UFZ HELMHOLTZ Zentrum für Umweltforschung

BIODT biodiversitydigitaltwin

Designing Dynamic Data-Driven Digital Twin Systems in Ecology

Taimur Khan (UFZ, Community Ecology)

ECEM23, September 2023

Funded by the European Union

Time	Activity
~ 45 mins	Introduction + Topical Lecture
~ 15 mins	Q&A
~ 10 mins	Pause
~ 45 mins	Exercise: Sample DT Design Schema
~ 10 mins	Close up

- No coding today, rather a conceptual discussion.
- Systems Design is subjective, there are always many ways to design things.
- Workshop website: https://biodt.github.io/dddas4dt/

- **Project name:** Biodiversity Digital Twin for Advanced Modelling, Simulation and Prediction Capabilities (BioDT)
- **Call title:** Next generation of scientific instrumentation, tools and methods (HORIZON-INFRA-2021-TECH-01)
- **Duration:** 1 June 2022 – 31 May 2025
- **Consortium:** 22 partners
  - 12 countries: Finland (FI), Italy (IT), Czech Republic (CZ), the Netherlands (NL), Estonia (EE), Sweden (SE), United Kingdom (UK), Germany (DE), Austria (AT), Denmark (DK), Norway (NO), Spain (ES)
    - Incl. one Affiliated Entity and three Associated Partners
- **Work Package (WP) members:** 140+
- **Coordinator:** CSC – IT Center for Science
- **Website:** www.biodt.eu

A digital twin is a virtual representation of real-world entities and processes, synchronized at a specified **frequency** and **fidelity***

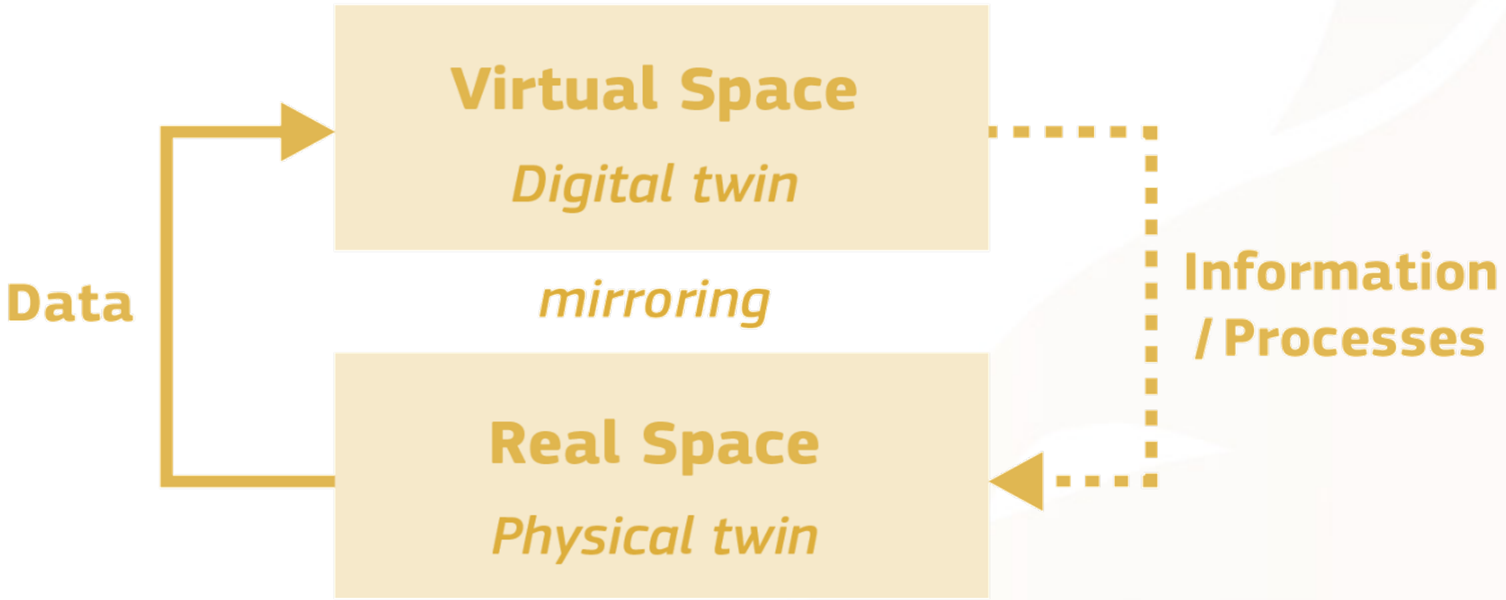


Image: digital-strategy.ec.europa.eu

**Here, fidelity refers to the level of precision captured by the DT in comparison with its physical counterpart.*

A **digital twin** (DT) is typically composed of:

- Data
- A model that is the representation in terms of behaviour and
- An application that connects the data and model in a way that makes the outputs of the model relevant, given the specific purpose of the DT

Since different scopes require different behaviour and fidelity, there cannot be a single twin answering all possible questions

Industrial DTs typically facilitate:

- Product design
- Operation of machinery

In **BioDT**, DTs used to:

- Mimic behaviour observed in nature
- Meet requirements of BioDT Use Cases
- Contribute toward EC goal of devising a full DT of the Earth

Use Cases split into four groups

Species response to environmental change

- Biodiversity dynamics
- Ecosystem services

Genetically detected biodiversity

- Crop wild relatives and genetic resources for food security
- DNA detected biodiversity, poorly known habitats

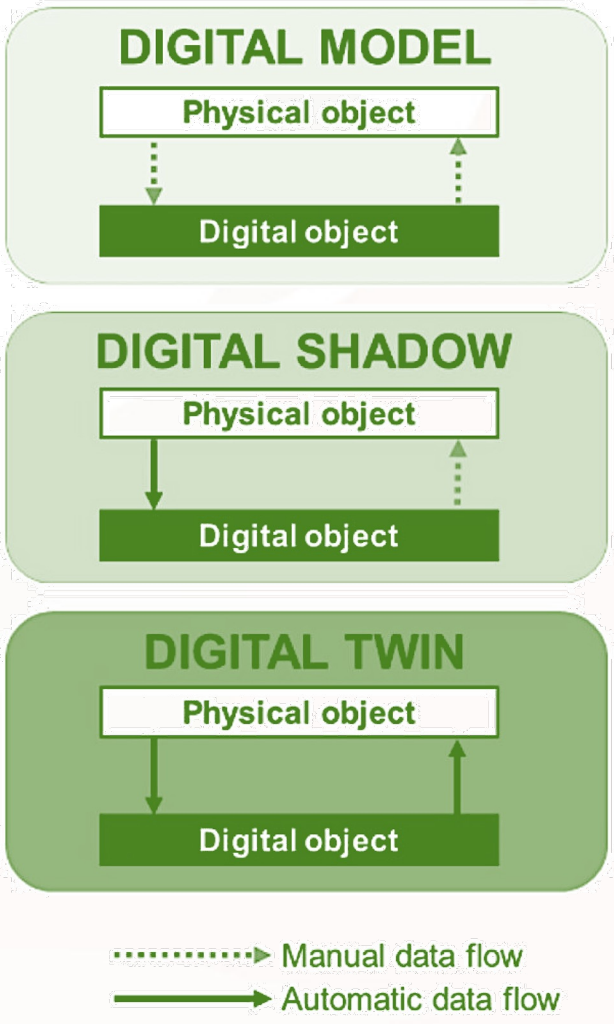
Dynamics and threats from and for species of policy concern

- Invasive species
- Endangered species

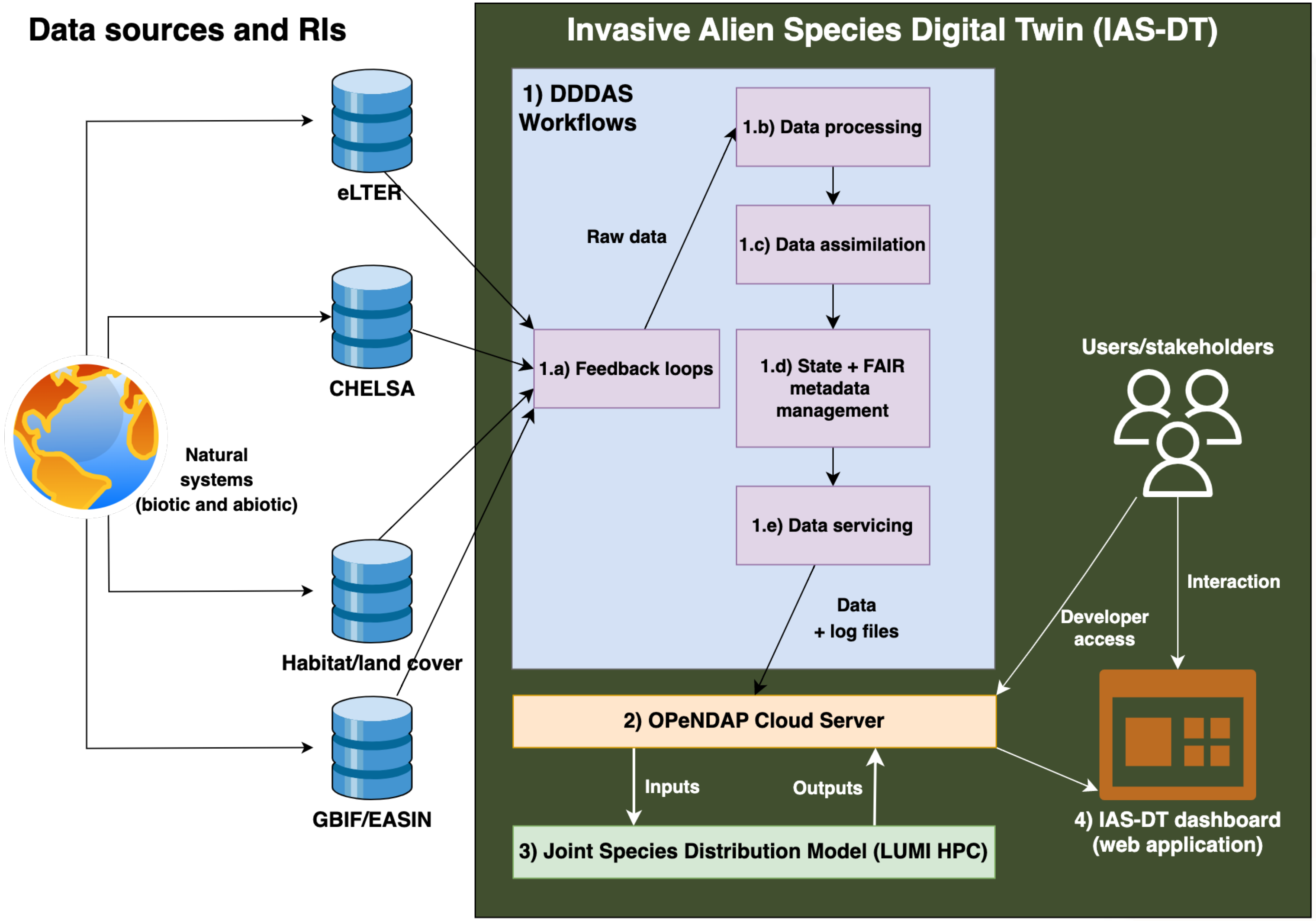
Species interactions with each other and with humans

- Disease outbreaks
- Pollinators

The main difference between **DTs**, **digital shadows** and **digital models** is the nature and direction of the data flow between the physical and virtual systems.

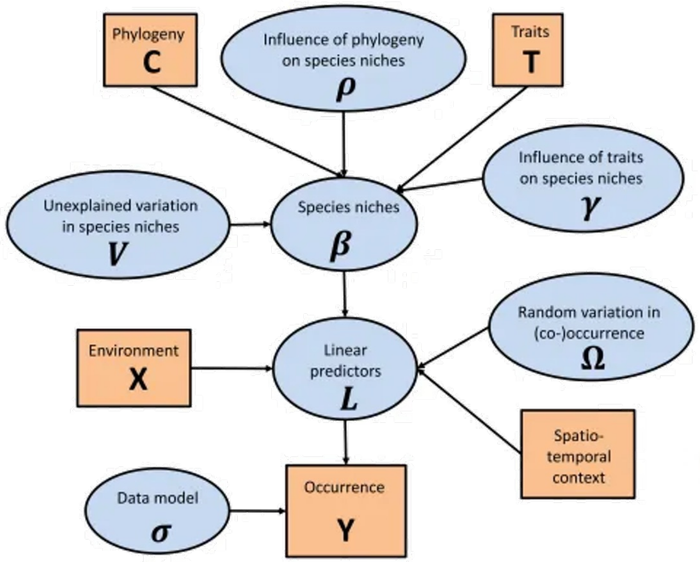


[Source: Open Engineering]

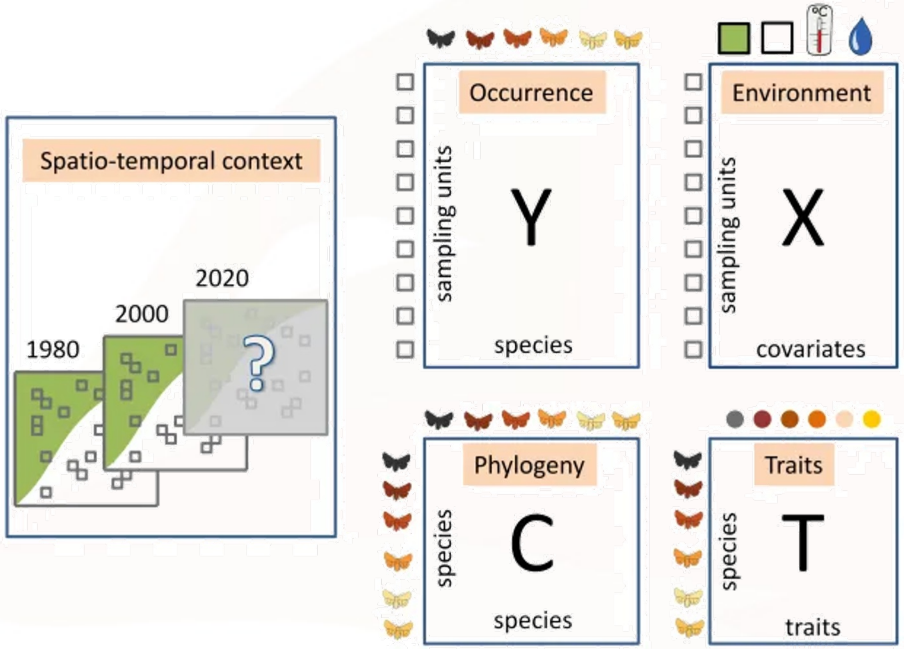


Architecture for Invasive Alien Species Digital Twin (IASDT). *Source: Taimur Khan*

- Predictive Digital Twin.
- State data ranging in ~ 100s of GBs.
- No direct data collection/sensor access.
- SDM = Hierarchical Modelling of Species Communities (HMSC).



Source: Ovaskainen et al. 2017a

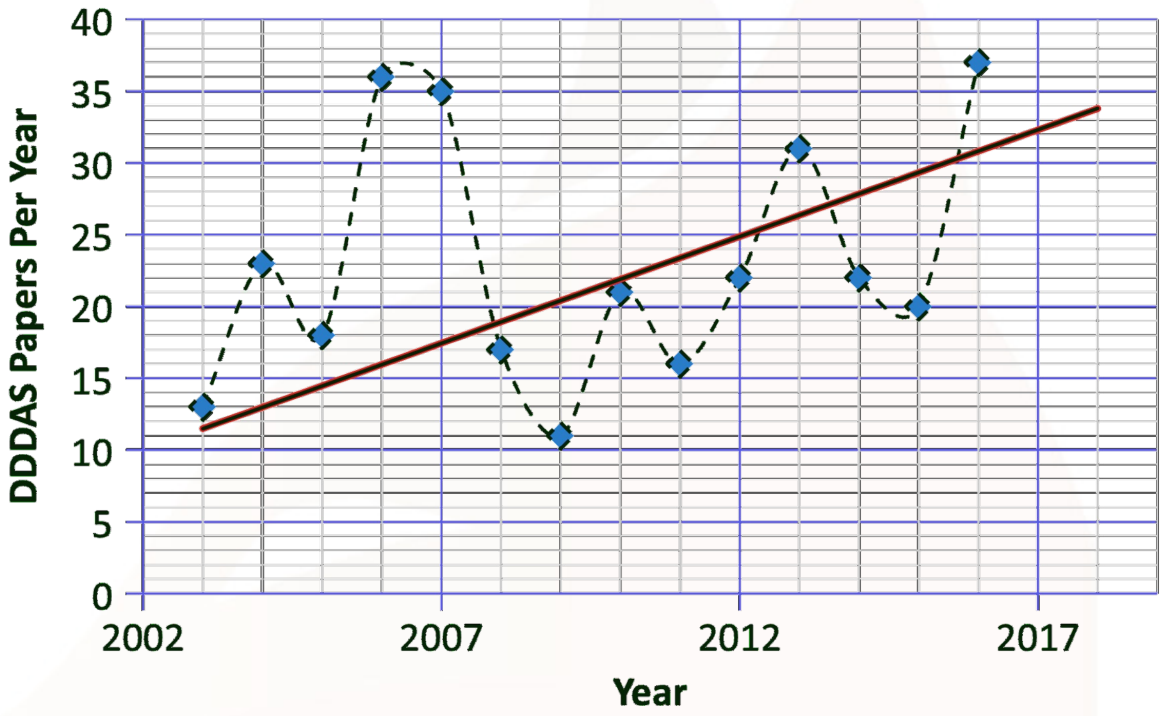


DT Challenges in Ecology

- No mature DT systems exist, hence a common design is a not clear.
- DT tools are limited to other niches.
- Literature is sparse for biodiversity DTs systems design.
- Datasets are updated infrequently, with often lots of heterogeneity.
- Researchers mostly working with “indirect” data collection.

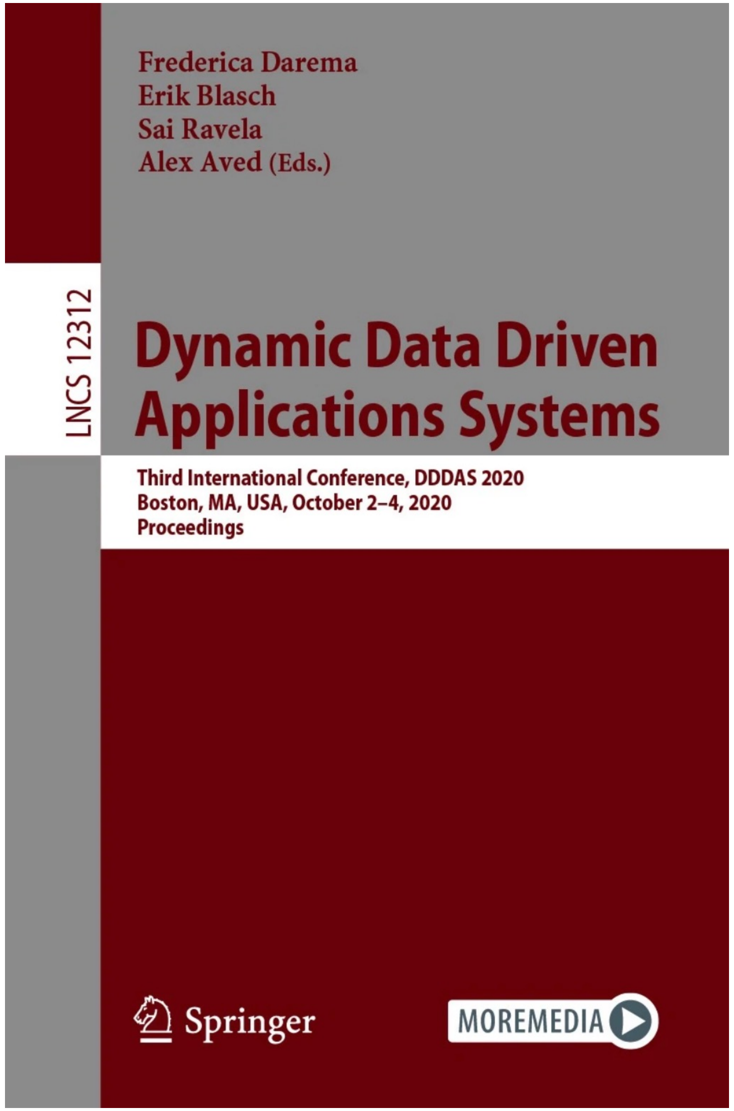
The Data Driven Applications Systems (DDDAS) concept entails "*the ability to dynamically incorporate data into an executing application simulation, and in reverse, the ability of applications to dynamically steer measurement processes*",

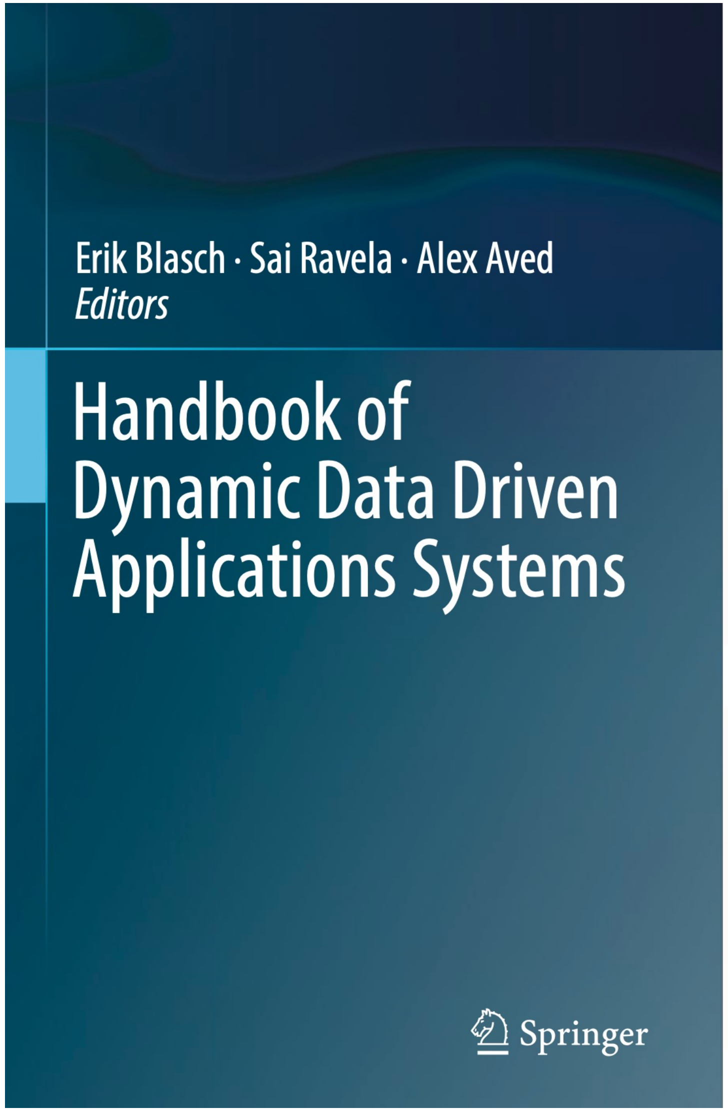
creating "*application simulations that can dynamically accept and respond to 'online' field data and measurements and/or control such measurement.*"

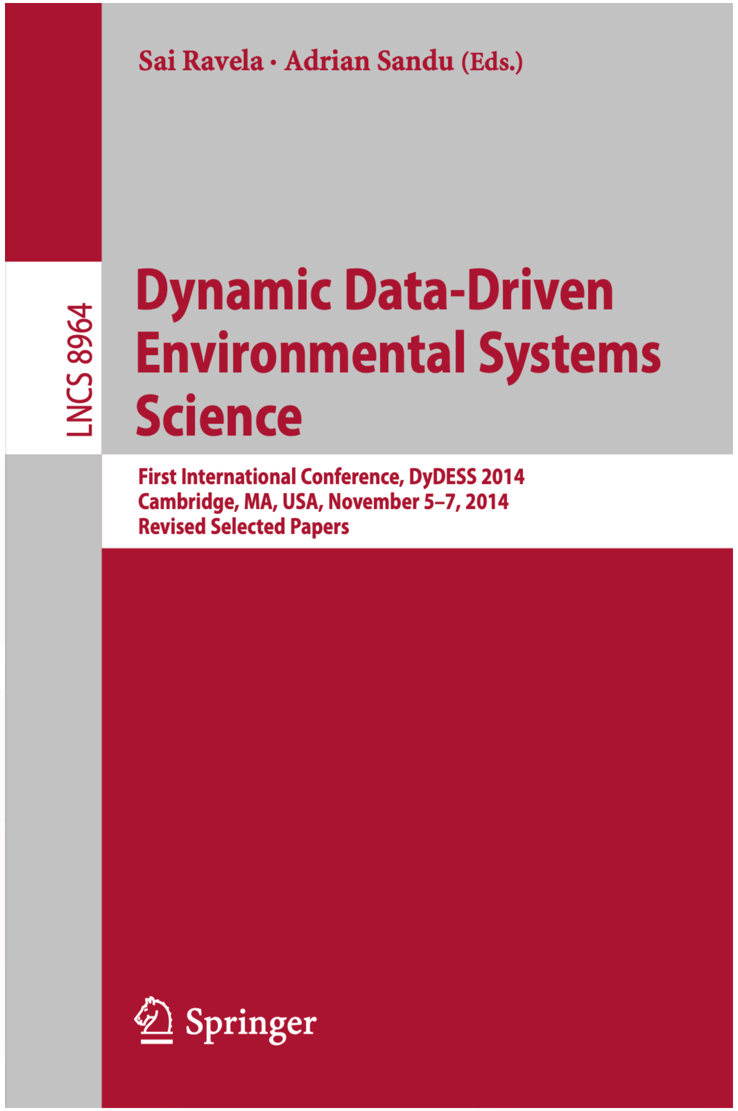


DDDAS papers per year. Source: Handbook of DDDAS (2018).

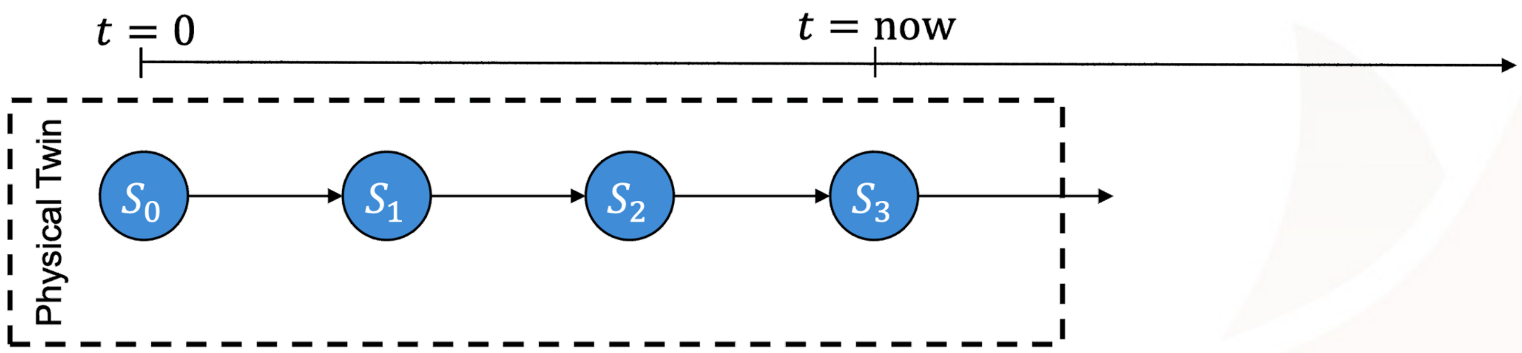
Why DDDAS?


Frederica Darema
Erik Blasch
Sai Ravela
Alex Aved (Eds.)
LNCS 12312
Dynamic Data Driven Applications Systems
Third International Conference, DDDAS 2020
Boston, MA, USA, October 2–4, 2020
Proceedings
Springer
MOREMEDIA


Erik Blasch · Sai Ravela · Alex Aved
Editors
Handbook of Dynamic Data Driven Applications Systems
Springer

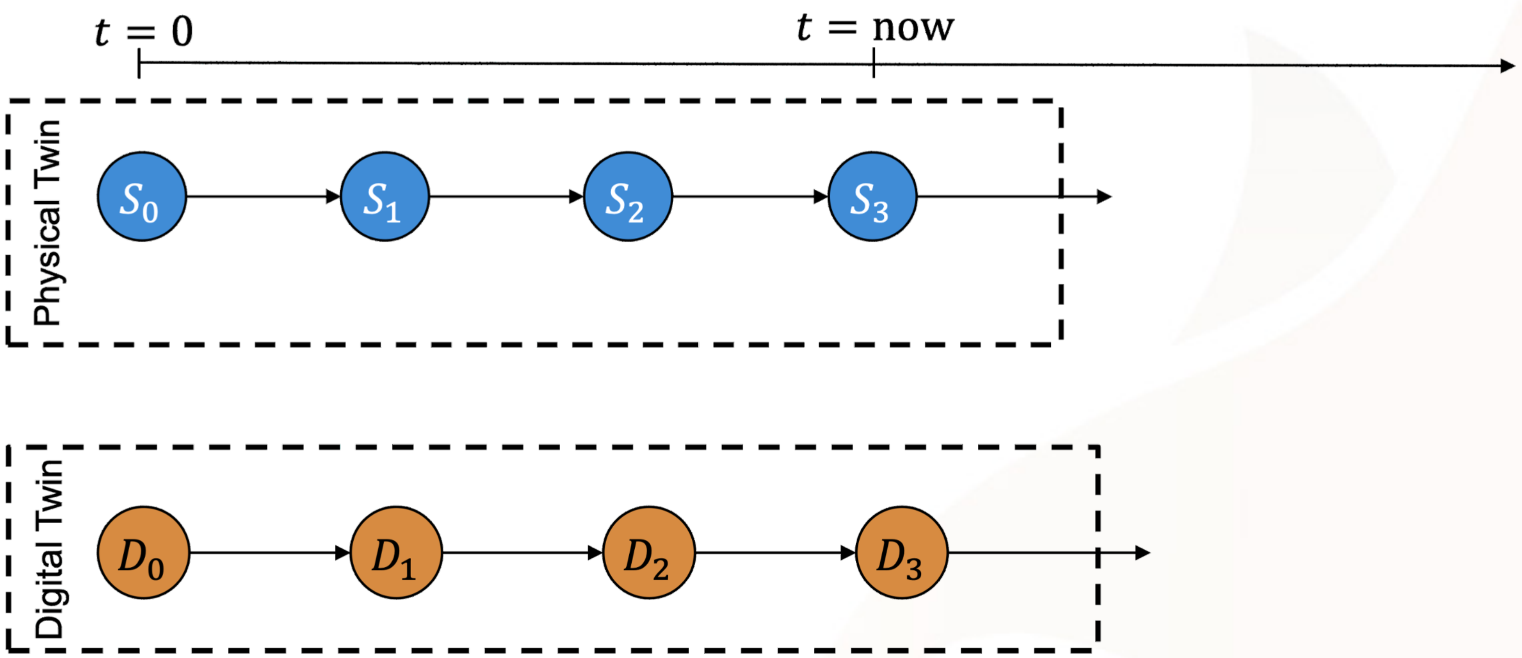

Sai Ravela · Adrian Sandu (Eds.)
LNCS 8964
Dynamic Data-Driven Environmental Systems Science
First International Conference, DyDESS 2014
Cambridge, MA, USA, November 5–7, 2014
Revised Selected Papers
Springer

Physical State, S:
Parameterized state of the physical asset



Physical State, S:
Parameterized state of the physical asset

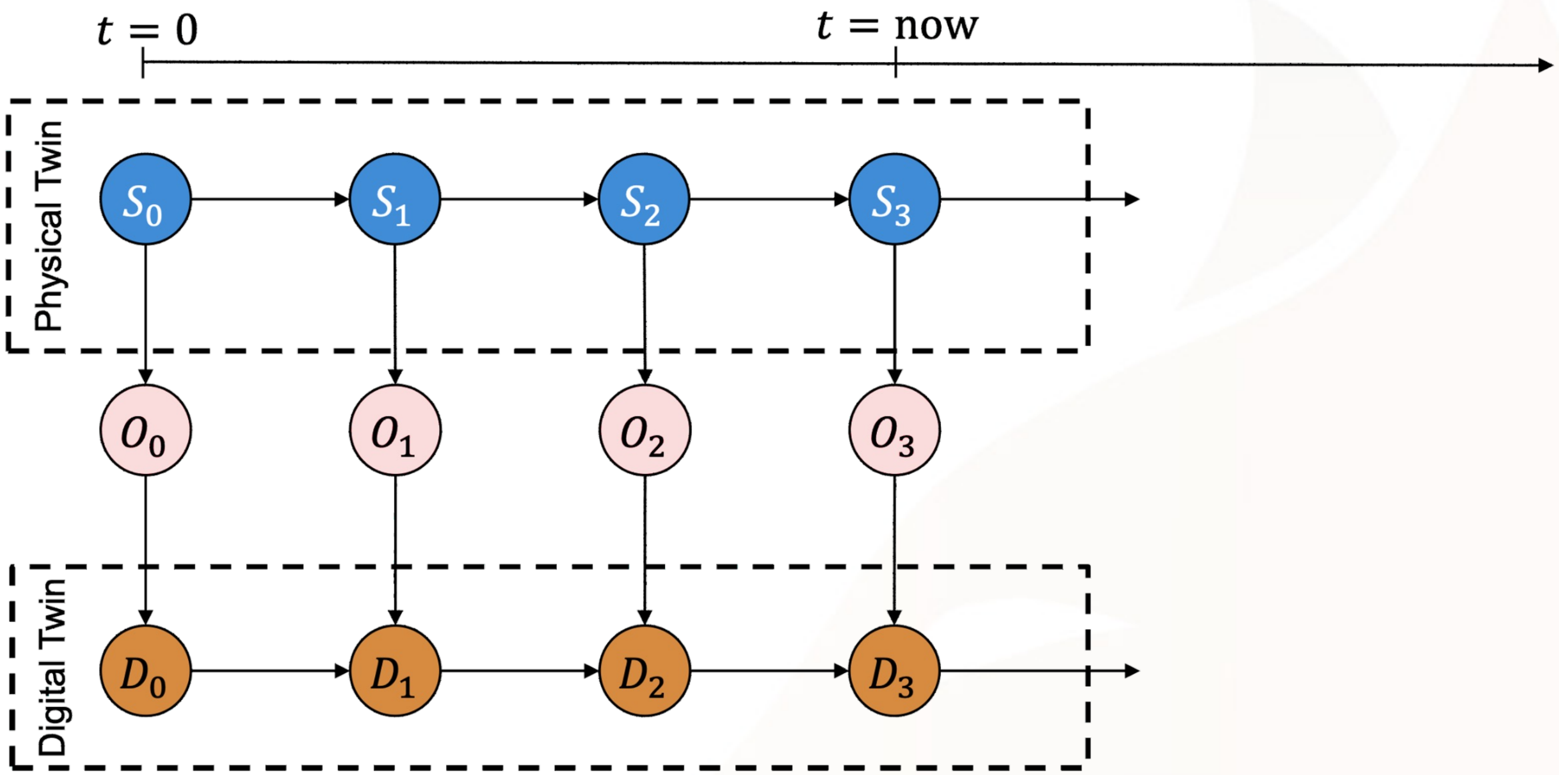
Digital State, D:
Parameters (model inputs) that define the computational models comprising the digital twin



Physical State, S:
Parameterized state of the physical asset

Digital State, D:
Parameters (model inputs) that define the computational models comprising the digital twin

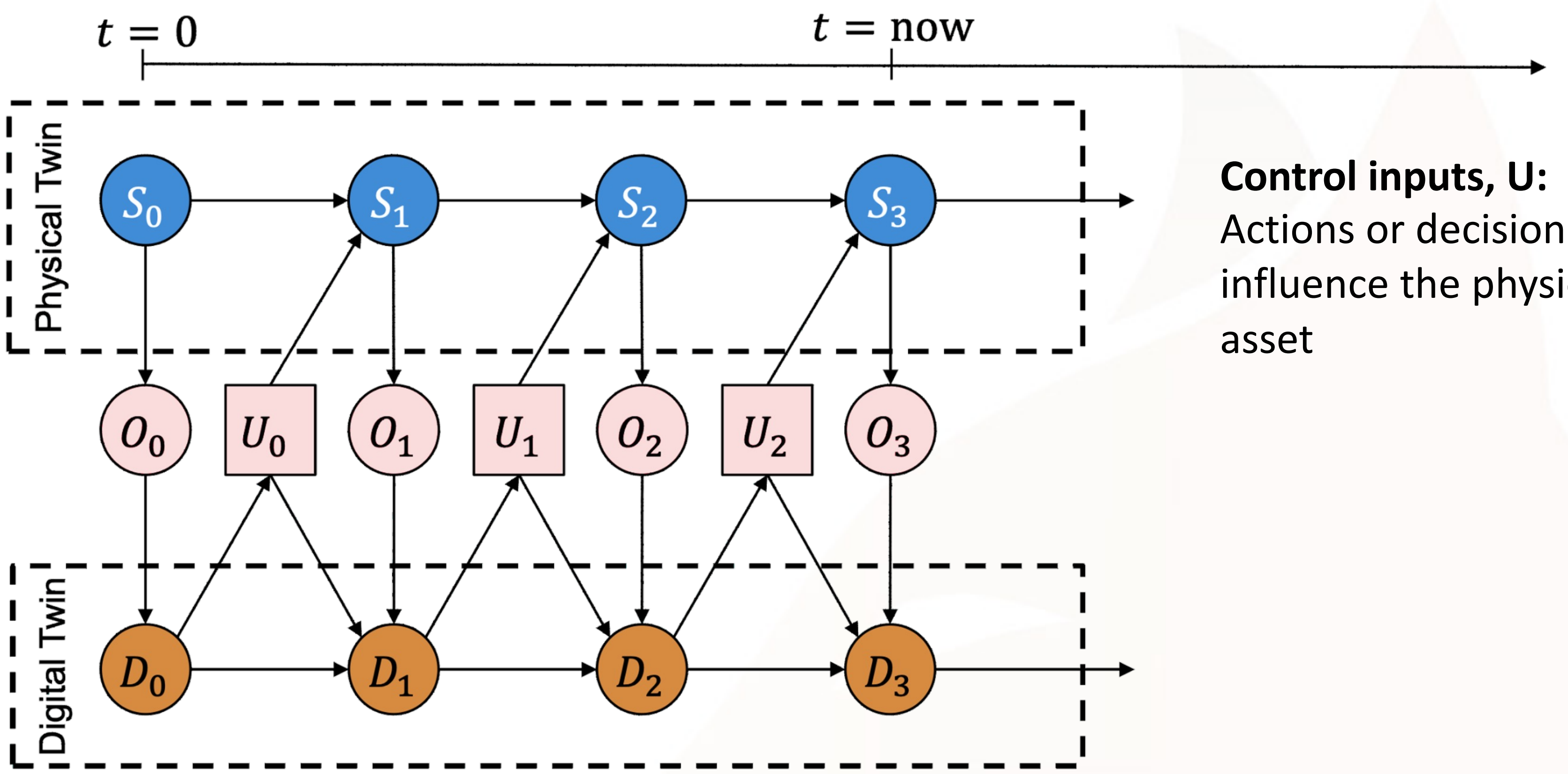
Observational data, O:
Available information describing the state of the physical asset



Physical State, S: Parameterized state of the physical asset

Digital State, D: Parameters (model inputs) that define the computational models comprising the digital twin

Observational data, O: Available information describing the state of the physical asset

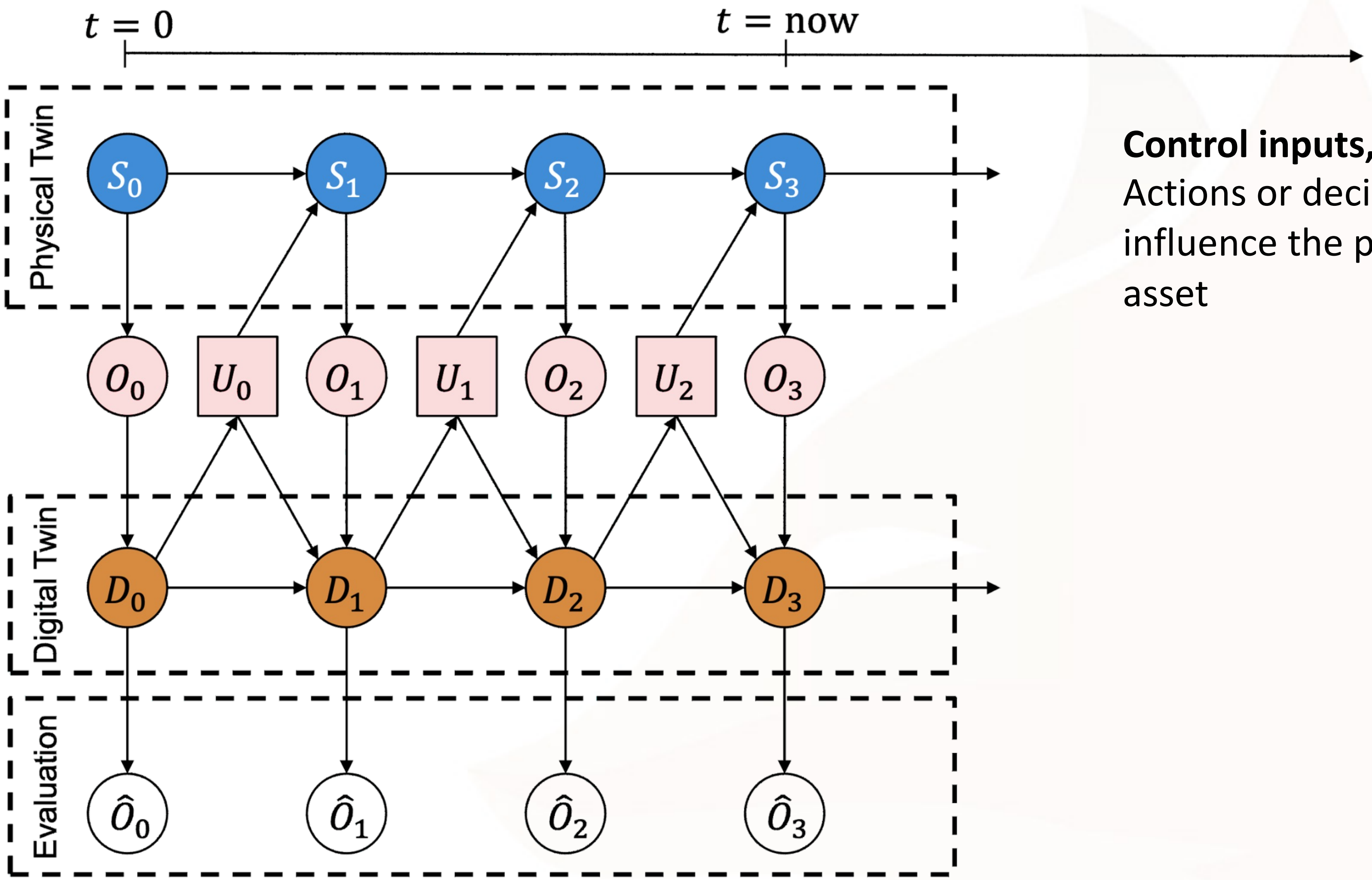


Control inputs, U: Actions or decisions that influence the physical asset

Physical State, S:
Parameterized state of the physical asset

Digital State, D:
Parameters (model inputs) that define the computational models comprising the digital twin

Observational data, O:
Available information describing the state of the physical asset

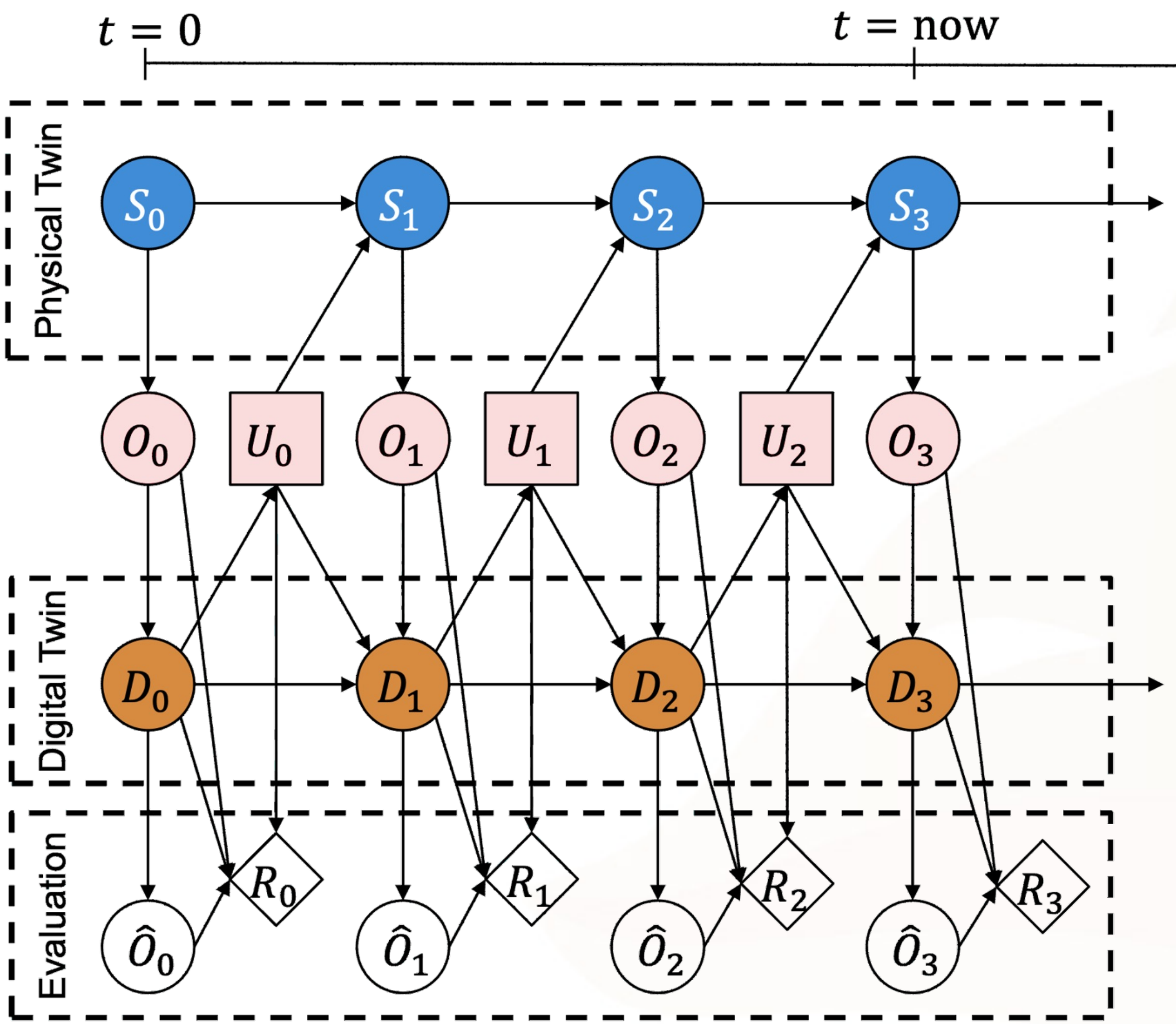


Control inputs, U:
Actions or decisions that influence the physical asset

Physical State, S: Parameterized state of the physical asset

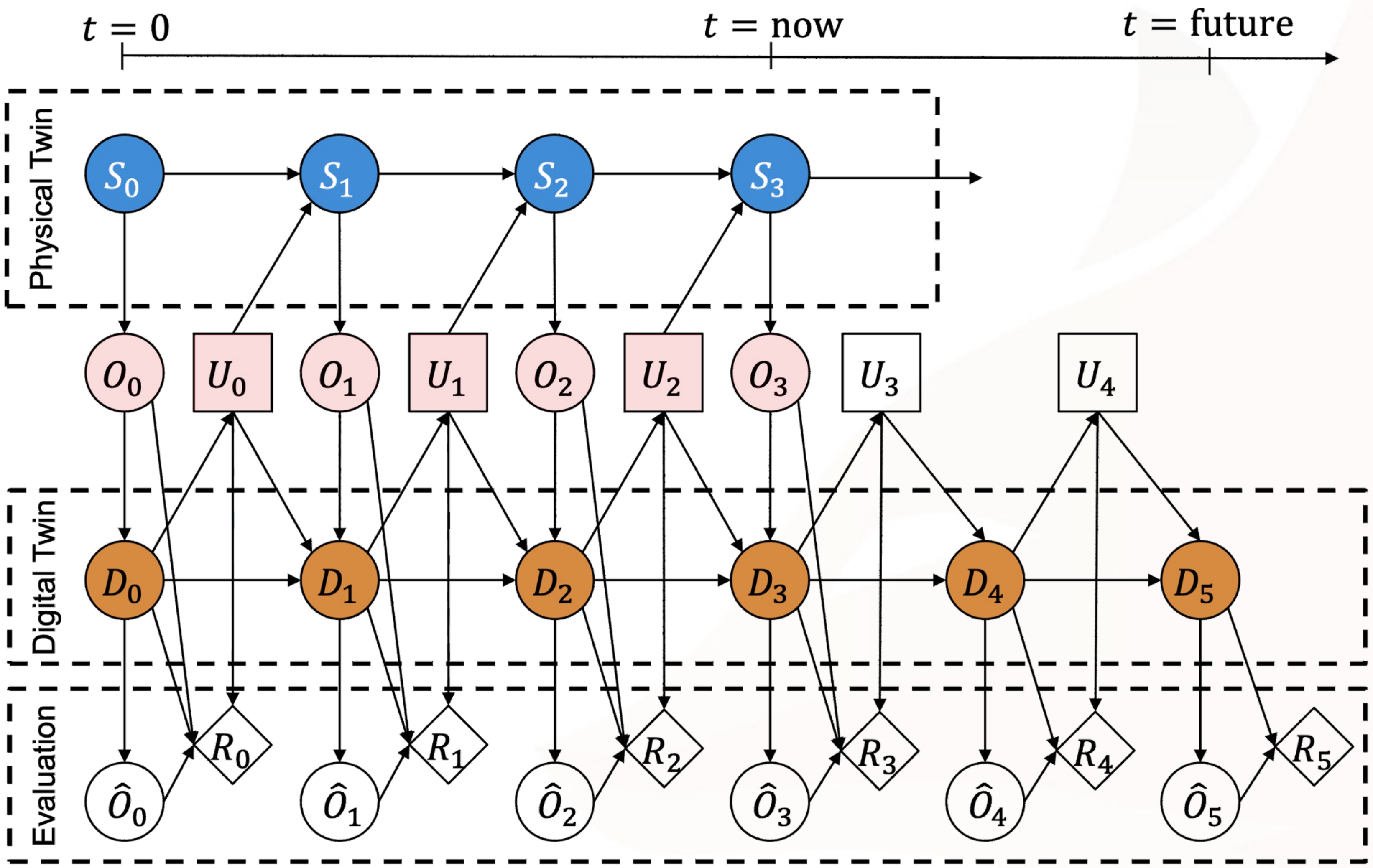
Digital State, D: Parameters (model inputs) that define the computational models comprising the digital twin

Observational data, O: Available information describing the state of the physical asset



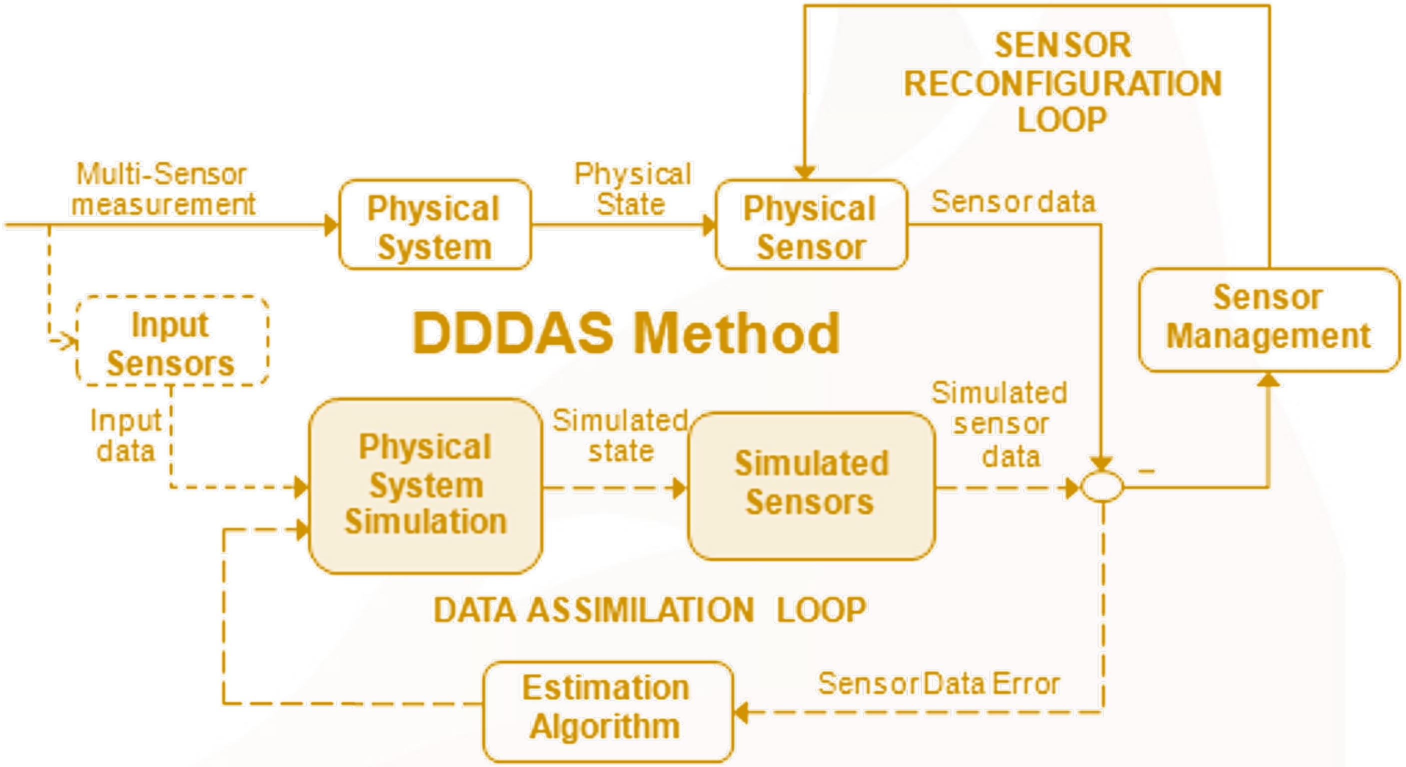
Control inputs, U: Actions or decisions that influence the physical asset

Reward, R: Quantifies preference of different states or trajectories of the asset-twin system



t = 0
t = now
t = future
Physical Twin
Digital Twin
Evaluation
S_0 S_1 S_2 S_3
O_0 U_0 O_1 U_1 O_2 U_2 O_3 U_3 U_4
D_0 D_1 D_2 D_3 D_4 D_5
R_0 R_1 R_2 R_3 R_4 R_5
$\hat{O}_0$ $\hat{O}_1$ $\hat{O}_2$ $\hat{O}_3$ $\hat{O}_4$ $\hat{O}_5$

- Based on Dynamic Data-Driven Application Systems(DDDAS) paradigm.
- DDDAS Components:
  - Feedback loop.
  - State management.
  - Sensor reconfiguration loop.
- Other components:
  - Data servicing.
  - Model.
  - User interface.



Sample DDDAS system. *Source:* 1dddas.org

DDDAS Workflow Components and their roles in Ecological DTs

Feedback Loop

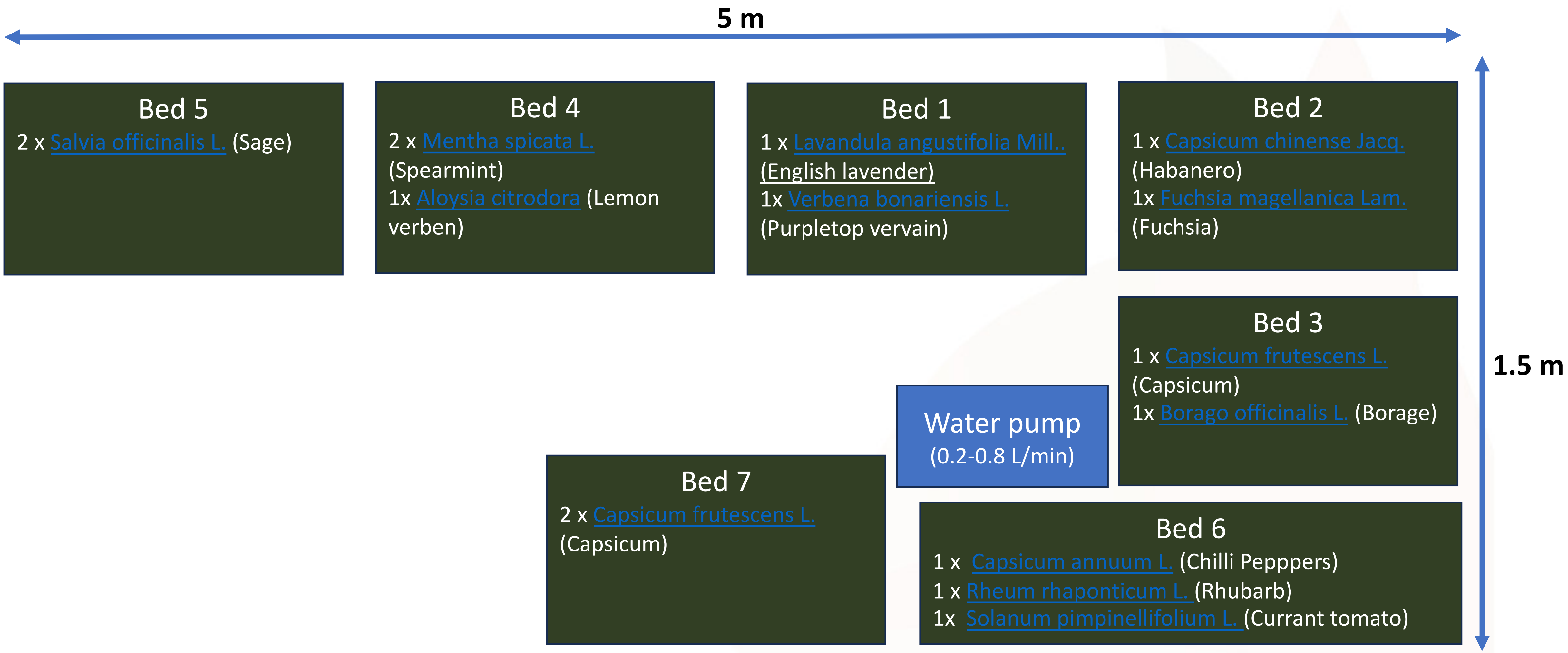
- **vSensors** - "Listen" for changes in data sources.
- **Intakers** - Pull new data.
- **Processors** - Process the data.
- **Assimilators** - assimilate the data into existing datasets.
- **Actuators** - change control inputs.
- **Loggers** - Assign FAIR metadata.

State Management

- **State Space** – defined states of the system.
- **Trackers** - Track state of the data.
- **Synchronizers** - State synchronization.
- **Sniffers** - Detect changes in state of the DT.

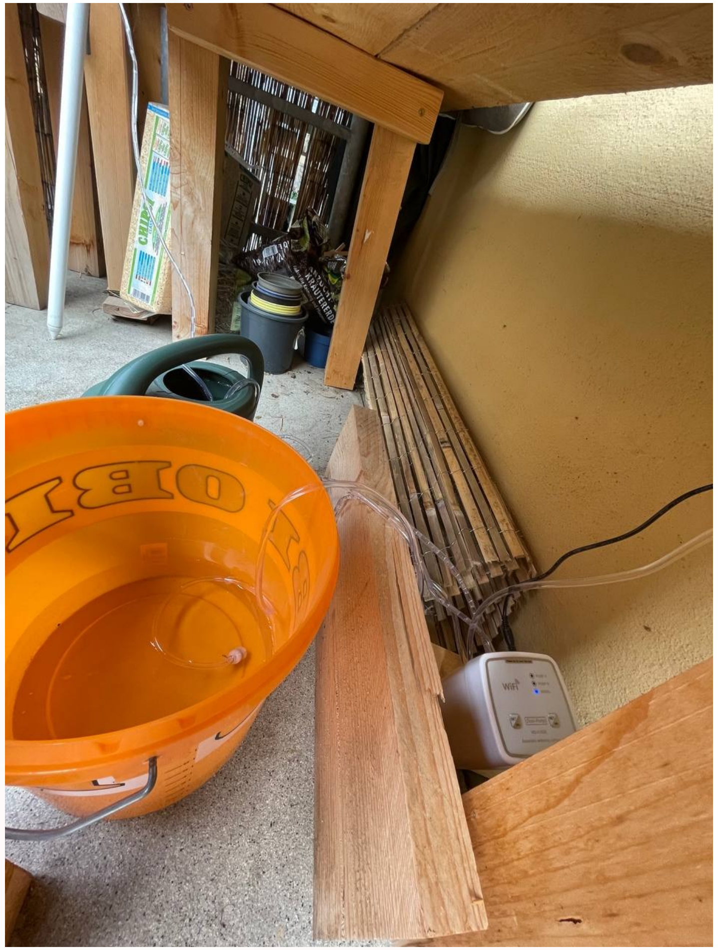
- **System:** Soil Watering DT
- *8 soil beds* with sensor network and watering system
- **Observational Data (from sensors):**
  - Soil moisture (%)
  - Soil temperature (F)
- **Control inputs:**
  - WiFi controlled water pumps with on/off states.
- **Model:** ? *(e.g. linear regression, rates-of-change)*
- Use the DT schema template on draw.io to create a DDDAS-based DT of the given Soil-Plant system that automates soil watering based on soil moisture and soil temperature data.

HINT: Think about what other data sources can be added, what is the state space, what type of model, is needed and what components would be required.

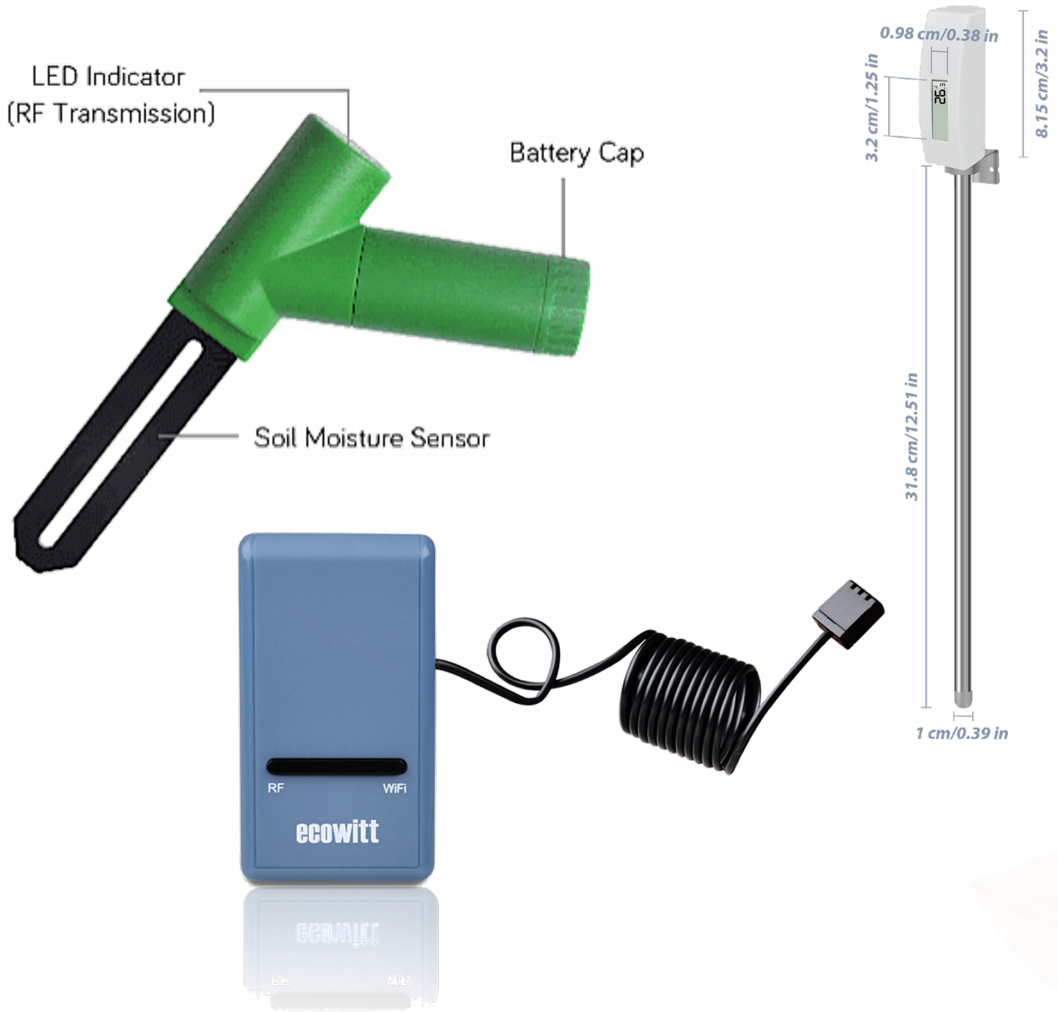


Relative positions and numbering of soil beds and water pump. *Source: Taimur Khan.*

Exercise: Design a DT around given system



Exercise: Design a DT around given system



Soil moisture sensor, soil temperature sensor, base station. *Source: ecowitt.com*

WiFi enabled dual motor water pump. *Source: Cikonielf.*

UFZ HELMHOLTZ
Zentrum für Umweltforschung

BIODT
biodiversitydigitaltwin

@BiodiversityDT

BioDT

Funded by
the European Union