Check for updates

RESEARCH ARTICLE

An extension of Thing Descriptions from the Web of Things for Digital Twins

[version 1; peer review: 3 approved with reservations]

Salvador González-Gerpe[1], Andrea Cimmino[1], Socorro Bernardos[1], Raúl García-Castro[1], María Poveda-Villalón[1], Kyriakos Katsigarakis[2], Georgios N. Lilis[2], Dimitrios Rovas[2]

[1]Universidad Politécnica de Madrid, Madrid, Spain
[2]University College London, London, UK

v1 **First published:** 21 Dec 2022, **2**:144
https://doi.org/10.12688/openreseurope.15280.1
Latest published: 21 Dec 2022, **2**:144
https://doi.org/10.12688/openreseurope.15280.1

Abstract

Background: Digital Twins are becoming ubiquitous in a range of domains, such as construction or e-health. Due to the numerous applications where Digital Twins can be adopted, they have to manage heterogeneous cross-domain data that is consumed by third-party service or domain experts. Digital Twins may allow for performing operational commands that are transformed into actions in the physical world; these operations can be grouped into processes. In the domain of construction, a European Project named COGITO was proposed, in which a Digital Twin is used to simulate the construction processes carried out in three railway stations in Europe (Spain, Denmark and Austria). The lack of interoperable data due to heterogeneity in the APIs of the physical layer, the complexity that entails managing these data at the digital model, and the difficulty of orchestrating the data exchanged among the advanced features have been the roadblocks in the project.
Methods: The problems have been addressed by adopting the W3C Web of Things (WoT) standard, by using Thing Descriptions. Although the challenges are specific for the COGITO project, they could be found in any Digital Twin proposal. Thus, the article presents how they are solved and addressed in COGITO.
Results: The results demonstrate that relying on a Knowledge Graph for the provenance, another for describing a distributed architecture of semantic interoperability services and a final one for data, allows to build a distributed Digital Twin over heterogeneous data sources that provide semantic interoperable data in each digital layer.
Conclusions: The lack of a Digital Twin standard highlights the necessity of adopting existing standards and extending them in order to adapt their suitability to the Digital Twin context. Hence, the article presents how WoT is helpful for addressing the challenges, and how it could be extended in the future to support Digital Twins.

Open Peer Review

Approval Status ? ? ?

	1	2	3
version 1 21 Dec 2022	? view	? view	? view

1. **Alessandro Ricci**, University of Bologna, Bologna, Italy
2. **Marco Picone**, Universita degli Studi di Modena e Reggio Emilia, Modena, Italy
3. **Riku Ala-Laurinaho**, Aalto University, Espoo, Finland

Any reports and responses or comments on the article can be found at the end of the article.

Keywords
Digital Twins, Linked Data, WoT, Thing Descriptions.



This article is included in the Horizon 2020 gateway.



This article is included in the Digital Twins collection.



This article is included in the Sustainable Places 2022 collection.

Corresponding author: Salvador González-Gerpe (salvador.gonzalez.gerpe@upm.es)

Author roles: González-Gerpe S: Conceptualization, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Cimmino A**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Bernardos S**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **García-Castro R**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Poveda-Villalón M**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Katsigarakis K**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **N. Lilis G**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Rovas D**: Conceptualization, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958310 (COGITO).

Copyright: © 2022 González-Gerpe S *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: González-Gerpe S, Cimmino A, Bernardos S *et al.* **An extension of Thing Descriptions from the Web of Things for Digital Twins [version 1; peer review: 3 approved with reservations]** Open Research Europe 2022, **2**:144 https://doi.org/10.12688/openreseurope.15280.1

First published: 21 Dec 2022, **2**:144 https://doi.org/10.12688/openreseurope.15280.1

Introduction

Digital Twins are currently growing in relevance across a number of domains, e.g. construction or e-health. Because of the wide variety of applications in which Digital Twins are used, they need to be able to handle heterogeneous data from different domains; i.e., the physical layer. Digital Twins need to harmonise these data in order to make it interoperable so it can be transparently consumed by domain experts through the digital layer provided by the Digital Twins. As a result, the data flow between the physical and the digital layer becomes bi-directional, entailing that changes at the digital layer may produce effects in the physical layer.

In the construction domain, particularly in railway stations, a European project called COGITO[1] (COnstruction-phase diGItal Twin mOdel) has been proposed. The COGITO approach is developing a Digital Twin for ingesting and man-aging data coming from a physical layer that consists on sensors, actuators, or static data files such as ifc, schedule or workers, and providing a digital layer containing the Digital Twin Platform and applications that use that data.

During the development of the COGITO project, three main challenges have been identified and addressed. I) Transparently consume the data coming from the physical layer due to the heterogeneity of the APIs within; to tackle this issue RDF generation services are used, which ingest data from a specific domain and produce semantic interoperable data that conforms the Digital Twin Knowledge Graph. II) Since the domains covered by COGITO are wide, the number of services generating RDF is numerous and, as result, given a specific piece of data choosing the correct service becomes challenging; to tackle this issue COGITO relies on the W3C standard Web of Things for generating a Knowledge Graph about the available Translation Services that is used for taking a suitable decision. III) The construction projects involve several sources from the physical layer and their related fragments in the Digital Twin Knowledge Graph. As a result, keeping track of all the data involved in a specific project from the layers physical and digital becomes a challenge; to tackle this issue COGITO relies on the W3C standard Web of Things for generating a Platform Knowledge Graph that relates for a given project all the sources from the physical layer and their corresponding RDF in the Digital Twin Knowledge Graph.

The fact that the physical layer contains data sources that provide heterogeneous data entails that the COGITO Digital Twin Platform is not able to transparently consume these data. Therefore, a semantic interoperable approach is required for providing a digital layer that is transparently consumed by domain experts. To this end, COGITO counts with an ontology and demands that all data be expressed according to it. The difficulty in orchestrating the data exchanged be-tween the advanced functions is another challenge, because the data belonging to different formats cannot be correctly translated into interoperable data, so the data to be translated have to be orchestrated to make the correct translation. Finally, the management of these data in the digital model is the last challenge due to its complexity, as it does not allow existing data within the platform to be referenced to other data belonging to the same particular process, making it impossible to differentiate between them.

In the following, the above-mentioned challenges will be solved in the section of “Our proposal”, later demonstrated by means of the examples made and the results obtained.

Our proposal

The COGITO project has been designed for ingesting numerous sources that belong to a wide range of domains, as shown in Figure 1; such as satellite images, 3D scans, devices and sensors, meteorological, spatial, task and cost data. For this scope in this project the COGITO Digital Twin Platform has been developed, which uses data from the physical layer, which is conformed by the heterogeneous sources, and provides a set of Knowledge Graphs at the digital layer. Thanks to this last layer that provides data expressed according to the COGITO ontology, third-party are able to transparently consume and develop applications for different domains, such as Health & Safety, Quality Control, Adaptative Workflow Management and GUI & support Apps.

Challenge I - Dealing with data heterogeneity from the physical layer

Due to the large amount of data belonging to different domains and their heterogeneity in terms of format and model, DTP cannot transparently consume these data. The transformation of these data into an equivalent version that is semantically interoperable is the first challenge identified in the COGITO project. The solution

[1] https://cogito-project.eu/

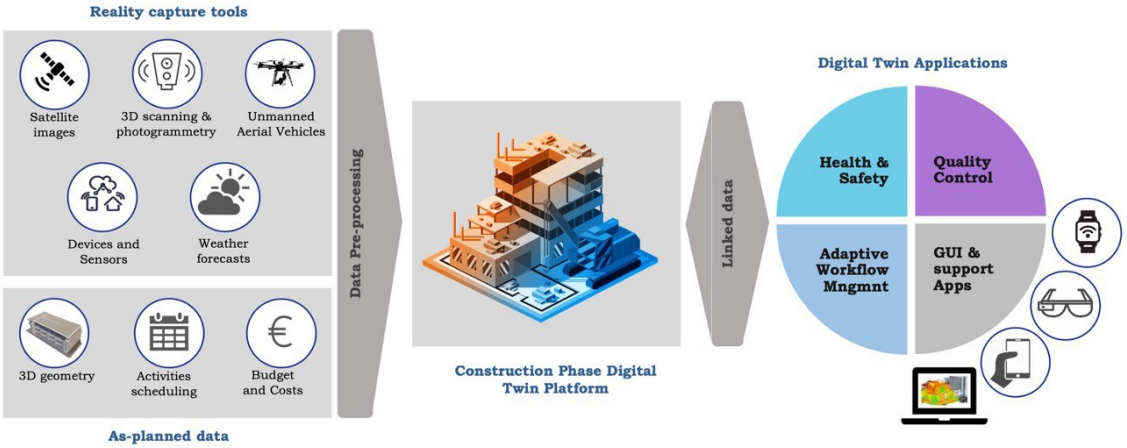

Figure 1. COGITO General Architecture.

to this challenge is provided by the component Knowledge Graph Generator[1], as depicted in Figure 2. This component is a service defined for a specific data source that first pre-processes the heterogeneous data, then translates these data into RDF expressed according to the COGITO Ontology[2], and finally validates the RDF. The data resulting from this process are stored in the COGITO Data Persistence Layer inside the Digital Twin Knowledge Graphs.

The Knowledge Graph Generator integrates the Helio tool[2], which relies on Mappings that hold a set of ad-hoc trans-lation rules for a specific source of data. For instance if a construction has schedules in XML and another in CSV two mappings are required, respectively. In addition, the Knowledge Graph Generator validates the RDF data using the W3C standard SHACL shapes[3] which are made using the Astrea tool[4].

Challenge II - Orchestrate distributed semantic interoperable architecture

Since the Knowledge Graph Generator must be defined for a specific data source and the physical layer counts with numerous, more than one instance of this component is required. Also, consider that the Knowledge Graph Generator may consume considerable computational power, entailing that a unique instance could become a bottleneck.

To build an efficient architecture, the DTP counts with several Knowledge Graph Generator instances, one per each source. This architecture leverages the computational power and the parallel processing by distributing in different nodes the Knowledge Graph Generators. The second challenge is due to this architectural choice and consists of know-ing for a given source which service needs to be invoked.

The solution to this challenge is provided by the Thing Manager, as shown in Figure 3. The W3C Web of Things paradigm is adopted in the Digital Twin Platform, which demands that the services must be described by a Thing Description[5]. In the Digital Twin Platform these services are the Knowledge Graphs Generators and their Thing Descriptions are stored in the Translation Services Knowledge Graph. As a result, the Thing Manager is able to discover the suitable service that must be invoked for the data coming from a specific source that must be processed.

Challenge III - Organize data provenance under construction project

The Digital Twin Platform provides the Digital Twin Knowledge Graph through its digital layer. However, these data must be organized by the construction projects to which they belong, so for a given project all its related data is retrievable. Furthermore, the Digital Twin Platform demands to know the sources from the physical layer related to the data provided in the Digital Twin Knowledge Graph. The third challenge consists on grouping into projects different fragments of the Digital Twin Knowledge Graph and, also, relating these fragments to the physical layer.

[2]https://cogito.iot.linkeddata.es/

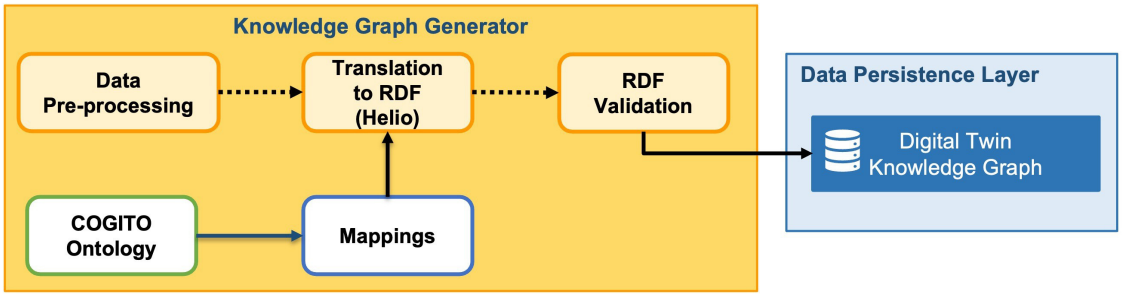

Figure 2. First challenge solution.

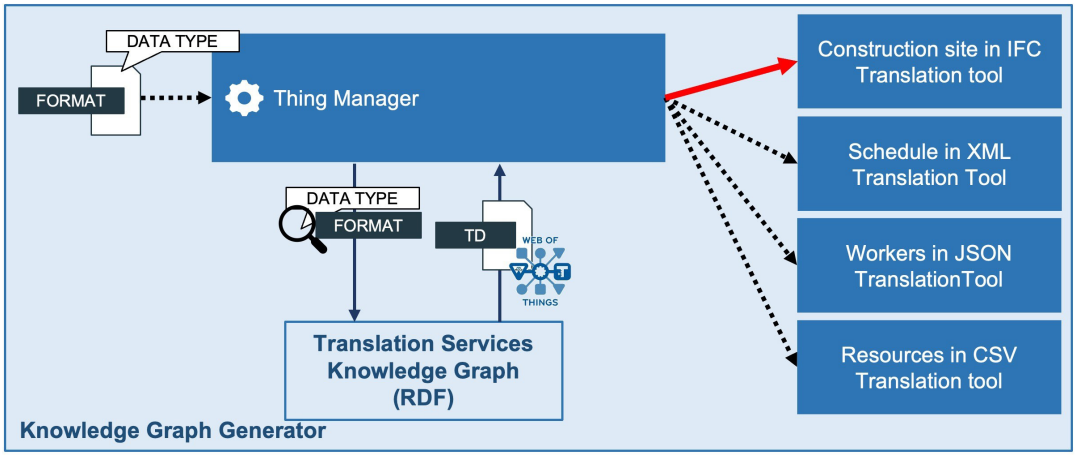

Figure 3. Second challenge solution.

The solution to this challenge is provided by the Thing Manager, as shown in Figure 4. The DTP creates a Thing Description for each new project. These Thing Descriptions are enriched with pointers to the data sources from the physical layer that are involved in that project. Then, the Thing Manager updates the Thing Description of a project with pointers to the Digital Twin Knowledge Graph anytime a Knowledge Graph Generator related to one of the sources described in its Thing Description is invoked.

As a result, the Thing Manager is able to discover the Thing Descriptions related to a specific project that contain the pointers to all the fragments in the Digital Twin Knowledge Graph related to it and, also, the pointers to the sources used to generate those fragments.

Experimental results

The COGITO project is implemented in pilots from different countries, namely: Austria, Denmark, and Spain. In these pilots, one of the most important data sources are those that provide data about the schedules. These contain a set of tasks that must be carried out in a construction project. These tasks must be assigned to the workers and associated with existing physical resources in the pilot. To showcase how the aforementioned challenges are addressed in COGITO, this section provides an insight of how the schedules are handled in the Digital Twin Platform.

On the one hand, in order to include schedule, workers, and resources data into the Digital Twin Knowledge Graph, three Knowledge Graph Generators must be deployed and used with their mapping. In order to find a suitable Knowledge Graph Generator for them the Thing Manager must perform a discovery over the existing ones in the Translation Services Knowledge Graph. Listing 1 shows the Thing Description that points to the three Knowledge Graph Generators, notice that each of them are deployed in different nodes.

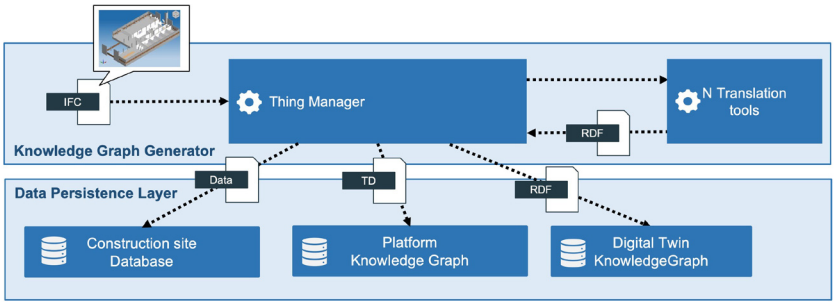

Figure 4. Third challenge solution.

```
{
  "@context":"https://www.w3.org/2019/wot/td/v1",
  "title":"Channel Thing Description",
  "id":"123456789",
  "description":"Channel Thing Description to retrieve channel for server sent events",
  "properties":{
      "Channels":{
         "forms":[
             {"href":"https://data.cogito.linkeddata.es/tm/stream?channel=file.workers","type":"workers"},
             {"href":"https://data.cogito.linkeddata.es/tm/stream?channel=file.schedule","type":"schedule"},
             {"href":"https://data.cogito.linkeddata.es/tm/stream?channel=file.res_types","type":"resource_types"}
         ]
      }
  },
  "security":"nosec_sc",
  "securityDefinitions":{"nosec_sc":{"scheme":"nosec"}}
}
```

Listing 1. Translation Services Knowledge Graph.

Once the suitable Knowledge Graph Generator is found, it must be invoked for generating the RDF. In this example, the schedule is an XML file[3], Listing 2 shows an excerpt. This file contains information about tasks, but not about workers and resources that are provided by different sources as shown by the Thing Description. For the sake of the article, the RDF of these sources is already present in the Digital Twin Knowledge Graph and the Platform Knowledge Graph.

```
<?xml version="1.0"?>
<Project xmlns="http://schemas.microsoft.com/project">
    <SaveVersion>14</SaveVersion>
    <BuildNumber>16.0.5134.1000</BuildNumber>
    <Name>rst_advanced_sample_project_MSP2010.xml</Name>
    <GUID>FE090C7C-7AEC-42FE-9452-AD0AA8643D28</GUID>
    <Title>msproj11</Title>
    <Company>Microsoft</Company>
    <Author>Frederic Bosche</Author>
    <CreationDate>2009-10-12T08:00:00</CreationDate>
    <LastSaved>2021-03-29T16:37:00</LastSaved>
    <ScheduleFromStart>1</ScheduleFromStart>
    <StartDate>2009-10-12T08:00:00</StartDate>
    <FinishDate>2010-07-02T17:00:00</FinishDate>
    <FYStartDate>1</FYStartDate>

    <Tasks>
    <Task>
       <UID>0</UID>
       <GUID>AD74C853-6199-4FE2-994B-9C5142AD8021</GUID>
       <ID>0</ID>
       <Name>msproj11</Name>
       <Active>1</Active>
       <Manual>0</Manual>
       <Type>1</Type>
       <IsNull>0</IsNull>
       <CreateDate>2009-10-12T11:25:00</CreateDate>
       <WBS>0</WBS>
       <OutlineNumber>0</OutlineNumber>
       <OutlineLevel>0</OutlineLevel>
       <Priority>500</Priority>
       <Start>2009-10-12T08:00:00</Start>
       <Finish>2010-07-02T17:00:00</Finish>
       <Duration>PT1520H0M0S</Duration>
       <ManualStart>2009-10-12T08:00:00</ManualStart>
```

Listing 2. Schedule XML Data Example.

[3]https://raw.githubusercontent.com/helio-ecosystem/helio-ecosystem/main/resources/mappings-data/COGITO/schedule.xml

For this example, the result of running the mapping for the schedule, and the mapping itself, can be found here. Once the translation process has been carried out, the result is stored in the Digital Twin Knowledge Graph and shown by Listing 3.

```
@prefix owl:<http://www.w3.org/2002/07/owl#>.
@prefix facility:<https://cogito.iot.linkeddata.es/def/facility#>
@prefix process:<https://cogito.iot.linkeddata.es/def/process#>.
@prefix xsd:<http://www.w3.org/2001/XMLSchema#>.
@prefix project:<http://data.cogito.iot.linkeddata.es/resources/project/>.
@prefix process:<http://data.cogito.iot.linkeddata.es/resources/process/>.
@prefix task:<http://data.cogito.iot.linkeddata.es/resources/task/>.
@prefix resource:<http://data.cogito.iot.linkeddata.es/resources/resource/>.

project:079637bb-87d5-4fef-9704-d556364d90ba
  a facility:Project;
  facility:isRelatedToProcess
    process:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28.
process:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28 a
  process:Process;
  process:hasCreationDate "2009-10-12T08:00:00"^^xsd:dateTime;
  process:hasTask
    task:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28_0,
    task:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28_1
    ...

task:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28_0 a
  process:Task;
  process:hasCreationDate ""^^xsd:dateTime;
  process:hasPriority 500;
  process:hasProgress "PT0H0M0S";
  process:hasStatus "PT0H0M0S";
  process:taskId "AD74C853-6199-4FE2-994B-9C5142AD8021";
  process:hasAssignedResource resource:079637bb-87d5-4fef-9704-d556364d90ba_Worker_28;
  process:hasAssignedResource resource:079637bb-87d5-4fef-9704-d556364d90ba_Truck_4;
  process:taskUid "0".

task:079637bb-87d5-4fef-9704-d556364d90ba_FE090C7C-7AEC-42FE-9452-AD0AA8643D28_1 a
  process:Task;
  process:hasCreationDate ""^^xsd:dateTime;
  process:hasPriority 500;
  process:hasProgress "PT0H0M0S";
  process:hasSubTask "1.1","1.2";
  process:taskId "64518854-E3AA-4FF7-B8BC-33DB2326571C";
  process:hasAssignedResource resource:079637bb-87d5-4fef-9704-d556364d90ba_Worker_24;
  process:hasAssignedResource resource:079637bb-87d5-4fef-9704-d556364d90ba_Truck_1;
  process:taskUid "1".
...
```

Listing 3. Schedule translation in interoperable data.

Notice that in Listing 3 the process of linking between workers (human resources) and resources are made by using the predicate *"process:hasAssignedResource"*.

Finally, to have all the information related to a project described in the Digital Twin in one place, a Thing Description is produced, as shown by Listing 4, in which the process is related to the endpoints of the previously translated information.

```
{
   "@context":[
      "https://www.w3.org/2019/wot/td/v1",
      {
          "facility":"https://cogito.iot.linkeddata.es/def/facility#"
      },
      "https://w3c.github.io/wot-discovery/context/discovery-context.jsonld"
   ],
   "security":[
       "nosec_sc"
   ],
   "title":"test1",
   "types":"https://cogito.iot.linkeddata.es/def/facility#Project",
   "registration":{
       "created":"2022-10-05T11:46:50.411Z",
       "modified":"2022-10-05T11:47:10.587Z"
   },
   "id":"079637bb-87d5-4fef-9704-d556364d90ba",
   "securityDefinitions":{
       "nosec_sc":{
          "scheme":"nosec"
       }
   },
```

```
24    "description":"Digital Twin process 079637bb-87d5-4fef-9704-d556364d90ba",
25    "properties":{
26        "hasKnowledgeGraph":{
27           "forms":[
28               {
29                   "href":"https://data.cogito.linkeddata.es/resources/079637bb-87d5-4fef-9704-d556364d90ba",
30                   "type":"text/turtle"
31               }
32           ]
33        },
34        "hasValueInGraph":{
35           "forms":[
36               {
37                   "href":"https://data.cogito.linkeddata.es/resources/project/079637bb-87d5-4fef-9704-d556364d90ba",
38                   "type":"text/turtle"
39               }
40           ]
41        }
42    },
43    "links":[
44         {
45            "href":"http://data.cogito.iot.linkeddata.es/api/things/9dbfc5ac71a0aeaa49d56b654cdb62e44b743b4abd5b6179f4979bebc1940afb",
46            "rel":"facility:hasSchedule",
47            "type":"application/td+json"
48         }
49    ]
50 }
```

Listing 4. Process Thing Description from Platform Knowledge Graph Example.

Conclusion and discussion

Digital Twins are a tool that is currently on the rise. They handle a large amount of data, that makes them sometimes unmanageable to obtain particular data. Due to this problem, three challenges have been identified. These were solved by means of semantic technologies such as translation tools to translate heterogeneous data into RDF expressed according to the ontology of COGITO, making such data semantically interoperable in the platform, or the use of Thing Descriptions with the aim of handling data scalability problems in the existing processes in the Digital Twin.

As explained earlier, a semantic interoperable based approach has been created, to handle the problem of describing and exchanging information between different systems. Also, to manage and orchestrate data between internal Digital Twins tools, the Thing Descriptions have been used, containing all the links that references to the translation tools, and the information belonging to the heterogeneous data stored at the filestorage and the semantic interoperable data stored at the triplestore.

Future work

As future work, it would be possible to create the linking of construction projects by means of Thing Descriptions by construction project type, in order to be able to acquire the information more efficiently if information pertaining to a set of processes is required.

Data availability

Zenodo: An extension of Thing Descriptions from the Web of Things for Digital Twins (Sustainable Places 2022), https://doi.org/10.5281/zenodo.7304368[6].

Software availability

Source code available from: https://github.com/oeg-upm/cogito_wrapper_module/tree/main/example

Archived source code at time of publication: https://doi.org/10.5281/zenodo.7304368[6]

License: Creative Commons Attribution 4.0 International

References

1. Katsigarakis K, Lilis GN, Rovas D, *et al.*: **A digital twin platform generating knowledge graphs for construction projects.** 2022.
Reference Source
2. Cimmino A, García-Castro R: **Helio: a framework for implementing the life cycle of knowledge graphs.** Semantic Web, 2022.
3. **Shapes constraint language (SHACL).** Technical report, W3C, 2017.
Reference Source
4. Cimmino A, Fernández-Izquierdo A, García-Castro R: **Astrea: automatic generation of shacl shapes from ontologies.** In: *European Semantic Web Conference.* Springer, 2020; **12123**: 497–513.
Publisher Full Text
5. Kaebisch S, Anicic D: **Thing description as enabler of semantic interoperability on the web of things.** In: *Proc IoT Semantic Interoperability Workshop.* 2016; 1–3.
Reference Source
6. González-Gerpe S, Cimmino A, Bernardos S, *et al.*: **An extension of Thing Descriptions from the Web of Things for Digital Twins (Sustainable Places 2022) (0.1).** [Data set]. Zenodo. 2022.
http://www.doi.org/10.5281/zenodo.7304368

Open Peer Review

Current Peer Review Status: ? ? ?

Version 1

Reviewer Report 27 June 2023

https://doi.org/10.21956/openreseurope.16522.r31880

© **2023 Ala-Laurinaho R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Riku Ala-Laurinaho
Aalto University, Espoo, Finland

The article is generally well-written and easy to understand, and the presented approach has scientific merits.

The major issue with the paper is that the literature review, including a state-of-the-art review of existing solutions, is missing. The novelty of the presented approach is not supported by analysis of the previous work. The author should also consider other possible options for describing digital twins, such as Microsoft DTDL and AAS, and justify the use WoT TD. In addition, the author's definition and view on digital twins could be added either to the introduction or literature review section.
In the experimental results section, the use case could be elaborated more: now you have shown the implementation, but how could this implementation be used in real-world use cases? Maybe also presenting the modeled information in graph format could be suitable in the presented case.

The conclusion section could be extended to highlight the relevance of the work, or I would suggest adding a separate discussion section to analyze the benefits of the presented method, compare it to the other alternatives and present practical and scientific impacts.

One minor comment: The link in "here" does not work in this sentence: "For this example, the result of running the mapping for the schedule, and the mapping itself, can be found here."

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital twins, industrial internet

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 15 June 2023

https://doi.org/10.21956/openreseurope.16522.r31877

© 2023 Picone M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Marco Picone
Universita degli Studi di Modena e Reggio Emilia, Modena, Emilia-Romagna, Italy

The paper analyzes an interesting topic related to the interoperability issues associated with Digital Twins and in particular in the context of their application to the Internet of Things ecosystem by proposing an extension to the WoT description to effectively support twins representation and effective description.

According to me, the existing open issues of the submitted manuscript are the following:

- The concept and definition of Digital Twin and how they are mapped into a WoT description is not clear. What are the fundamental characteristics and responsibilities of a DT and how they can be mapped into a WoT-TD?

- While the paper centers on the concept of DT, it lacks clarity regarding the specific model and architecture adopted for twins. The reader is left with insufficient information to understand the structure of the DT model and how it is designed. Consequently, there is a need for the authors to provide a more comprehensive explanation of the chosen DT model, including its underlying architecture in terms of relationships with both the physical and digital worlds. This would greatly enhance the reader's comprehension and allow for a more robust evaluation of the paper's contributions in the context of DTs.

- Analysis of the existing and recent scientific literature on the topic can be definitely

improved in order to properly highlight the benefits and advantages of the proposed contribution.

- Authors in the paper report the concept of DTs Knowledge Graph and of the Platform Knowledge Graph but unfortunately, the structure of these components is still unclear

- Existing enterprise-level products such as Microsoft Azure or open source approaches like Eclipse Ditto proposed solutions to model and describe DTs? Which are the differences with the proposed contribution?

As a general comment, in the present condition, the significance of the contribution remains unclear and the manuscript appears not yet solid from a scientific perspective. It is not clear how DTs have been modeled (also with respect to the state of the art), developed, and deployed in order to be effectively used by external digital applications. Consequently, it is difficult to evaluate the value of the proposed approach and the introduced benefits.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital Twins, Web of Things, Internet of Things

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 14 February 2023

https://doi.org/10.21956/openreseurope.16522.r30609

© 2023 Ricci A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Alessandro Ricci
Computer Science and Engineering Department (DISI), University of Bologna, Bologna, Italy

As the paper pointed out, digital twins are getting a momentum in providing a uniform way to mirror strategic physical assets of the physical worlds into synchronized software models and technologies that make it possible to track and act upon them, as well as predict their future behaviour. Nowadays we are moving from single digital twins to ecosystems of digital twins.

Constructions and buildings -- which is the context of the EU project COGITO -- are nowadays one of the domains where digital twins are mostly used. A main reference for this was e.g. the National Digital Twins programme that ran in recent years in the UK (https://www.cdbb.cam.ac.uk/what-we-did/national-digital-twin-programme).

In this context, the paper aims to present how some specific challenges about mirroring a physical layer which is characterised by strong heterogeneity have been tackled, by means of adopting semantic technologies (RDF) to build proper uniform knowledge graphs.

Generally speaking, this is an important issue to deal with when building digital twins. Nevertheless, in the current state, it is not fully clear what the relevance of the contribution is from a research point of view. By reading the paper, generally speaking, it is not easy to get a clear idea about what are -- at the end of the day -- the digital twins that are developed, how do they work and how they are exploited by applications on top. Given this, it is not easy to assess the value/complexity of the specific aspect that is focused in the paper, about dealing with data heterogeneity at the physical layer. So the paper would deserve a strong revision making it more clear all these things, both the basic picture and the reason why the approach adopted can be considered relevant from a research of view (compared to existing state of the art and related work).

Some more detailed comments follow:

- Even if the paper is focused on "digital twins" (DT), the specific model adopted for DTs - and their architecture as well - appear unclear. The paper talks about the "Digital Twins Knowledge Graph" and of the "Platform Knowledge Graph", but it is not fully clear what the structure of these KGs is. So in order to fully appreciate the contribution, it would be important to give, briefly, a more clear and rigorous idea of what are the digital twins considered in the project, what is the meta-model adopted to describe them, what is the architecture of the platform adopted to run them and what's the structure of the corresponding KGs adopted to describe them.

- Related to the previous point, it is not clear what the complexity is of the physical assets that are considered in the project and of the corresponding digital twins. The current description in the introduction and proposal sections appear very high-level and somewhat superficial. A more clear and rigorous description of these aspects -- eventually including also some concrete examples - would be important to assess the value from a research point of view

of the challenges and solutions proposed in the paper.

- The problem described in the paper -- about interoperability, data heterogeneity from the physical layer, etc. -- is not new indeed, as well as solutions based on the adoption of semantic technologies. For instance, interoperability is a main issue tackled by e.g. the Web of Things (WoT) initiative (which is used as a main reference in the paper) and Thing Description in WoT supports semantic annotations. Commercial technologies such as Azure DT already supports the definition and implementation of digital twins of heterogenous physical assets, involving heterogeneous data, and DTs can be represented in terms of knowledge graphs based on semantic technologies. Accordingly, it would be important to remark what makes the challenge and solution discussed in the paper relevant compared to the state of the art.

- As a follow-up of the previous point, the paper is missing a related work section. This is critical: it is not clear how authors situate their work in the existing research context about digital twins and -- orthogonally -- about the key challenges that are considered, about dealing with data heterogeneity and the use of semantic technologies to this purpose. There literature about all this is quite large.

- In the specific case studies considered in "Experimental Results" it is not fully clear how they related to "digital twins".

- It talks about "schedules", that contain a set of tasks that must be carried out in a construction project, that are meant to be assigned to workers and associated with existing physical resources. The section explains how schedules are handled in the DT Platform, in using translators to deal with heterogeneous data formats to get a uniform data representation.
- It is not clear why a DT platform is managing these schedules, which is typically the role of other software systems (workflow management systems, business process management systems, etc) and why the problem is relevant from a Digital Twin perspective.

- The clarification mentioned before (about what are digital twins and the corresponding physical assets, how they are synchronised, etc.) may help to better understand then also these results.

Typos:
- In the paper there are many words broken (e.g. "trans-lation", "know-ing",..).

Is the work clearly and accurately presented and does it cite the current literature?
No

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital twins (modelling, engineering); Web of Things; Intelligent Agents and Multi-Agent Systems;

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.