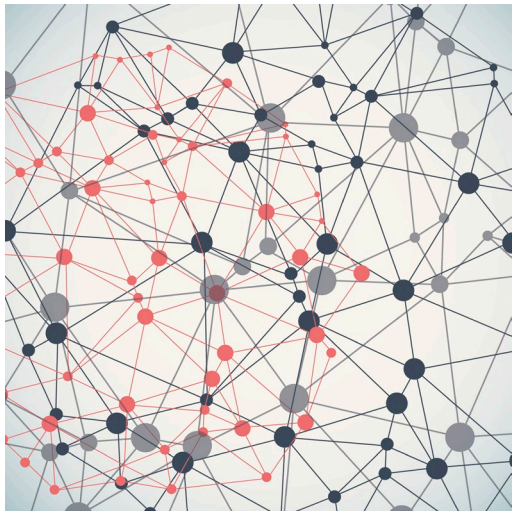
It‘s like trying to connect a gramaphone to bluetooth!

Potential Solutions

- Centralised data store and data model
- Identify legacy information which can be connected
- Translate legacy information to the data model
- Now Update the ecosystem

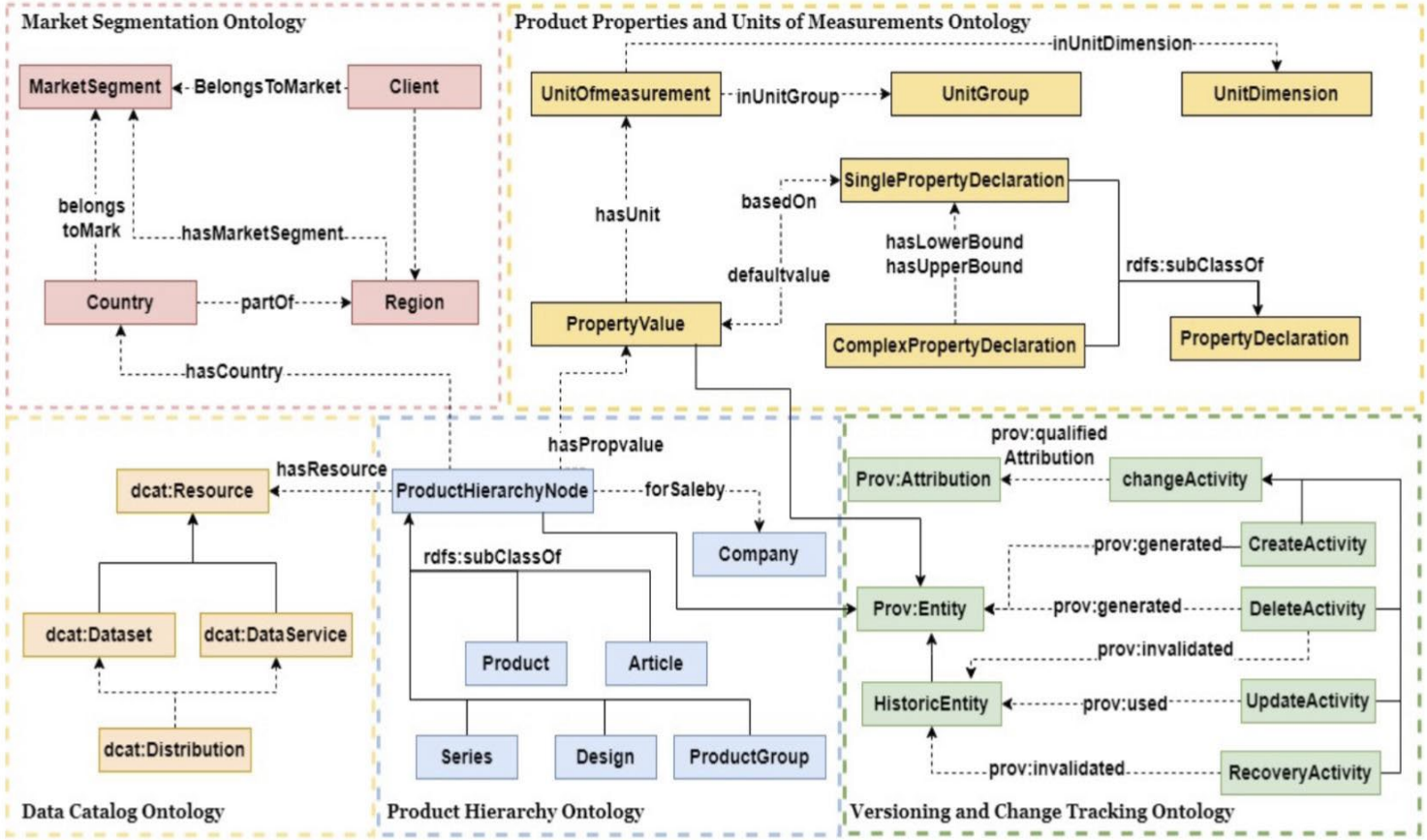
Legacy Data

Data and Knowledge Engineers



Centralised Knowledge Base

PIM Data Model



Expectations

- Representing the entire ecosystem with a standard data model
- Harmonising legacy information with the PIM data model
- Five set of ontologies to improve maintainability and keep heterogeneity

Translation to PIM Data Model

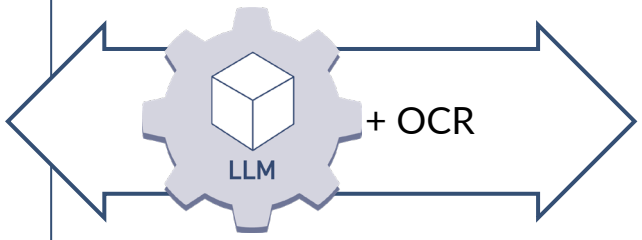
Operating & Installation Manual

for Grey Water Treatment Systems

Series: PowerClear

Firmenlogo

Type: PowerClear 1500
PowerClear 3000
PowerClear 4500
PowerClear 6000
PowerClear 10000

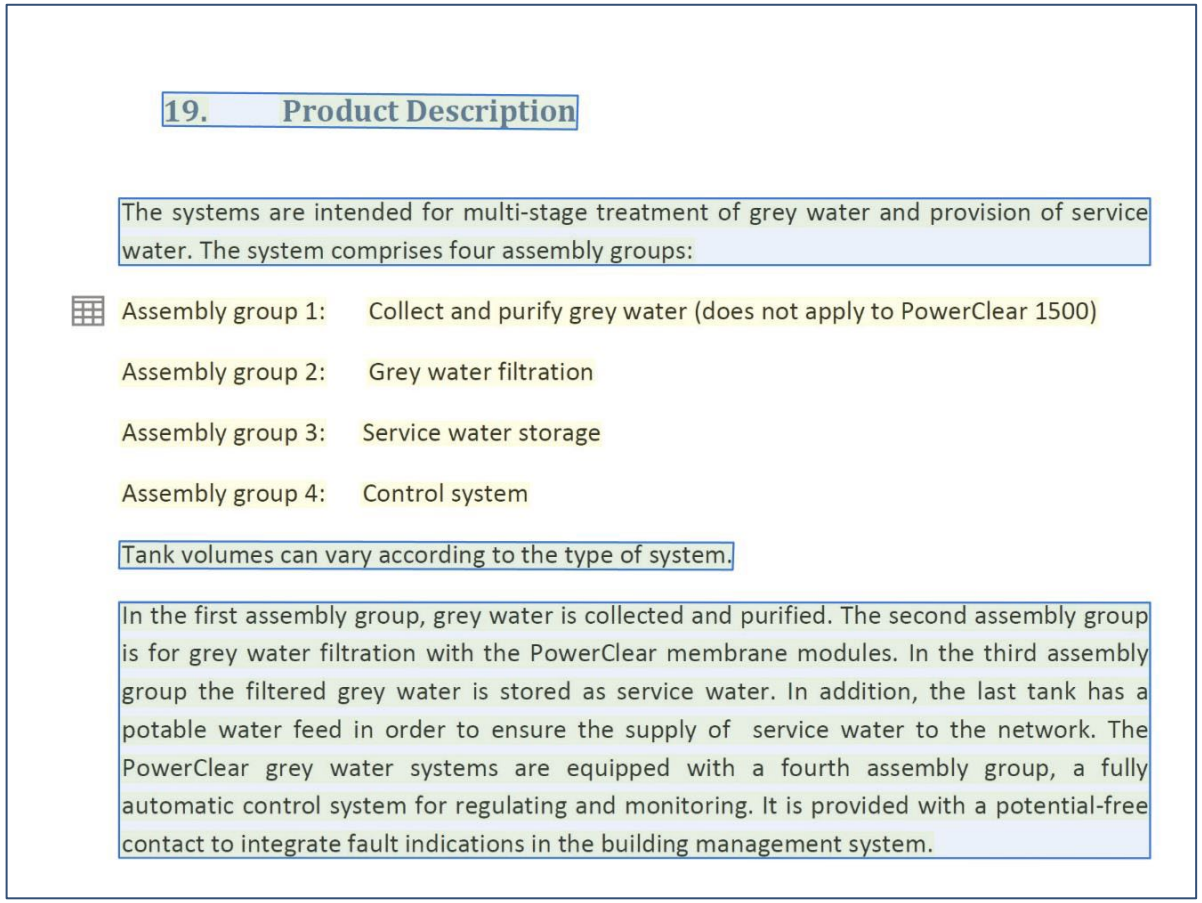


Title: Operation and Installation Manual for Grey Water Treatment Systems

Series: PowerClear

Type: PowerClear 1500
PowerClear 3000
PowerClear 4500
PowerClear 6000
PowerClear 10000

Translation to PIM Data Model



19. Product Description

The systems are intended for multi-stage treatment of grey water and provision of service water. The system comprises four assembly groups:

Assembly group 1:	Collect and purify grey water (does not apply to PowerClear 1500)
Assembly group 2:	Grey water filtration
Assembly group 3:	Service water storage
Assembly group 4:	Control system

Tank volumes can vary according to the type of system.

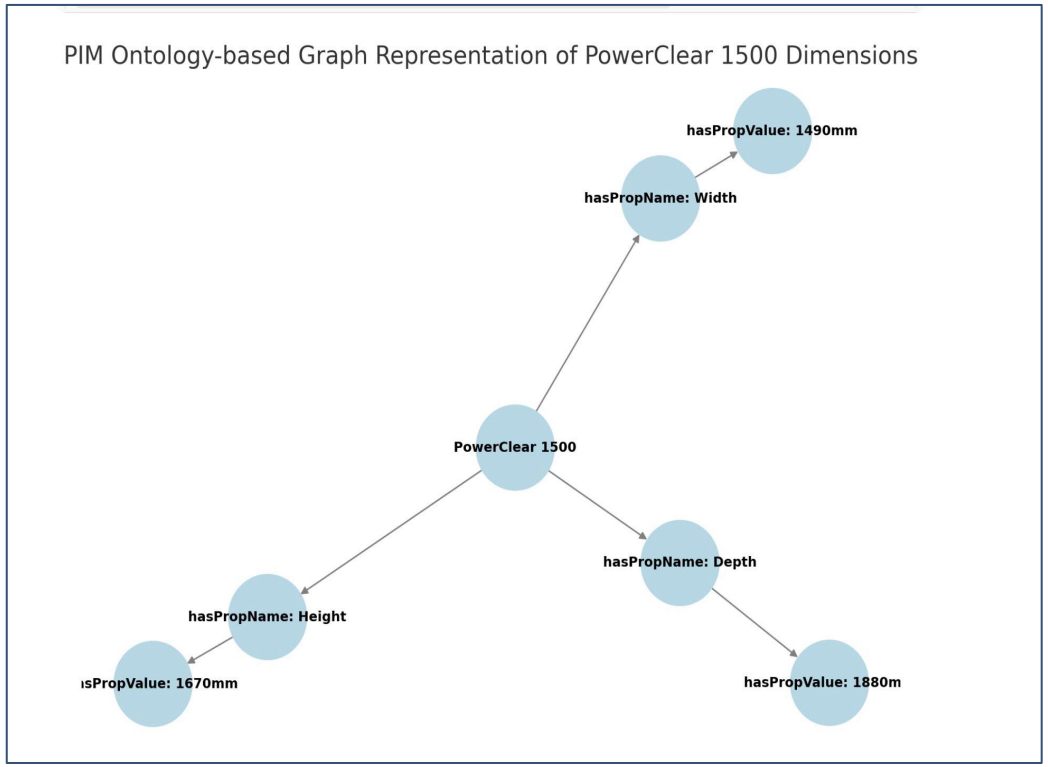
In the first assembly group, grey water is collected and purified. The second assembly group is for grey water filtration with the PowerClear membrane modules. In the third assembly group the filtered grey water is stored as service water. In addition, the last tank has a potable water feed in order to ensure the supply of service water to the network. The PowerClear grey water systems are equipped with a fourth assembly group, a fully automatic control system for regulating and monitoring. It is provided with a potential-free contact to integrate fault indications in the building management system.

- Extract Product Description
- Extract functioning of each assembly group
- Extract semantics of the description

Mapping to PIM Ontology

	PowerClear 1500	PowerClear 3000
Height	approx. 1670mm	approx. 1770mm
Width	approx. 1490mm	approx. 2740mm
Depth	approx. 1880mm	approx. 2020mm
Weight filled max.	approx. 2850kg	approx. 5750kg
Connection grey water inflow	DN110	DN160
Connection grey water overflow	DN110	DN160
Connection service water overflow	DN50	DN75
Connection potable water backfeed	1“ DN 20mm	1“ DN 20mm
Connection service water transfer	2“	2“
Membrane surface	3.2m²	6.4m²
Treatment capacity	approx. 1500l/day	approx. 3000l/day
Tank volume assembly group 1	approx. 1350l	approx. 1800l
Tank volume assembly group 2	approx. 1400l	approx. 1900l
Tank volume assembly group 3		approx. 1900l
Supply voltage	230V/50Hz	230V/50Hz
Back-up fuse	16A	16A
Max. power consumption	approx. 300W	650W

mapping



Building a Question Answering System

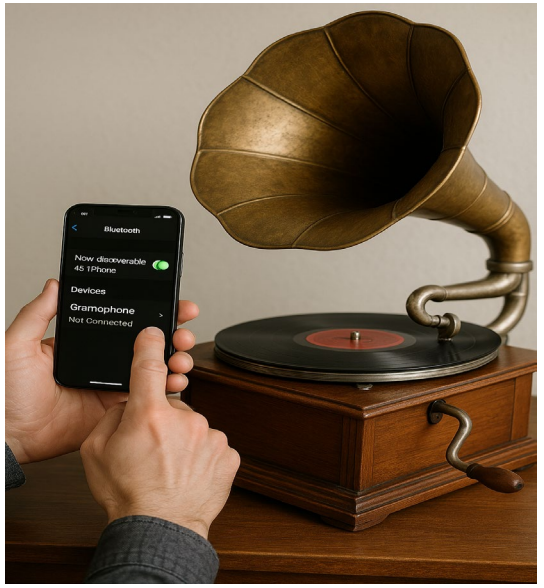
Question: What are the height, width, and depth of PowerClear 1500?

Generated Query

```
PREFIX pim: <http://example.com/pim#>

  SELECT ?propName ?propValue
  WHERE {
         ?product a pim:Product ;
          pim:hasPropName ?propName ;
          pim:hasPropValue ?propValue .
          FILTER (?product = "PowerClear 1500" && (?propName = "Height" || ?propName = "Width" || ?propName = "Depth"))
     }
```

Benefits



Imagine successfully integrating a gramaphone to all modern tech!
Let‘s discuss benefits?

Key Learnings

- Everything is evolving, but one constant remains – the quality of information
- Less is more – Mindset shift to smaller manageable entities
- Legacy information

… can be of real high value to improve processes
… can broaden scope of comparison
… can enhance reusability

PANTOPIX

Thank you!

www.pantopix.com

Feedback

PANTOPIX

Location Lindau

Josephine-Hirner-Straße 2
88131 Lindau

Location Munich

Klugstraße 47A
80638 Munich

Location Romania

30 Infrăţirii Street
400393 Cluj-Napoca

www.pantopix.com
info@pantopix.com