

# 5 rules to ensure sustainable data management platform deployment

---

Building a Robust Data Foundation



## What we do:

IOSight – an ‘All in one’ software Solution Provider for Water / Energy facilities of all types. We collect, store, model, present and analyze data from all available sources within the facility.

The iGreen system embeds smart analysis using specific domains algorithms.

Employs a quality group of engineers, software & algorithm experts with vast experience, based on more than 100 installations, world-wide.

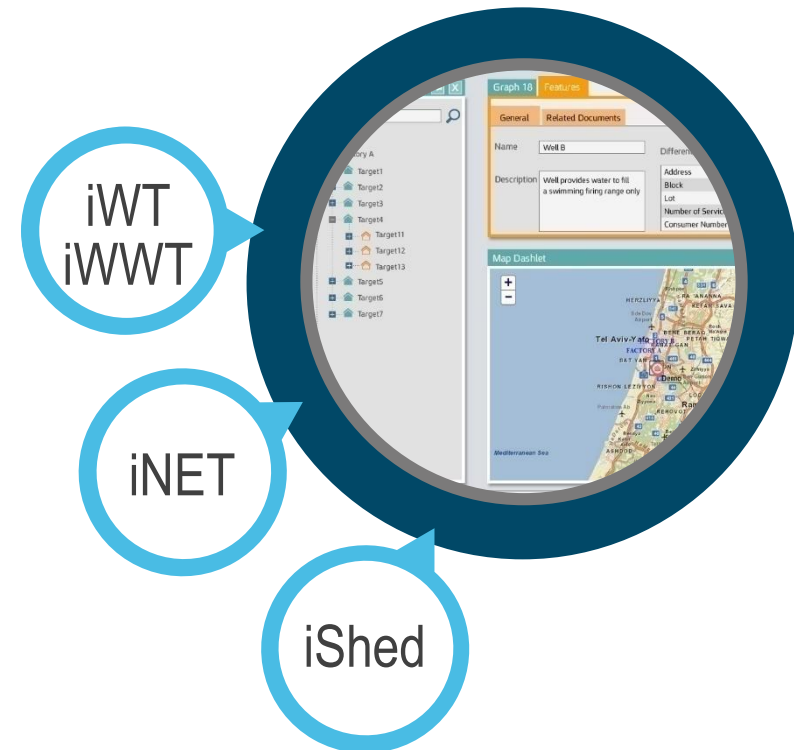
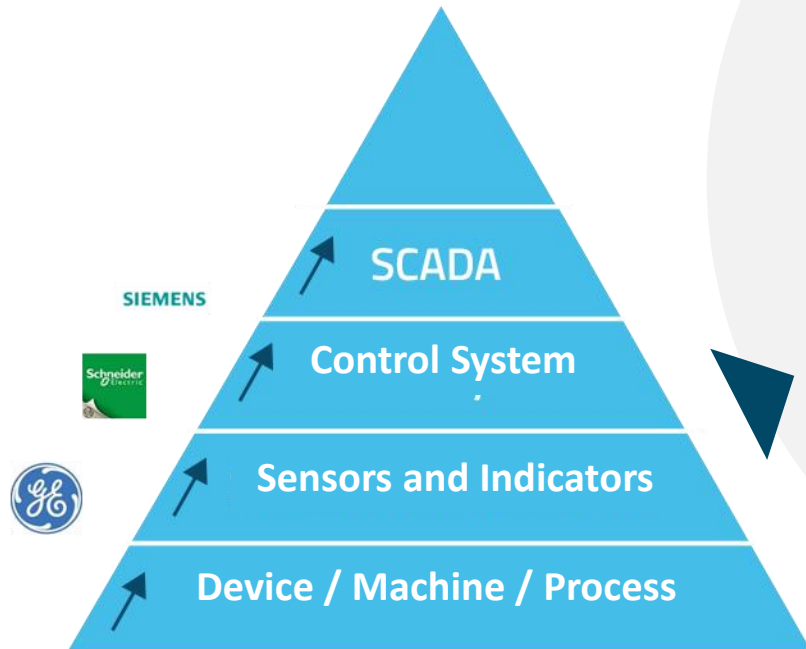


Delivered successfully to leading player in the water market in the US, China, EU and Israel

# Solution Overview

- Historian
- Sensors Devices
- Control System
- SCADA
- AMR
- GIS
- Lab
- Manual
- ERP
- CRM

Data collected from all available sources in the organization



Algorithm-Based Expert Systems


Cleansing | Visualization  
Analysis | Decision Support

## Why do projects fail?

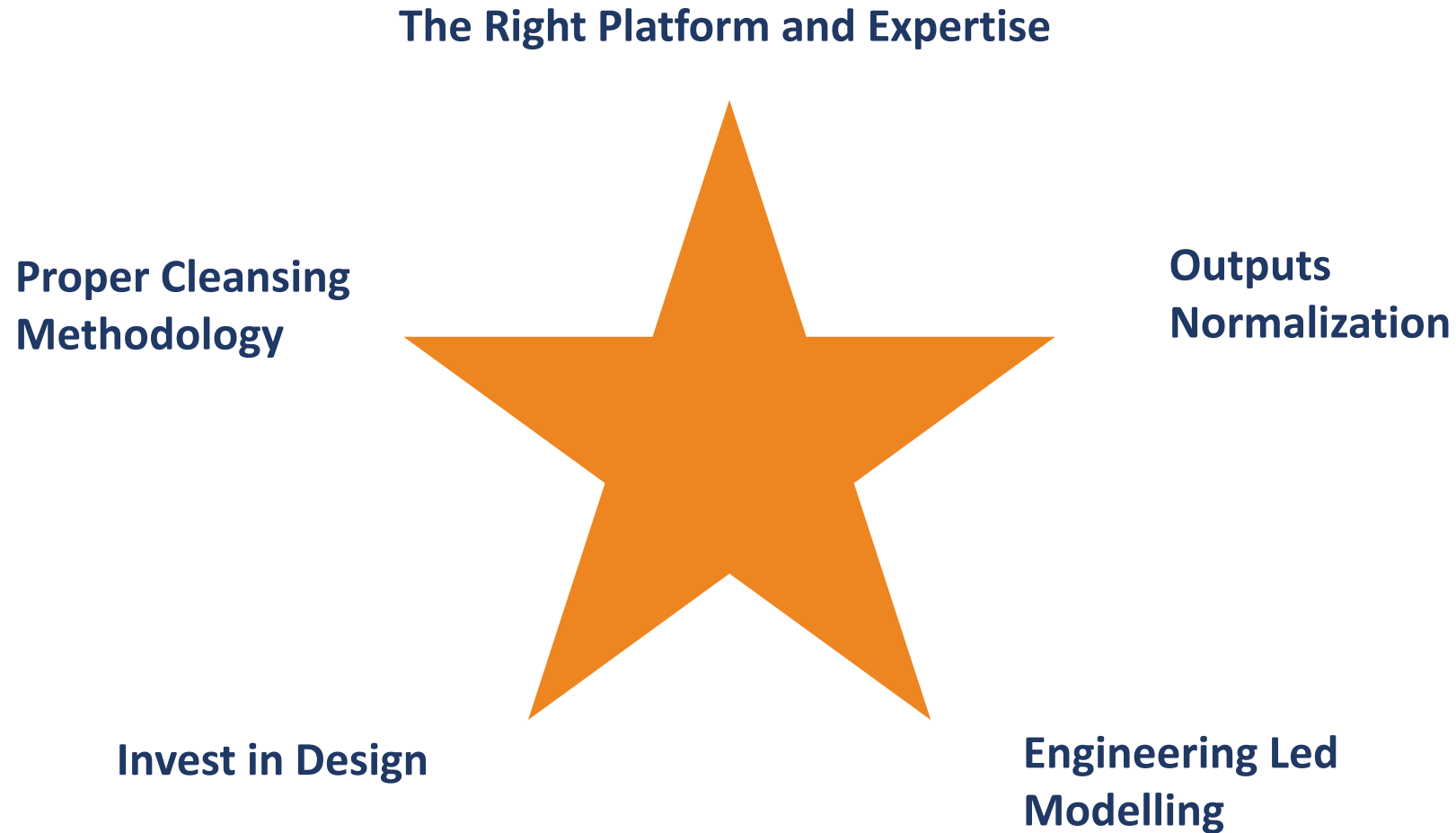
- An **incidental approach** to data science
- A **bottom-up** attitude



## Doing it right

- Data should be **reliable**
  - The system should be **accessible and easy to use**
  - **Long-term performance** should be taken under consideration from the inception of the project
  - **Team and management** - An ongoing maintenance of the system requires a trained team, data integrity 'watch dogs' and mainly – continuous involvement from the user
- 
- **Coherency** - The BI system (e.g. reporting/ dashboarding, analytics) should be integrated in the DB and must reflect the data model
  - **Future analytics** must go through the **validated data base**

# 5 Rules for successful data projects



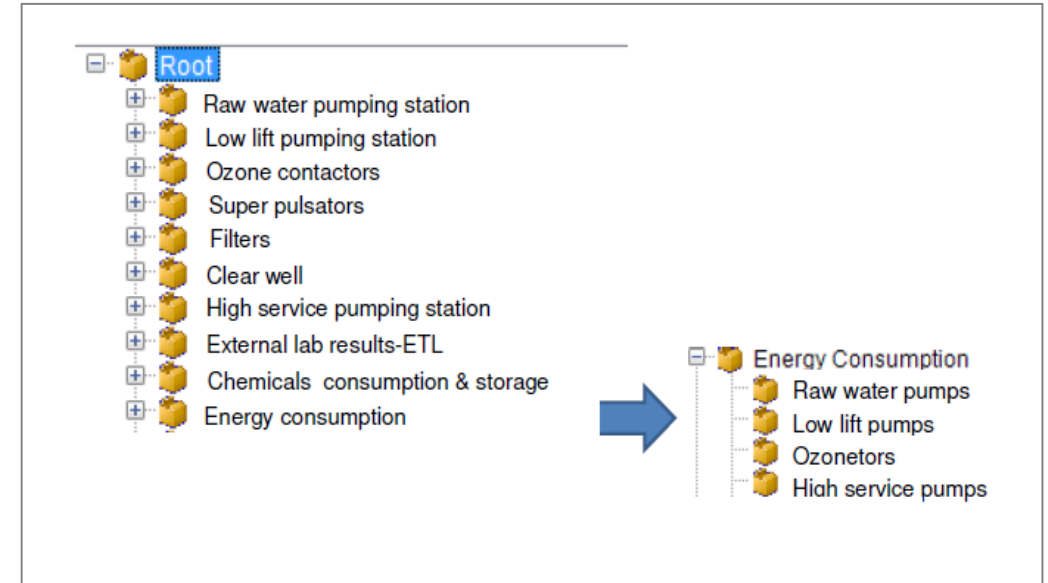
# Rule # 1

- Invest in the design stage:
  - Prepare a **thorough base of design spec**
  - Define a product tree the will be **intuitive for the users**
  - **Agree** with the customer on the basic reasons/ **triggers** that initiated the project and secure a continuous use
  - **Agree** with the customer on the **initial outputs**

Level 1	Level 2	Level 3	Level 4	Level 5	Tag Name	Engineering Units		
HBAF	Feed	Raw Water Tank Level					m	
		Pumps	Feed Pump A Status					
			Feed Pump A Current					A
			Feed Pump B Status					
			Feed Pump C Status					
			Feed Pump D Status					
		Feed Pump E Status						
		Blowers	Blower A Status					
			Blower A Current					A
			Blower B Status					
	Blower B Current					A		
	Blower C Status							
	Group A	Group A Feed Flow					m <sup>3</sup> /hr	
		HBAF 1 Level					m	
HBAF 2 Level					m			
Group A Outlet DO					mg/l			
Group B Feed Flow					m <sup>3</sup> /hr			
HBAF 3 Level					m			
Group B	HBAF 4 Level					m		
	Group B Outlet DO					mg/l		
	MF Feed Tank Level					m		
	MF Feed Turbidity					NTU		
MF	Feed	MF Feed Pressure				MPa		
		Pumps	Pump A Status					

## ➤ Engineering led modelling

- Prepare a **product tree** that reflects the **engineering aspects** on the system (rather than the automation aspects)
- Collect only the necessary tags
- Virtualize missing tags or unreliable processes (like Pump in rest status while flow parameter showing results)
- Standardize unit and tag properties (no of digits, categories, formulas etc.)
- Minimize sampling rate





- **Use a proper cleansing methodology**
  - **Pre-processing** based on a predefined rule base
  - **Avoid** values that are in **mis-match** with the process (e.g. concentration of less the 0 , Etc)
  - **Reject** data collection of repeating values with **no significant** change (dead band, change)
  - **Create calculated tags** that are based on a validated tested algorithm
  - **Compare** aggregated data and multiple sources of data that refers to the same entity (e.g Manual Meters VS AMI), and document the logical 'right' value



## ➤ Outputs normalization

- Try to **be consistent** with formulas online VS reporting and dashboarding rule base
- Build **frequently used calculated tags** to be used on multiple outputs
- **Be coherent** when building balances and complicated formulas and reports
- **Do not create dependency** on data that has to be extracted repeatedly with growing volumes – it will eventually miss the original performance objectives
- **Do not launch** complex analytics oriented projects before ensuring the data is 100% reliable and normalized!



## Choose the right platform and implementation team

- It should be **all in one** – collect- cleanse- store- present and analyze
- It must be **easy to use and configurable** so that the direct user will be able to take control on a certain point of time and take it forward independently
- The implementors must be **engineers that can understand** the need to the extent of possessing basic know how of what's best for the client
- **Data replication does not go together with normalized data!**
- Make sure the system **alerts** on missing datasets, anomaly results and duplications

**Remember: Preparing the actual report or dashboard should be a straight-forward procedure – if you invest more than that – there is either something wrong on the deployment methodology or in choosing the platform**

# Thank You

---

