Article

SCRO: A domain ontology for describing steel cold rolling processes towards Industry 4.0

Sadeer Beden*, **Qiushi Cao and Arnold Beckmann**

Department of Computer Science, Swansea University, United Kingdom
* Correspondence: sadeer.beden@swansea.ac.uk

Abstract: This paper introduces the Steel Cold Rolling Ontology (SCRO) to model and capture domain knowledge of cold rolling processes and activities within a steel plant. A case study is set up that uses real-world cold rolling data sets to validate the performance and functionality of SCRO. This includes using the Ontop framework to deploy virtual knowledge graphs for data access, data integration, data querying, and condition-based maintenance purposes. SCRO is evaluated using OOPS!, the ontology pitfall detection system, as well as feedback from domain experts from Tata Steel.

Keywords: Industry 4.0; Steelmaking; Cold rolling; Ontology; Ontop

1. Introduction

The fourth industrial revolution, also known as Industry 4.0, is full of new concepts, technologies, and innovations with the goal to optimise, digitize, and autonomize industrial processes [1]. It is a vision where machines, products, and processes are connected intelligently and are able to derive meaning from data to make autonomous decisions.

Presently, large industrial machines follow rigid automation protocols which generate vasts amount of data. This data is often not machine-understandable, and stored in data silos that are often not interconnected yet contain data that is semantically related [2]. A fundamental task to enable Industry 4.0 is to enrich data with semantics to make the data interoperable and machine-understandable. The steel industry is one of many manufacturing domains that are working towards this goal [3–5].

Meanwhile, ontologies have become a prominent methodology for knowledge modelling and capturing domain knowledge, as well as addressing and improving data semantics in various domains. By developing an ontology, we are in essence building a knowledge base within a specific domain [6,7]. In the domain of smart manufacturing, ontologies can play a key role as they are able to provide machine-understandable vocabularies and data exchange between different individuals and processes. Ontologies provide additional functionalities such as stream reasoning which infer new knowledge, and ontology-based data access which allows data to be queried without being physically integrated.

Cold rolling is one of many different steel-making processes within a steel factory. Rolling in general processes the greatest tonnage of metals than any other metal working technique [8]. The purpose of cold rolling is to compress steel to produce steel coils. During the cold rolling process, the material undergoes deformation, and is compressed by a pair of rolls that rotate in opposite directions under a heavy force. There is a gap between the two rolls that is smaller than the material, thus forcing the material to decrease in size as it passes through the rolls.

Due to strong forces being involved, these rolls are affected by roll wear where the roll service life and the quality of the product are significantly impacted [9]. To avoid this, the rolls are refurbished regularly, where the diameter of the rolls are marginally reduced

Citation: Beden, S.; Cao, Q.; Beckmann, A. Title. *Information* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

to remove the worn surface. One long-term aim of our research is to use the semantically interoperable data to optimise the life of the rolls, improving their total tonnage and yield. In addition, accidents and anomalies that occur, such as overloading, spalling, and incorrect grinding operation [10], can be avoided preemptively once achieving better semantic interoperability.

The goal of this paper is to develop an ontology that focuses on modelling the cold rolling processes that occur during steel-making. Thereby, this paper introduces the Steel Cold Rolling Ontology (SCRO) that acts as a knowledge base for cold rolling processes within a steel manufacturing plant. This includes the relevant systems, facilities, hardware, software, and inventory of a cold rolling mill. To validate and evaluate the usefulness and accuracy of the SCRO ontology, we perform a case study that aligns the ontology with real-world data sets of a cold rolling mill provided by Tata Steel Europe[1]. In this case study, we exploit Virtual Knowledge Graphs (VKG) to access and query the data sets to obtain valuable knowledge.

The remainder of the paper is structured as follows. In Section 2, we provide a literature review that focuses on two key topics: ontologies for Industry 4.0, and ontologies for the steel industry. We also introduce our selected design methodology of ontology development. In Section 3, we describe in detail the SCRO Ontology, including its classes and main concepts. In Section 4, we demonstrate the usefulness of the ontology on an application that uses real-world data. In Section 5, we discuss the validation of the SCRO Ontology to ensure that the knowledge is accurate. Finally, we reflect over our work and end with a conclusion and future work in Section 6.

2. Literature Review

The W3C have developed a formal ontology language named The Web Ontology Language (OWL)[2] to model concepts and relations within ontologies. OWL is a component of SemanticWeb that allows for explicit representations of the meaning of terms in vocabularies and the relationships between those terms. These representations and their interrelations form an ontology. In the following subsections, we review relevant existing OWL ontologies and their rule-based extensions.

2.1. Ontologies for Industry 4.0

There have been numerous ontologies developed in recent years to tackle and achieve aspects of Industry 4.0. The Reference Architecture Model for Industry 4.0 (RAMI 4.0) [11], a model that highlights the fundamental requirements for achieving Industry 4.0, has introduced the fundamental concept of an Asset Administration Shell (AAS) as a way for storing and communicating data between assets. A core requirement to enable the AAS concept is to enhance assets with rich data semantics and make them interoperable. As a result, one research direction shifted towards ontology development to capture domain knowledge and concepts to achieve this goal. In our review, we structured the literature into three categories: *product-related* concepts, *process-related* concepts, and *resource-related* concepts.

Firstly, when looking at product-related concepts, Vegetti et al. [12] developed the PRoduct ONTOlogy (PRONTO) to model *Complex Products* which consider different abstraction levels of product concepts such as *Family* and *Variant*. This approach has benefits and drawbacks. One benefit is that it extends conventional product structure representations, and considers composition and decomposition structures of products from a wide range of different manufacturing environments. One drawback is that there is a lack of capability to refer to existing international standards related to the modelling of product structure, processes, and features. Further research in this direction has been lead by Panetto et al. [13] as they developed the ontological model ONTO-PDM

[1] https://www.tatasteeleurope.com/ts/
[2] https://www.w3.org/TR/owl2-overview/

which overcomes these shortcomings. This ontology uses the knowledge related to the product technical data to formalize heterogeneous information that is scattered across different organizations [13]. ONTO-PDM also incorporates different standardisation initiatives, including the International Electrotechnical Commission (IEC) standards and International Organization for Standardization (ISO) standards. Another example of product-related concept modelling includes the MASON ontology, developed by Lemaignan et al [14] to create a common semantic net for Industry 4.0. It models three core concepts: *Entities*, *Operations*, and *Resources*, and specifies the product information as *Geometric Entities*, *Raw Material*, and *Cost Entities*. Using the proposed semantic net, they accurately link the product-related concepts with the description of manufacturing process and resources.

Secondly, some ontologies focus on *resource-related* concepts within Industry 4.0. *Resources* in this context are defined as the physical objects within an Industry 4.0 environment that are capable of executing a range of different operations. The MASON ontology mentioned above also studies the notion of *Resources* and deconstructs it into four sub-notions: *Machine-tools, Tools, Human Resource, and Geographical Resources*. The modelling of resources enables estimations of total costs for certain manufacturing activities. Additionally, Borgo and Leitão defines *Resource* as "an entity that can execute a certain range of jobs, when it is available, as long as its capacity is not exceeded" in [15]. The authors used the Java Agent Development Framework (JADE) framework to implement their ontology as a part of a multi-agent control system, and concluded that an ontology is a core requirement in handling heterogeneous data generated by manufacturing control applications.

Finally, some ontologies address *process-related* concepts within Industry 4.0. These processes are generally a linear sequence of activities in which raw materials undergo some treatment such as assembly and integration before converting into the final product. The Process Specification Language (PSL) Ontology [16] was developed by Grüninger et al. to facilitate different methods of exchanging process information between manufacturing systems. Using PSL and first-order logic theories, the authors formalize the concept of a *process*. This formalisation has been widely adopted in many different domain applications such as process modelling and process monitoring [16]. Another ontology that focuses on process-related concepts was developed by Cao et al. [17] which formalises essential concepts and relationships related to condition monitoring. Their ontology contains three sub-modules: *Manufacturing*, *Context*, *Condition Monitoring* which is used within a Cyber Physical System to enable a case study to model real-time predictive maintenance. The same authors developed a new ontology named *Manufacturing Predictive Maintenance Ontology (MPMO)* in [18] which uses Semantic Web rule Language (SWRL) rules to enable ontology reasoning. Using a real-world data set, this ontology is able to detect and predict possible anomalies within an Industry 4.0 manufacturing process.

2.2. Ontologies for the Steel Industry

In the steel industry, ontologies are used as an effective and intelligent knowledge management tool for conceptual modelling and information integration. Leveraging the strong modelling and reasoning capabilities of ontologies, process knowledge regarding steelmaking is structured and inferred to facilitate decision making.

Developed as a core component of a Big Data Knowledge Management System (BDAKMS), the ontology introduced in [19] is used to model domain knowledge of steelmaking and enhance the usability and interoperability of BDAKMS. The developed ontology is further used together with SWRL [20] rules to infer knowledge regarding the demand of raw materials. In [21], a shared global supply chain ontology is designed to manage the heterogeneous internal and external decision knowledge of steel companies. Similar to the previous literature, semantic rules are also used to perform ontology reasoning. The goal of ontology reasoning is to facilitate the decision making of business strategies of steel companies. In this way, senior managers can use the ontology to retrieve useful implicit decision knowledge such as pricing strategies, partner selection strategies, and product development strategies.

Ontologies are also used for planning and scheduling of steel production. In [22], an ontological approach is proposed for the goal of optimal planning and scheduling. Within the proposed approach, a set of ontologies are integrated to form an ontological framework. A core meta-ontology and different domain specific ontologies for primary steelmaking are integrated with ANSI/ISA-S95 standard to construct the main body of the framework. Another ontology is introduced in [23] to help with the conceptual design of steel structures. During the ontology design phase, required knowledge elements are identified using intelligent agents. The proposed ontology is reused in other projects such as Agent-Based Collaborative Design of Light Industrial Buildings (ADLIB) and Automated Agent Learning (AAL).

2.3. Ontology Development Methodology

Using an extension of the *eXtreme Design methodology* [24], The SCRO ontology is designed using *Ontology Designing Patterns* (ODPs) [25]. We conclude that this design approach offered numerous evident advantages for developing ontologies, including: a faster ontology design process, more flexible design choices, improved interoperability and ontology quality [26].

3. SCRO: Steel Cold Rolling Ontology

Most of the domain knowledge mentioned in this section was obtained from a case study with Tata Steel, at the Cold Rolling Mill in the Port Talbot plant. SCRO models the fundamental structure and operations of the rolling processes in the case study. Although SCRO is initially designed for the processes and machines at Tata Steel, it could potentially be reused by other steel manufacturers for knowledge modelling. In this section, we describe the SCRO ontology in detail, beginning with the encoding and classes.

3.1. Coding

SCRO was developed using the free, open-source ontology editor and framework called Protégé [27]. We used the latest version to date, Protégé 5.5.0, that offers a unique interfaces for creating and maintaining ontologies for intelligent systems. Protégé supports the commonly used ontology language, OWL, which enables us to model concepts, as well as their relations and attributes through classes, object properties, and data properties [28]. Figure 1 displays the structure and the architecture of SCRO, whereas Figure 2 displays the classes, object properties, and data properties.

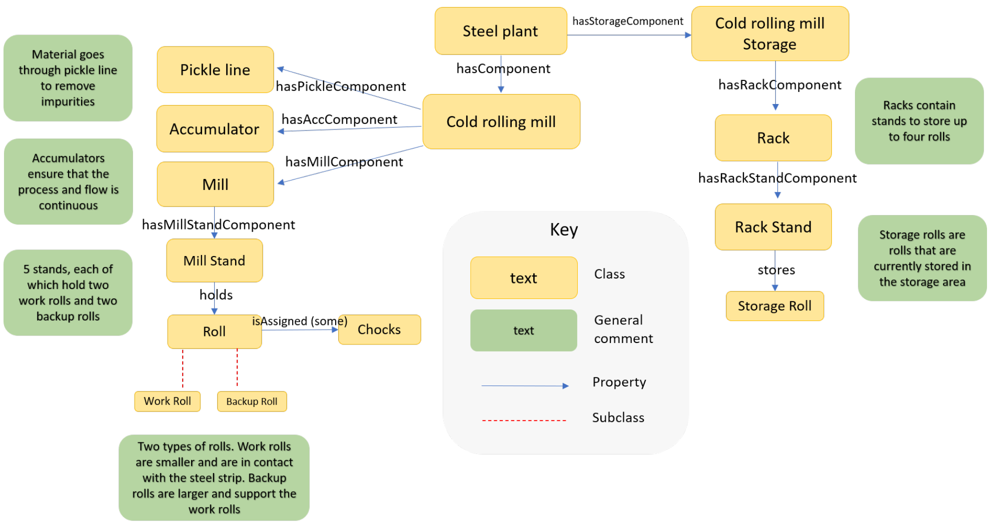

Figure 1. Structure of the SCRO Ontology

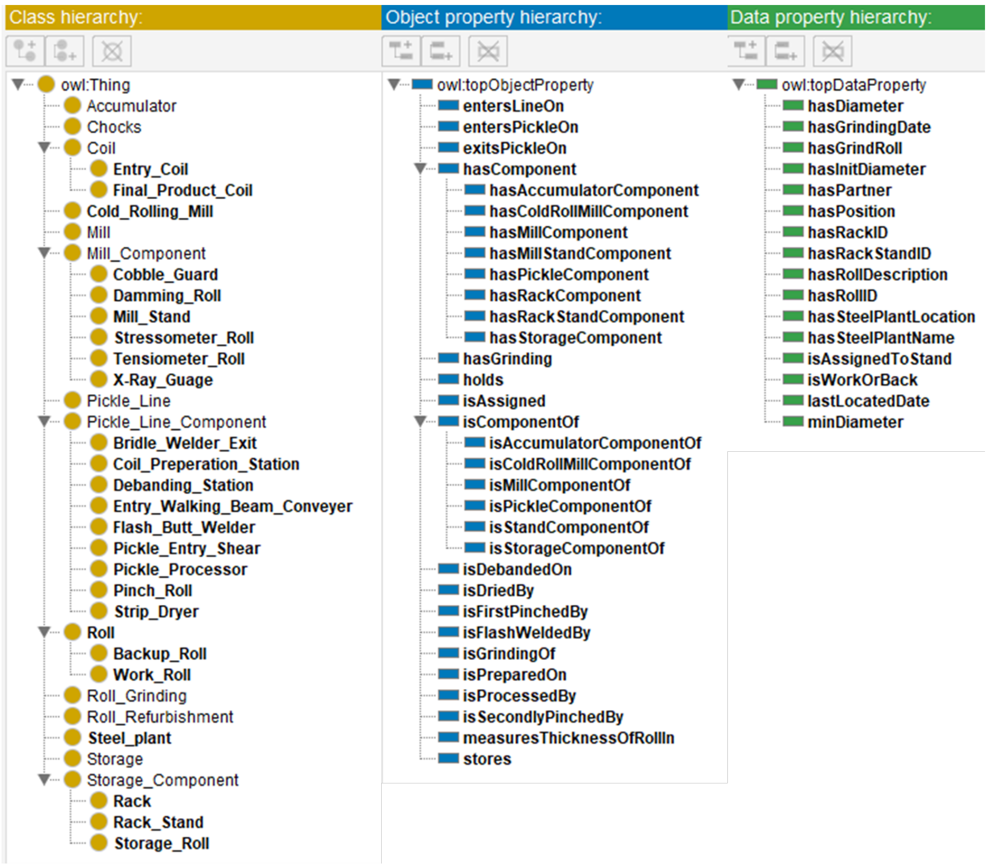

Figure 2. Classes, object properties and data properties of the SCRO Ontology

3.2. *Reusing Existing Ontologies*

An extensive amount of data within the domain of steel manufacturing is generated and read through sensors. Generally, these sensors run on timestamp data to record the continuous flow of dynamic data. Therefore, we have imported the *Time* ontology created by W3C that supports the use of timestamp data [29]. These are excluded from Figure 2 but play an important role in SCRO.

3.3. *Classes*

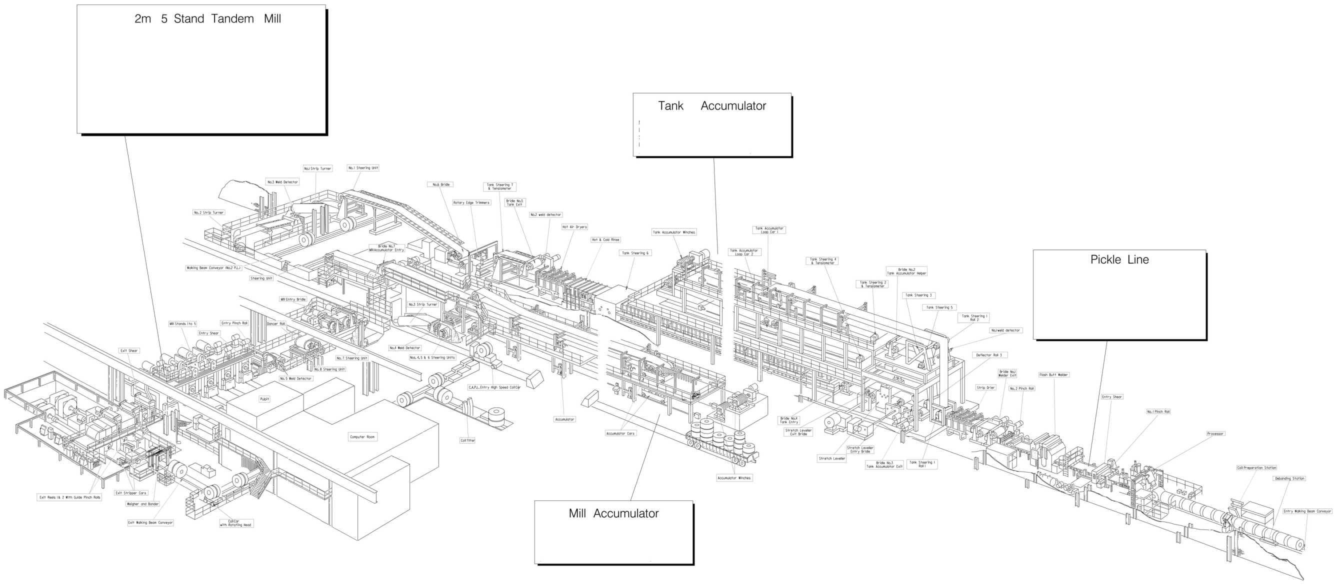

Figure 3. The big picture of the cold rolling processes at the Port Talbot plant provided by Tata Steel. Copyright © 2021 All Rights Reserved

There are many processes and components on the shop floor that are fundamental for cold rolling as depicted in Figure 3. We create classes for each one respectively. The cold rolling mill processes are divided into three sub-processes: the pickle line, accumulators, and the mill.

Firstly, the process of steelmaking creates undesirable oxidations on the material. To counter this, the material, *entry coil*, undergoes surface treatment on the pickle line. The process of Pickling cleanses the entry coil by using acid to eliminate impurities and oxidations, providing a smoother surface. The class ***:Pickle_Line*** denotes this process whereas the superclass ***Pickle_Line_Component*** contains the necessary pickle line components on the shop floor as subclasses. Table 1 defines these components.

Both the pickling and mill processes are continuous and run at different speeds. Often one of these processes is required to stop while the other is still in operation. For example, when introducing a new coil into the pickling process, the pickle line is paused to weld/stitch the new coil while the mill process is still running at a constant speed. An *Accumulator* between these two processes is able to facilitate such activities through movable rolls that are able to control the amount of material in that intermediate section, ensuring the whole cold rolling process to be continuous from beginning to end. The class ***:Accumulator*** denotes this process.

Finally, the material is passed through the mills where its thickness is reduced. The class ***:Mill*** denotes this process whereas the superclass :***Mill Component*** contains the necessary mill components on the shop floor as subclasses; these components are also defined in Table 1.

The rolls are fundamental components of the cold rolling process. The rolls are the physical entities that rotate to reduce the thickness of the steel trip. These are denoted by the superclass ***:Roll*** and its two nodes ***:Work Roll*** and ***:Backup Roll***. These rolls are assigned some chocks which allow for rotation within a mill; these chocks are denoted as ***:Chocks*** in the ontology. In addition, we have included ***:Storage Roll*** which are rolls that are out of the mill and are in the storage area. This storage area is denoted by the class ***:Storage***, and the superclass ***:Storage Component*** contains the components of the storage as subclasses.

Finally, the ontology contains other classes such as ***:Steel Plant***, ***:Cold_Rolling_Mill***, ***:Roll Refurbishment*** and ***Roll Grinding*** which are briefly described in Table 1. Figure 4 displays the hierarchy of all the classes, generated by the protégé tool.

Table 1. Description of SCRO classes.

SCRO Classes	**Description**
Accumulator	Manage the speed of the rolling processes to ensure flow is continuous
Chocks	Attached to rolls. Chocks contain bearings that allow rolls to rotate
Coil	**Superclass of the material and final product**
Entry_Coil	Denotes the steel strip that enters the cold rolling mill
Final_Product_Coil	The final product sold to customers
Cold_Rolling_Mill	Denotes the shop floor of the cold rolling mill
Mill	Process of the cold rolling mill where thickness of the steel strip is reduced
Mill_Component	**Superclass of all Mill component**
Cobble_Guard	Component that reduces chance of producing cobbles
Damming_Roll	Component that restrains the outward flow of coolants
Mill_Stand	Stand that fits two work rolls and two backup rolls
Stressometer_Roll	Measures the flatness of the steel strip
Tensiometer_Roll	Measures the tension of the steel strip
X-Ray_Guage	Measures the thickness of the steel strip
Pickle_Line	Process where the entry coil undergoes surface pickling
Pickle_Line_Component	**Superclass of all Pickle component**
Bridle_Welder_Exit	Mill exit equipment that the strip uses to exit the pickling process
Coil_Preparation_Station	Station where the entry coils are entered
Debanding_Station	Station where the entry coils are debanded
Entry_Walking_Beam_Conveyor	Conveyor where entry coils are first placed
Flash_Butt_Welder	Machine that presses together and welds the ends of the workpiece
Pickle_Entry_Shear	Machine that cuts rolls to desired size
Pickle_Processor	Processes the coil and minimizes the tendency for coils to break
Pinch_Roll	Machine that holds and moves the strip
Strip_Dryer	Removes excess water from the strip to prevent rusting
Roll	**Superclass of the two types of rolls at a cold rolling mill**
Backup_Roll	Larger roll that support a work roll during milling
Work_Roll	Smaller roll that rotates to reduced thickness of steel during milling
Roll_Grinding	Contain previous grinding data of rolls
Roll_Refurbishment	Process where rolls are sent to be refurbished
Steel_Plant	Denotes the whole steel plant
Storage	Section of the cold rolling mill where assets (e.g unused rolls) are stored
Storage_Component	**Superclass of the Storage component**
Rack	Ccontain stands for rolls to be stored
Rack_Stand	Store one storage roll
Storage_Roll	A roll that is not currently being used and stored away

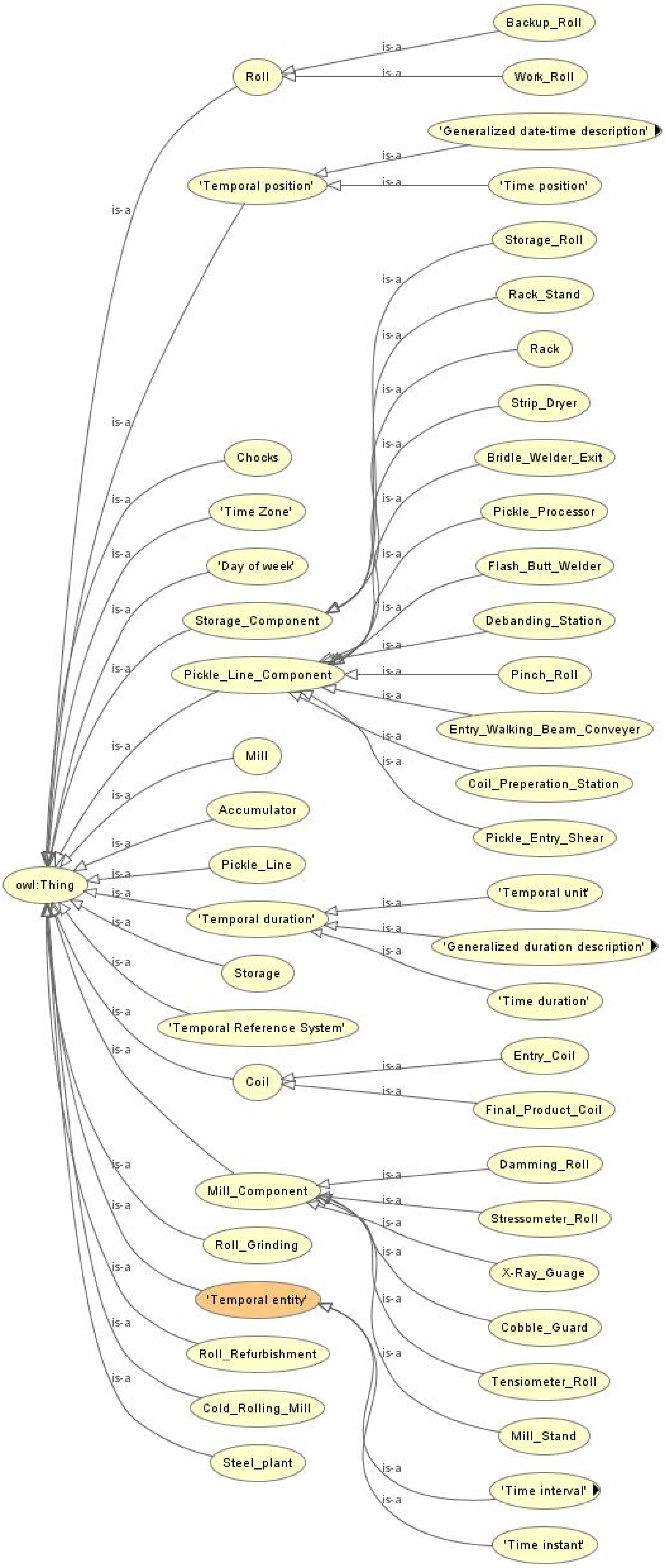

Figure 4. Hierarchy of all the classes in the SCRO Ontology.

3.4. Object and Data Properties

To semantically describe the relations between classes, it is important that we specify the domain and ranges of the properties. These properties are clarified below:

- entersLineOn(object1, object2) where object1 is an *Entry_Coil* and object2 is an *Entry_Walking_Beam.*
- entersPickleOn(object1, object2) where object1 is an *Entry_Coil* and object2 is a *Pickle_Entry_Shear.*
- exitsPickleOn(object1, object2) where object1 is an *Entry_Coil* and object2 is a *Bridle_Welder_Exit.*
- hasComponent(object1, object2) where object1 and object2 are left undefined as this is the superclass for all hasComponents mentioned below.
- hasAccumaltorComponent(object1, object2) where object1 is a *Cold_Rolling_Mill* and object2 is an *Accumulator.*
- hasColdRollMillComponent(object1, object2) where object1 is a *Steel_plant* and object2 is a *Cold_Rolling_Mill.*
- hasMillComponent(object1, object2) where object1 is a *Cold_Rolling_Mill* and object2 is a *Mill.*
- hasMillStandComponent(object1, object2) where object1 is a *Mill* and object2 is a *Mill_Stand.*
- hasPickleComponent(object1, object2) where object1 is a *Cold_Rolling_Mill* and object2 is a *Pickle_Line.*
- hasRackComponent(object1, object2) where object1 is a *Storage* and object2 is a *Rack.*
- hasRackStandComponent(object1, object2) where object1 is a *Rack* and object2 is a *Rack_Stand.*
- hasStorageComponent(object1, object2) where object1 is a *Steel_Plant* and object2 is a *Storage.*
- hasGrinding(object1, object2) where object1 is a *Roll* and object2 is a *Roll_Grinding.*
- holds(object1, object2) where object1 is a *Mill_Stand* and object2 is a *Storage_Roll.*
- isAssigned(object1, object2) where object1 is a *Roll* and object2 are *Chocks.*
- The superclass *isComponentOf* which is the inverse of *hasComponent*, as well as all of its subclasses.
- isDebandedOn(object1, object2) where object1 is an *Entry_Coil* and object2 is a *Debanding_station.*
- isDriedBy(object1, object2) where object1 is a *Entry_Coil* and object2 is a *Strip_Dryer.*
- isFrstPinchedBy(object1, object2) where object1 is a *Entry_Coil* and object2 is a *Pinch_Roll.*
- isFlashWeldedBy(object1, object2) where object1 is a *Entry_Coil* and object2 is a *Flash_Butt_Welder.*
- isPreparedOn(object1, object2) where object1 is a *Entry_Coil* and object2 is a *Coil_Preperation_Station.*
- isProcessedBy(object1, object2) where object1 is a *Entry_Coil* and object2 is a *Pickle_Processor.*
- isGrindingOf(object1, object2) where object1 is a *Roll_Grinding* and object2 is a *Roll.*
- MeasuresThicknessOfRollIn (object1, object2) where object1 is a *X-Ray_Guage* and object2 is a *Mill_Stand.*
- stores(object1, object2) where object1 is a *Rack_Stand* and object2 is a *Storage_Roll.*

Similarly with the data proprieties in the ontology:

- hasDiameter(object, datatype) where object is Roll and datatype is xsd:double.
- hasGrindingDate(object, datatype) where object is *Time instant* and datatype is xsd:date.
- hasGrindRoll(object, datatype) where object is *Roll_Grinding* and datatype is xsd:integer.
- hasInitDiameter(object, datatype) where object is *Roll* and datatype is xsd:double.
- hasPartner(object, datatype) where object is *Roll* and datatype is xsd:integer.

- hasPosition(object, datatype) where object is *Roll* and datatype is xsd:string.
- hasRackID(object, datatype) where object is *Rack* and datatype is xsd:integer.
- hasStackStandID(object, datatype) where object is *Rack_Stand* and datatype is xsd:integer.
- hasRollDescription(object, datatype) where object is *Storage_Roll* and datatype is xsd:String.
- hasRollID(object, datatype) where object is *Roll* and datatype is xsd:integer.
- hasSteelPlantLocation(object, datatype) where object is *Steel_Plant* and datatype is xsd:String.
- hasSteelPlantName(object, datatype) where object is *Steel_Plant* and datatype is xsd:String.
- isAssignedToStand(object, datatype) where object is *Roll* and datatype is xsd:integer.
- isWorkOrBack(object, datatype) where object is *Roll* and datatype is xsd:string.
- lastLocatedDate(object, datatype) where object is *Time instant* and datatype is xsd:dateTime.
- minDiameter(object, datatype) where object is *Roll* and datatype is xsd:double.

4. Application

4.1. Data set

We test and evaluate SCRO through a real-world industrial application. Within this industrial application, a collection of real-world data sets have been provided by Tata Steel. These data sets come specifically from their five stand Tandem Cold Rolling Mill at their Port Talbot plant.

Firstly, static data related to the rolls, roll storage, and roll refurbishment have been collected. These data sets are stored in a database where the values of these rolls are always updated manually from someone at the plant. This data is considerable in quantity and located in different tables within their database. For our research, we focused on three specific tables: the *Roll*, *Roll Grinding* and *Roll Storage* tables. These tables contain many fields of data that we have chosen not to include in SCRO. Instead, we only include the fields we acknowledged as the core fields such as *RollID* and *diameter* but not *SupplierID*. The domain experts from Tata Steel agreed with this approach. Table 2 describes the tables in the database, including the fields, data types and descriptions.

Secondly, the data sets also contain dynamic data from the cold rolling mills that are read through sensors and stored in a database. These sensors record the condition of rolls in short intervals, thus, creating huge amounts of industrial data. The data includes the chemistry of the rolls, temperature, pressure, and much more.

Table 2. Description of all three tables from the data sets.

Table and fields	Data type	Description
Rolls	**Table**	**Contains static data relevant to the Rolls**
Roll_ID	Integer	Unique identifier of the roll. Primary Key
Diameter	Double	Stores the value of the diameter of the roll.
Position	String	*Top* or *Bottom* to denote their position in mill
Partner_ID	Integer	Unique identifier of the roll's partner
Work_Backup	String	Identifier to specify whether a roll is a work or backup roll
Last_Loc_Date_Time	Date	Timestamp of the date when the roll was last located
Last_Stand_ID	Integer	The last stand this roll was placed in
Roll_Grinding	**Table**	**Table that stores the previous grindings of each roll**
Roll_ID	Integer	Non-unique identifier to specify which roll
Diameter	Double	Stores the value of the diameter of the roll
Grind_date	Date	Timestamp of the date when that roll was grinded
Stand_ID	Integer	The last stand this roll was placed in
Roll_Storage	**Table**	**Table that stores the data of rolls that are currently not in use**
Rack_Location	Integer	Non-unique identifier of the location of the racks
Single_Rack_ID	Integer	Unique identifier of the rack
Roll_ID	Integer	Unique identifier of the roll that is stored on a rack
Status_description	String	The status of the roll, i.e. it's a new roll or damaged roll
Actual_Diameter	Double	Stores the value of the diameter of the roll

Note: these tables are not interconnected but contain fields that are semantically related. For example, *Roll_ID* appears in all three tables. To effectively use the data, integration is required. However, it can be costly to join, clean, and homogenize the data. To avoid this, in recent years, VKG have been developed as a paradigm for data integration and access by exploiting data virtualization [2]. This is achieved by creating graphs on top of relational databases where the data is not physically moved to another database and instead kept and viewed at a virtual level [30]. virtualization is achieved by creating an ontology, and linking the data sources to the ontology via *Mappings*. These mappings enable the ability to query data at a virtual level without paying the cost of integration. Numerous applications have been developed to support the VKG approach. Some examples include Mastro [31], Morph [32], and Ontop [30]. For our approach, we have adopted the Ontop framework.

4.2. Ontop Framework

The Ontop Framework[3] is an open-source VKG (previously known as Ontology-based Data Access) framework developed by the Free University of Bozen-Bolzano. We have chosen Ontop over the other VKG approaches as Ontop supports all the W3C languages and recommendations including RDF, OWL, SPARQL, R2RML, and SWRL [33]. Additionally, it supports widely used standards including: (1) Ontologies: Ontop supports *OWL 2 QL* ontology language which runs on description logics; (2) Mappings: Ontop supports its own *Ontop Mapping Language* as well as the W3C recommendation *R2RML mapping language*; (3) Data Source: Ontop supports the major commercial and free structed databases such as *MySQL, H2*, and *PostgreSQL*; (4) Querying: Ontop supports the latest version of the *SPARQL* querying language, which includes many features such as aggregation and negation [34].

4.3. Mappings

Mappings are created to link ontology classes and properties with data from the relational data sources to produce RDF triples. R2RML is the standard mapping language used in the semantic web [35]. For our mappings, as mentioned above, we used the Ontop mapping language which is fully interoperable with R2RML [33].

Mapping engineering is considered a difficult and time-consuming activity that requires strong knowledge of not only the domain of interest, but also the rigid structure of databases and their schemas. Presently, there are several contributions working towards this direction to automate the process. There are two main approaches to mapping engineering. The first is using Mapping Bootstrappers (MB) which automatically generate a mapping for a data source [2]. These mappings follow a set of rules based on the W3C Direct Mapping specification to generate RDF graphs [36]. Ontop bootstrapper and BootOX [37] are two examples of existing MBs. A benchmark suite named Relational-to-Ontology Data Integration (RODI) [38] has been developed to evaluate and compare MBs. Using an MB has both benefits and drawbacks. The key benefit is that it is fast and automatic; whereas the biggest drawback is that it lacks flexibility when having numerous data sources as the generated vocabulary becomes restricted to data-source specific data. The second approach is to use mapping editors to manually write mappings. For our approach, we manually wrote our mappings using a text editor that is available in the Protégé IDE.

[3] https://ontop-vkg.org//

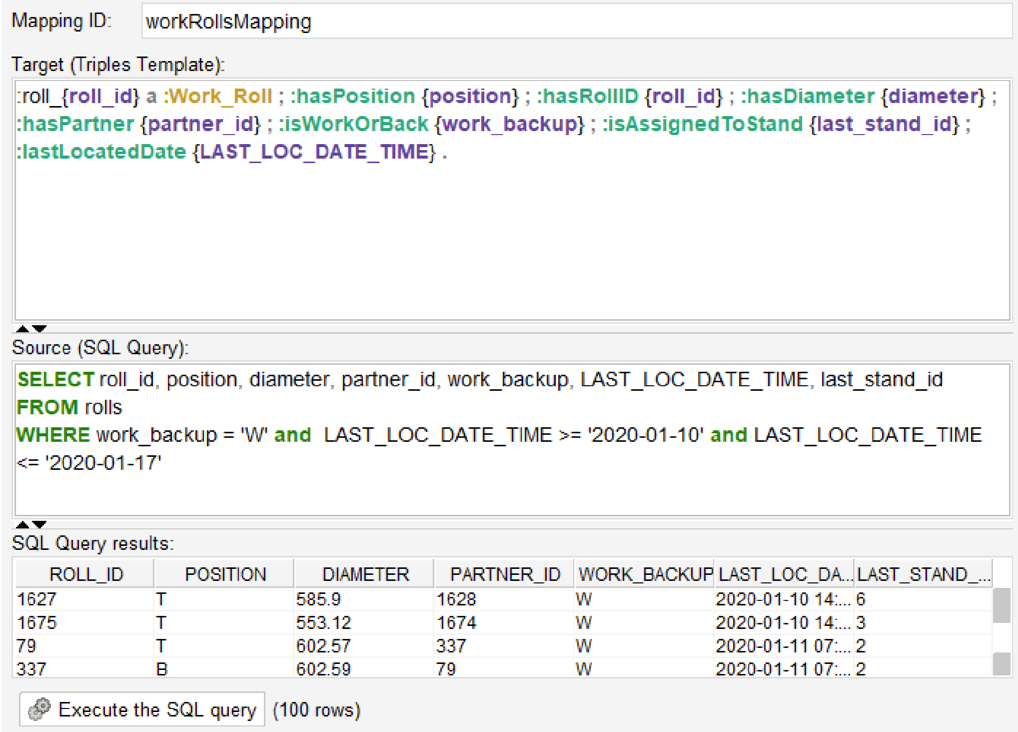

Figure 5. Ontop Mapping for work rolls.

Figure 5 shows a mapping between the *Work_Roll* class in SCRO and the *Rolls* table in the SQL database. The bottom half of the figure illustrates the source, in the form of an SQL query that allows us to specify and filter the data we want to map. Like with all SQL queries, we use the *SELECT* clause to select the necessary fields from the database, followed by the *FROM* clause to select the table name. Finally, we use the *WHERE* clause to refine the query. As seen in Figure 5, we are interested in the *roll_id, position, diameter, partner_id, work_backup, last_loc_date_time, and last_stand_id* values from the *rolls* table where the *work_backup* field is *'W'* which denotes work rolls. We use the *AND* clause to further refine the query to restrict the *last_loc_date_time* timestamp value to a seven day period. We can then click the "Execute the SQL query" provided by the Ontop Mappings plugin in Protégé to print and verify the results of the query. To conclude, the SQL query returns all work rolls that were last located between the 10^{th}-17^{th} of January 2020.

Secondly, we create a mapping *target* which maps the selected fields from the database onto the classes in the SCRO ontology. The target section is written using Turtle-like syntax[4]. The first part ***:roll_{roll_id}*** is a variable name of the individual, and the subject of the RDF triples being generated. Here, we used the primary key *roll_id* from the SQL query to create a unique IRI for each individual roll. For example, the roll with *roll_id* of 500 in the database will be named *roll_500*. The second part **a :Work_Roll** specifies that this individual and RDF triple will be an instance of the *Work_Roll class*, followed by a semi-colon. Note, by using a semi-colon instead of a fullstop, Ontop is able to map numerous fields from the SQL query to the data properties in the ontology without having to specify the initial subject and class each time. The syntax for these mappings are shown in Figure 5. For example, **:hasPosition {position}** implies *:hasPosition* is a data property from the ontology where the value of this property is mapped to the *{position}* field from the SQL source.

Similarly, we have a comparable mapping for the backup rolls. The key difference is the **:roll_{roll_id} a :Work_roll** becomes **:roll_{roll_id} a :Backup_roll** and the *work_backup* field in the SQL *WHERE* clause is set to equal *'B'*.

4 https://www.w3.org/TR/turtle/

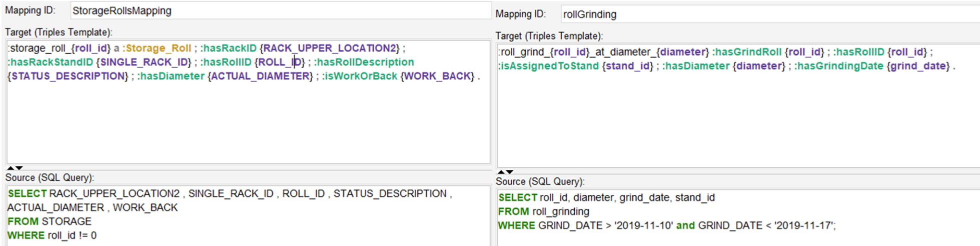

Figure 6. Ontop Mapping for Grindings and Storage Rolls.

Figure 6 depicts two other mappings. The mapping on the left manages and links the SCRO ontology with the *roll_storage* data set, whereas the mapping on the right manages historical grinding values of rolls from the *roll_grinding* data set.

4.4. SPARQL

We use SPARQL[5] to query the data for condition-based maintenance of rolls and information retrieval purposes. SPARQL is a well known querying language within the semantic web. The difference between SPARQL and SQL is that SQL queries on structured databases, whereas SPARQL queries on RDF triples [35]. As described above, the RDF triples are generated by the Ontop mappings that are depicted in Figure 5 and Figure 6, which enable us to query the data with SPARQL.

There are applications being developed to aid the assistance of SPARQL query formulation. An example includes the OptiqueVQS tool [39], which provides an interactive interface that generates components to build SPARQL queries. However, we decided to write our SPARQL queries manually using a text editor provided by the Protégé software. Below are some queries that we developed to query the data.

Listing 1: Diameter values which appear in more than two rolls.

```
PREFIX : http://www.semanticweb.org/sadee/ontologies/2021/1/SCRO#
PREFIX time: http://www.w3.org/2006/time#

SELECT ?diameter
WHERE {
    ?roll :hasDiameter ?diameter .
    MINUS {
    ?roll :hasGrindRoll ?grind .
    }
}
group by ?diameter
having (count(?diameter) > 2)
```

Listing 1 is a query that outputs the diameter values that have three or more rolls that share that diameter. Rolls in operation are always paired with other rolls that have the same diameter value, thus, each diameter should appear twice in the rolls data set. In contrast, rolls from the storage data set have yet to be paired. By limiting our search to only return diameter values that appear three or more times, this type of query can be used to discover rolls that have matching diameter values to other rolls from either data set. Given a scenario where a roll gets damaged, we can use this query to see if there are other rolls in both the storage data set and roll data set that contain the same diameter of the damaged roll.

To construct this query, it is a requirement to specify the *prefixes* of the ontologies we wish to use. As shown in the first two lines of Listing 1, and for most of our queries,

[5] https://www.w3.org/TR/sparql11-overview/

we have declared two prefixes: an empty prefix to denote our SCRO ontology and a *time* prefix to denote the time ontology that we have imported.

Then the main body of a SPARQL query is structured similarly to an SQL query. We start the query with the *Select* clause to select the fields we are interested in. In SQL, this would be one or more fields from a specific table. In SPARQL, we simply enter a variable name that will hold our results. Note that all variables begin with a question mark. As shown in Listing 1, we have chosen to select a variable called *?diameter* to denote the result of the SPARQL query will be related to the diametric value of the rolls. Then, we use the *WHERE* clause to condition our results. In our query, we specify that we are interested in the RDF triples whose subject contain the property *:hasDiameter*, where the *:hasDiameter* property can be any value. This subject is then stored in the *?roll* variable, and the actual *:hasDiameter* property values are stored in the *?diameter* variable. The *Minus* clause removes the subjects that also contain the *:hasGrindRoll* property as we are not interested in the historical rolls grindings data that previously contained this diameter. We then use "*Group by*" which creates columns for the fields we have selected. Generally, these will always be the same variables in our *Select* clause. In this example, we are only printing out the diameter variable.

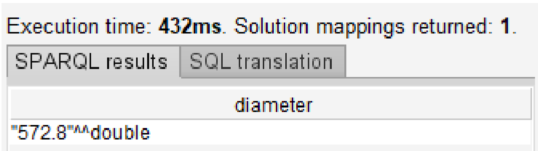

Figure 7. SPARQL result from Listing 1.

Figure 7 displays the results of this SPARQL query. The results show that *572.8* is the only diameter value that has three or more rolls that were last located between the 10^{th}-17^{th} of January 2020. We create another query to print out these rolls in Listing 2.

Listing 2: All rolls that have the diameter of 572.8

```
PREFIX time: http://www.w3.org/2006/time#
PREFIX : http://www.semanticweb.org/sadee/ontologies/2021/1/SCRO#

SELECT ?roll ?rollid ?partner ?diam
WHERE {
    ?roll :hasRollID ?rollid .
    ?roll :hasDiameter ?diam .
    OPTIONAL {
     ?roll :hasPartner ?partner .
    }
    MINUS {
     ?roll :hasGrindRoll ?grind .
    }
    FILTER (?diam = "572.8"^^xsd:double)
}
GROUP BY ?roll ?rollid ?partner ?diam
```

Listing 2 is a query written to display all the rolls that have the specific diameter of *572.8*. Similarly, we first select the ontologies we wish to use by declaring their prefixes. These are identical to our previous query. This time, however, our *Select* and *Group By* clauses contain the variables *?Roll*, *?Rollid*, *?partner*, and *?diam* which will be the columns containing our results. Once more, we use the *Where* clause to filter our results.

We created the variable *?roll* to store all the subjects that contain both the *:hasRollID* and *:hasDiameter* properties. The value of these properties are not specified and thereby can be any value. Each of these ?roll subjects may contain the optional property *:hasPartner*, but must not contain the *:hasGrindRoll* property.

Then, we filtered the *?diam* value to only return rolls that contained the diameter value of *572.8* which was the result from the first SPARQL query in Listing 1. Figure 8 displays the query result. Here we can see that *roll_1678* and *roll_1679* are partners that contain the diametric value of *572.8*. We can also see that there is a roll in storage with ID of *4631* that has the same diametric value and has no assigned partner. This type of query can be used to identify replacement rolls in case a roll gets damaged or needs replacing. Storage roll data is stored separately from active roll data, so this query skips the need for integration.

Execution time: **1.32s**. Solution mappings returned: **3**.

SPARQL results | SQL translation

roll	rollid	partner	diam
roll_1678	1678	1679	"572.8"^^double
roll_1679	1679	1678	"572.8"^^double
storage_roll_4631	4631		"572.8"^^double

Figure 8. SPARQL Result from Listing 2.

5. Ontology Validation

Ontology validation is a fundamental requirement when developing ontologies. It is essential to ensure that the quality of an ontology is adequate and the knowledge representation is accurate. There are many ways to validate ontologies; examples include task-based validation, criteria-based validation, data driven validation and expert knowledge validation [40]. In addition, a well known ontology validation tool known as "OntOlogy Pitfall Scanner" (OOPS!) [41] has been developed to validate ontologies by detecting common pitfalls aligned to a dimension classification developed in [42]. We use a combination of these approaches to validate SCRO.

5.1. OntOlogy Pitfall Scanner

Different pitfalls have different impacts and importance. Because of this, OOPS! categorises the evaluated results into three different levels: *critical, important and minor*. When evaluating SCRO, OOPS! displayed zero critical pitfalls, two important pitfalls, and a handful cases of minor pitfalls. The two important pitfalls are results from the P11 specification *"missing domain or range in properties"*. These include our object properties *"hasComponent"* and *"isComponentOf"*. However, according to [43], when using OWL, it is best practice not to specify the domain and ranges of superclasses but instead mention them in their respected subclasses. This is because the domain and ranges in OWL should not be viewed as constraints as this may cause unexpected classification and side effects [43], but rather viewed as *axioms* for reasoning. As the result of this, we have concluded to explicitly not specify the domain and ranges of these properties, but have included the domain and ranges of all the subclasses of these properties. For example, the object property *hasComponent* does not include a domain and range, but its subclass *hasPickleComponent* contains the domain *Cold Rolling Mill* and the range *Pickle Line*. On the other hand, *Minor* pitfalls include some elements missing annotations, or not explicitly declaring the inverse relationships of such object properties. These minor pitfalls do not affect the usability and consistency of the ontology and thus, remain as low-priority future changes.

5.2. Expert Knowledge Validation

As this work in linked closely with industry, we have validated our ontology with knowledge experts from Tata Steel. We set up a demonstration where we presented the SCRO Ontology to the domain experts where we received positive verbal feedback and small suggestions that have been implemented.

6. Conclusions

To conclude, this paper presents a novel Steel Cold Rolling Ontology that models and structures domain knowledge of cold rolling processes and activities within a steel plant. The purpose of the ontology is to improve data semantics and interoperability within the domain of smart manufacturing, which are the first steps towards achieving Industry 4.0. To our knowledge, this work is the first to develop an ontology for the cold rolling processes within a steel plant. The domain knowledge we have captured comes primarily from a case study with Tata Steel of their Port Talbot plant in the UK. We focus on capturing the knowledge for the Pickle line, Accumulators, and Mill sub-processes which are core to a cold rolling mill.

The ontology was developed using the eXtreme Design Methodology which includes using Ontology Design Patterns. We set up a case study that used real-world cold rolling data sets that were provided by the domain experts which validated the performance and functionality of SCRO. These data sets included roll data, roll refurbishment data, and roll storage data, all of which were in different tables and not integrated. We used the Ontop framework to deploy virtual knowledge graphs for data integration, data access, data querying, and condition-based maintenance purposes. SCRO was evaluated by both the ontology pitfall detection system *OOPS!* and domain experts from Tata Steel. OOPS! confirmed that there were no critical errors or inconsistencies in SCRO, and the domain experts confirmed that the knowledge in SCRO was uniform and accurate.

The domain knowledge encoded in SCRO is aligned with the processes and assets from the Port Talbot plant, which may differ from other plants from other companies. A key future goal will be to look at more cold rolling plants and compare any differences in processes and machinery to generalize the ontology, and add flexibility. Another future goal is to enhance the logic axioms for formalization of the knowledge. Presently, we have only mentioned basic axioms that show the relationships between classes and their properties. This paper does not include any logical constraints or logical connectives, whereas the ontology currently contains a few constraints, such as work rolls and backup rolls classes being disjoint. One future goal is to finish developing a full set of constraints for SCRO classes and properties. Finally, another future goal is to use SWRL rule reasoning techniques together with SCRO to perform rule-based reasoning for predictive maintenance purposes.

Author Contributions: Conceptualization, S.Beden. and Q.Cao.; methodology, S.Beden.; software, S.Beden.; validation, S.Beden., Q.Cao. and A.Beckmann.; formal analysis, S.Beden.; investigation, S.Beden.; resources, S.Beden.; data curation, S.Beden.; writing—original draft preparation, S.Beden. and Q.Cao.; writing—review and editing, S.Beden. and Q.Cao. and A.Beckmann.; visualization, S.Beden.; supervision, A.Beckmann.; project administration, A.Beckmann.; funding acquisition, A.Beckmann.

Funding: S. Beden was supported by the Engineering and Physical Sciences Research Council [grant number EP/T517537/1] and by Tata Steel. Q. Cao and A. Beckmann (in part) were supported by the Engineering and Physical Sciences Research Council [grant number EPSRC EP/S018107/1].

Acknowledgments: We would like to acknowledge Steve Thornton as the domain expert from Tata Steel during the development of the ontology.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Horvath, D.; Szabo, R. Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities? *Technological Forecasting and Social Change* **2019**, *146*, 119–132. doi:10.1016/j.techfore.2019.05.021.
2. Xiao, G.; Ding, L.; Cogrel, B.; Calvanese, D. Virtual Knowledge Graphs: An Overview of Systems and Use Cases. *Data Intelligence* **2019**, *1*, 201–223. doi:10.1162/dint_a_00011.
3. Peters, H. How could Industry 4.0 transform the steel industry. Future Steel Forum, 2017.
4. Miśkiewicz, R.; Wolniak, R. Practical Application of the Industry 4.0 Concept in a Steel Company. *Sustainability* **2020**, *12*, 5776.

5. Naujok, N.; Stamm, H. Industry 4.0 in steel: Status, strategy, roadmap and capabilities. Future Steel Forum, 2017.
6. Noy, N.F.; McGuinness, D.L. Ontology Development 101: A Guide to Creating Your First Ontology. Technical report, 2001.
7. Cao, Q.; Zanni-Merk, C.; Reich, C. Ontologies for manufacturing process modeling: A survey. International Conference on Sustainable Design and Manufacturing. Springer, 2018, pp. 61–70.
8. Roberts, W.L. *Cold Rolling of Steel*; 1978.
9. Schroeder, D.K.H. A BASIC UNDERSTANDING OF THE MECHANICS OF ROLLING MILL ROLLS **2003**.
10. Ray, A.; Mishra, K.; Das, G.; Chaudhary, P. Life of rolls in a cold rolling mill in a steel plant-operation versus manufacture. *Engineering Failure Analysis - ENG FAIL ANAL* **2000**, *7*, 55–67. doi:10.1016/S1350-6307(99)00004-7.
11. Zezulka, F.; Marcon, P.; Vesely, I.; Sajdl, O. Industry 4.0 – An Introduction in the phenomenon. *IFAC-PapersOnLine* **2016**, *49*, 8–12. 14th IFAC Conference on Programmable Devices and Embedded Systems PDES 2016, doi:https://doi.org/10.1016/j.ifacol.2016.12.002.
12. Vegetti, M.; Henning, G.P.; Leone, H.P. Product ontology: definition of an ontology for the complex product modelling domain. Proceedings of the Mercosur Congress on Process Systems Engineering, 2005.
13. Panetto, H.; Dassisti, M.; Tursi, A. ONTO-PDM: Product-driven ONTOlogy for Product Data Management interoperability within manufacturing process environment. *Advanced Engineering Informatics* **2012**, *26*, 334–348.
14. Lemaignan, S.; Siadat, A.; Dantan, J.Y.; Semenenko, A. MASON: A proposal for an ontology of manufacturing domain. IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06). IEEE, 2006, pp. 195–200.
15. Borgo, S.; Leitão, P. Foundations for a core ontology of manufacturing. In *Ontologies*; Springer, 2007; pp. 751–775.
16. Grüninger, M. Using the PSL ontology. In *Handbook on Ontologies*; Springer, 2009; pp. 423–443.
17. Cao, Q.; Giustozzi, F.; Zanni-Merk, C.; de Bertrand de Beuvron, F.; Reich, C. Smart condition monitoring for industry 4.0 manufacturing processes: An ontology-based approach. *Cybernetics and Systems* **2019**, *50*, 82–96.
18. Cao, Q.; Samet, A.; Zanni-Merk, C.; de Bertrand de Beuvron, F.; Reich, C. Combining chronicle mining and semantics for predictive maintenance in manufacturing processes. *Semantic Web* **2020**, pp. 1–22.
19. Bao, Q.; Wang, J.; Cheng, J. Research on ontology modeling of steel manufacturing process based on big data analysis. MATEC Web of Conferences. EDP Sciences, 2016, Vol. 45, p. 04005.
20. Horrocks, I.; Patel-Schneider, P.F.; Boley, H.; Tabet, S.; Grosof, B.; Dean, M.; others. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Member submission* **2004**, *21*, 1–31.
21. Wang, X.; Wong, T.; Fan, Z.P. Ontology-based supply chain decision support for steel manufacturers in China. *Expert Systems with Applications* **2013**, *40*, 7519–7533.
22. Dobrev, M.; Gocheva, D.; Batchkova, I. An ontological approach for planning and scheduling in primary steel production. 2008 4th International IEEE Conference Intelligent Systems. IEEE, 2008, Vol. 1, pp. 6–14.
23. Ugwu, O.; Anumba, C.J.; Thorpe, A.; Arciszewski, T. Building knowledge level ontology for the collaborative design of steel frame structures. Advances in Intelligent Computing in Engineering—Proceedings of 9th International Workshop of the European Group of Intelligent Computing in Engineering (EG-ICE), August, 2002, pp. 01–03.
24. Presutti, V.; Daga, E.; Gangemi, A.; Blomqvist, E. eXtreme design with content ontology design patterns. Proc. Workshop on Ontology Patterns, 2009, pp. 83–97.
25. Gangemi, A.; Presutti, V., Ontology Design Patterns; 2009; pp. 221–243. doi:10.1007/978-3-540-92673-3_10.
26. Blomqvist, E.; Presutti, V.; Daga, E.; Gangemi, A. Experimenting with eXtreme Design. Knowledge Engineering and Management by the Masses; Cimiano, P.; Pinto, H.S., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; pp. 120–134.
27. Musen, M.A. The protégé project: a look back and a look forward. *AI Matters* **2015**, *1*, 4–12. doi:10.1145/2757001.2757003.
28. G., A.; van Harmelen F.. *Web Ontology Language: OWL*; 2004. doi:10.1007/978-3-540-24750-0_4.
29. Hobbs, J.R.; Pan, F. Time ontology in OWL. *W3C working draft* **2006**, *27*, 133.
30. Xiao, G.; Lanti, D.; Kontchakov, R.; Komla-Ebri, S.; Güzel-Kalaycı, E.; Ding, L.; Corman, J.; Cogrel, B.; Calvanese, D.; Botoeva, E. The Virtual Knowledge Graph System Ontop. The Semantic Web – ISWC 2020; Pan, J.Z.; Tamma, V.; d'Amato, C.; Janowicz, K.; Fu, B.; Polleres, A.; Seneviratne, O.; Kagal, L., Eds.; Springer International Publishing: Cham, 2020; pp. 259–277.
31. Calvanese, D.; Giacomo, G.D.; Lembo, D.; Lenzerini, M.; Poggi, A.; Rodriguez-Muro, M.; Rosati, R.; Ruzzi, M.; Savo, D.F. The MASTRO system for ontology-based data access. *Semantic Web* **2011**, *2*, 43–53.
32. Priyatna, F.; Corcho, O.; Sequeda, J. Formalisation and Experiences of R2RML-Based SPARQL to SQL Query Translation Using Morph. Proceedings of the 23rd International Conference on World Wide Web; Association for Computing Machinery: New York, NY, USA, 2014; WWW '14, p. 479–490. doi:10.1145/2566486.2567981.
33. Calvanese, D.; Cogrel, B.; Komla-Ebri, S.; Kontchakov, R.; Lanti, D.; Rezk, M.; Rodriguez-Muro, M.; Xiao, G. Ontop: Answering SPARQL queries over relational databases. *Semantic Web* **2016**, *8*. doi:10.3233/SW-160217.
34. Bagosi, T.; Calvanese, D.; Hardi, J.; Komla-Ebri, S.; Lanti, D.; Rezk, M.; Rodríguez-Muro, M.; Slusnys, M.; Xiao, G. The Ontop Framework for Ontology Based Data Access. The Semantic Web and Web Science; Zhao, D.; Du, J.; Wang, H.; Wang, P.; Ji, D.; Pan, J.Z., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2014; pp. 67–77.
35. Rodríguez-Muro, M.; Rezk, M. Efficient SPARQL-to-SQL with R2RML mappings. *Journal of Web Semantics* **2015**, *33*, 141–169. Ontology-based Data Access, doi:https://doi.org/10.1016/j.websem.2015.03.001.
36. Sequeda, J.f.; Tirmizi, S.h.; Corcho, O.; Miranker, D.p. Review: Survey of Directly Mapping Sql Databases to the Semantic Web. *Knowl. Eng. Rev.* **2011**, *26*, 445–486. doi:10.1017/S0269888911000208.

37. Jiménez-Ruiz, E.; Kharlamov, E.; Zheleznyakov, D.; Horrocks, I.; Pinkel, C.; Skjæveland, M.G.; Thorstensen, E.; Mora, J. BootOX: Practical Mapping of RDBs to OWL 2. The Semantic Web - ISWC 2015; Arenas, M.; Corcho, O.; Simperl, E.; Strohmaier, M.; d'Aquin, M.; Srinivas, K.; Groth, P.; Dumontier, M.; Heflin, J.; Thirunarayan, K.; Staab, S., Eds.; Springer International Publishing: Cham, 2015; pp. 113–132.
38. Pinkel, C.; Binnig, C.; Jiménez-Ruiz, E.; May, W.; Ritze, D.; Skjæveland, M.; Solimando, A.; Kharlamov, E. RODI: A Benchmark for Automatic Mapping Generation in Relational-to-Ontology Data Integration. 2015, pp. 21–37. doi:10.1007/978-3-319-18818-8_2.
39. Soylu, A.; Kharlamov, E.; Zheleznyakov, D.; Jimenez-Ruiz, E.; Giese, M.; Skjaeveland, M.G.; Hovland, D.; Schlatte, R.; Brandt, S.; Lie, H.; Horrocks, I. OptiqueVQS: A visual query system over ontologies for industry. *Semantic Web* **2018**, *9*. The final publication is available at IOS Press through http://dx.doi.org/10.3233/sw-180293, doi:10.3233/SW-180293.
40. Brank, J.; et al.. Gold standard based ontology evaluation using instance assignment. IN: PROC. OF THE EON 2006 WORKSHOP, 2006.
41. Poveda-Villalón, M.; Gómez-Pérez, A.; Suárez-Figueroa, M.C. OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)* **2014**, *10*, 7–34.
42. Gómez-Pérez, A. Ontology evaluation. In *Handbook on ontologies*; Springer, 2004; pp. 251–273.
43. Horridge, M.; Knublauch, H.; Rector, A.; Stevens, R.; Wroe, C. *A Practical Guide To Building OWL Ontologies Using The Prot'eg'e-OWL Plugin and CO-ODE Tools*; 2004.