Performance evaluation of collaboration in the design process: Using interoperability measurement

Amir Pirayesh Neghab a,*, Alain Etienne a, Mathias Kleiner b, Lionel Roucoules b

a LCFC Laboratory, Arts et Métier ParisTech, Metz Campus, France
b LSIS Laboratory, Arts et Métier ParisTech, Aix-en-Provence Campus, France

Abstract

A design process, whether for a product or for a service, is composed of a large number of activities connected by data and information exchanges. The quality of these exchanges, called in this paper collaboration, requires the ability to exchange useful, understandable and unambiguous data and information to the different designers involved. In this paper, a global framework is first set for process/product performance management. Then, the research question focuses on the definition and evaluation of the performance of collaborations, and by extension, of the design process in its entirety. This performance evaluation requires the definition of several key elements such as object to evaluate, the performance criteria, indicators and action variables. In order to define the object of evaluation, this paper relies on a literature study on collaboration resulting in an ECORE meta-model of collaborative processes. The collaboration performance measurement is for its part based on the concept of interoperability. This measure estimates the technical and conceptual interoperability of the different pairwise collaborations. The paper is concluded by proposing a toolbed methodology for collaborations' performance evaluation including two main phases: process modeling and interoperability measurement. Tooling is provided through the Eclipse Modeling Framework (EMF) using its (meta-) model edition, constraint validation and model comparison features. The applicability of the methodology is also illustrated using a case study in design.

1. Introduction

Given the studies on product design such as [1,2], the goal of the design process is to develop a product that meets customer needs. Indeed, in this stage of the product life-cycle, the data, characterizing customer needs, must be processed in several artifacts in order to obtain knowledge concerning the product to design [3], and thus create one or more representations of the latter [4]. Howard [5] studies the phases and milestones of several design process, from the needs expressed to the product definition.

In the current design processes, collaboration is an intrinsic characteristic, particularly in the context of multidisciplinary design [6,7]. In fact, the design process is moving towards collaborative design process, which requires and encourages collaborations between designers from the beginning of product life cycle [8]. According to [9,10], there exist a plethora of collaboration definitions describing several types of interactions. These interactions in the design process mainly occur when designers communicate the results of their activities that support product data exchange [11,12]. According to Chiu [11], communications and data exchanges are prerequisites for collaboration. In addition, the author did an experiment in an academic environment that indicates that the designers have spent almost half of their time focusing on communication. It is also stated that effective communication is critical to the project participants of collaborative design.

It is observed that designers may encounter problems in their collaborations such as technological problems in data transmission or in understanding the transmitted data [13]. According to [17], some of these problems, particularly the different interpretations of data, are results from the use of this data in heterogeneous environments. These heterogeneities are based on the differences between computing environments, languages, techniques, tools [14], and data sources [15,16], in different areas of expertise. Such heterogeneities are noted by [17], [18] and [19] as sources of problems in collaborations. The heterogeneity problems often cause irrelevant [20], inadequate, imprecise or ambiguous data [21,20]. These problems might also cause failures in the individual tasks of designers [22]. This is the reasons why for more than a
decade, a large number of references indicate the need for studies on collaborations and the management of their performance [23–25]. Despite these indications, very few theories explain how to manage such performance particularly in a situation where information plays a key role [26].

The word performance can be applied in different terms, such as “performance management”, “evaluation or measurement of performance” and “performance assessment” to describe the process of capturing performance [27]. In this paper, the research approach is built around the key elements that form a problem of performance evaluation. Based on [28], it is considered that the evaluation of the performance of an object (e.g., a company, a design process or a product) comprises the following elements: criteria, indicators, action and measurement variables. Among these elements, the criterion choice plays a significant role since it reflects the objective of the evaluation. The performance assessment can be based on classical criteria such as cost, time and quality; in this paper, considering the significance of collaboration, it is decided to retain a criterion that intrinsically belongs to this concept and addresses the quality aspect of collaboration (see Fig. 1). This criterion is called Interoperability.

Several classifications and evaluation approaches of collaboration were studied to identify such criteria. Husted and Michailova [29] classify collaboration as: infantile, repeated and mature based on collaboration history, where some classifications address the collaboration results. In these classifications, the collaborations can be successful, significant and strong or failure, non-significant and weak [30–33]. The third category of classification provides one or more criteria based on the difficulties encountered by collaborative actors. For example, Girard [34] classifies collaboration as: free, encouraged or forced based on the freedom of actors to collaborate. Indeed, in this classification, organizational difficulties (i.e., authorizations and motivations) are underlined. It is observed that the classifications and assessment approaches, such as [35,36], which are based on a criterion called “Interoperability”, are more suited to evaluate the collaboration performance. In fact, this concept covers a larger spectrum of capacities required in collaboration. Regarding the definitions proposed by [37,38], the goal of interoperability is to overcome the problems in data exchanges [22].

In this paper, in order to clarify this concept with the objective of finding at least one of its indicators, the Chen’s framework is adopted [39]. This framework, as illustrated in Fig. 2, structures interoperability in three dimensions: Interoperability barriers, concerns and approaches.

Concerning the first two dimensions of Chen’s framework, this paper studies the technological and conceptual barriers of interactions of the type data exchange between designers. This paper also addresses the third dimension since any interoperability approach or solution, applied in a collaborative process, should be identified and studied in the evaluation of the collaborations’ performance. This scope is located in Chen’s framework (see Fig. 2). Considering this scope, the present paper focuses only on the problems of data adequacy. According to [26,40] such problems are in expansion. In this paper, it is primarily considered that the data are always produced correctly and are understood by the designers. This assumption allows to better focus on the inadequate data since this problem is disassociated from the accuracy and comprehensibility problems. Data adequacy is then considered as an indicator of interoperability. Here it is considered that this indicator reaches its maximum value when the data produced by the actors: (1) can be accessed by others (this condition corresponds to the technical aspect of this indicator), and (2) are sufficient and useful for the others. This condition corresponds to the semantic aspect of this indicator.

The design process in this paper is considered to be in its evaluation and validation phase. Therefore, the objective is to evaluate this process before its execution. In this case, only the backbone of the process, including the sequence of activities and resources, is available. In such process, product data are not yet instantiated. However, some information is available on them: file exchange formats and file abstractions (the file contents and structure of data) in case of availability of these abstractions. The extraction of the abstraction is usually a tedious or impossible task, especially when the format is not open or is not based on an open standard. Therefore, this paper is limited to the terms used for file formats when the abstractions are not available. Having this information on the data files, data adequacy conditions are translated as the Compatibility of file formats and contents.

Considering the fact that in this paper collaboration performance is defined as the ability of actors to exchange adequate data, the following questions are raised here: is it possible to quantify and improve the collaboration performance? How the latter is affected by the elements of collaboration (actors and their resources) or by their heterogeneities? Facing these questions, it is first considered that collaboration performance, as a part of process performance, impacts product performance. Such impacts are mentioned previously in the field of risk analysis, particularly in [41,42], which study the links between process and product performances based on product failures.

As illustrated in Fig. 3, the performance evaluation based on Interoperability (i.e., 1a) and the parallel evaluations based on other criteria (i.e., 1b) can be positioned in a global framework supporting the management of the design process. This framework is founded on the link between the results of the process and product performances through product failure analysis (i.e., 2). This analysis shows that all the results of process performance evaluation can be analyzed by focusing on the problems causing product failures. In the generating step (i.e., 3), alternative processes can be generated by modifying the elements of collaborations identified as critical in the previous step. During the selection step (denoted 4), an optimization model is used to maximize the value of the product. Considering the Fig. 3 and
facing the research questions, in this paper the proposition for step 1a of this process is discussed. This proposition is a tool to methodology for performance evaluation of collaboration based on data adequacy as one indicator of interoperability.

2. Literature review

This section presents at first the literature of interoperability measurement. Then interoperability solutions are discussed.

2.1. Interoperability measurement

A considerable amount of work has been done on the need for interoperability inside and outside organizations. A statistical study of this concept indicates that it has been used as a keyword in the literature for nearly 50 years [43]. Kasunic and Anderson [44] believe that interoperability is a broad and complex topic and the development of specific measures and their application is a difficult task. However, these authors recognize that assessing interoperability with well-chosen measures is essential to identify priorities to consider within business collaboration. Many researchers have been interested in such assessments and many approaches are proposed in the literature. Some of them are studied in [36]. These approaches can be studied in two categories: evaluation of the interoperability based on independent criteria such as cost, time or quality and evaluation of the level of realization of interoperability based on criteria more depending to this concept such as degree of coupling or compatibility and the ones evaluating. Approaches proposed by [45–47] are based on cost, time or quality. Such approaches evaluate directly these criteria in environments requiring interoperability or the impact of interoperability barriers on these criteria. For instance, Chen et al. [47] propose an approach for quantitative evaluation of the cost, time and quality in an interoperating process. In the approach proposed by [35,47], the interoperability potential is evaluated through a set of provisions that may impact the interoperability, such as the flexibility and openness of the actors of collaboration. These provisions can be assessed the interoperability of an object without prior knowledge of its collaborative partners. This assessment identifies the risk of encountering problems during the establishment of collaboration. Four criteria are applied in this assessment: openness, coupling, centralization and configurability. Each criterion has two quantifiable values. For example, an object can be assessed as open or close, based on the first criterion.

Other authors consider that the coupling degree between collaborative objects (actors and systems) indicate their interoperability. Bianchini [48] proposes an approach that firstly expresses the semantic relationships between these objects since each object is modeled as a collection of service providing processes. Thus, the interoperability evaluation is performed based on the degree of coupling between these processes. In the approach proposed by [49], the degree of coupling is measured by extent and intensity of information indicators. Actors are considered not coupled or strongly coupled when these indicators are respectively 0 or 1. In the first case no information exchange between actors of collaboration are realized and in the second case, all information are available to be shared and used by all these actors.

Ducq and Chen [50] consider the compatibility measurement in their interoperability evaluation approach. In this approach, a questionnaire is applied which takes into account interoperability barriers and concerns introduced before in section 1. The results of this questionnaire are linked to a table including the values of compatibility measurement for each pair of barrier-concern: 1 when non incompatibility is detected and 0 otherwise. Ford [51] proposes a methodology in two steps to quantitatively measure the interoperability. The first step is called Measuring I-score (Interoperability-score) through the quantification of the intrinsic property of a given network called interoperability spin. This property can qualify the value of interaction between pairs of objects: 1 when interoperability can be achieved without man or machine intervention for information transfer, 0 and −1 respectively in case of machine or human intervention. The second step is the measurement of optimum I-score if modifications are possible which affect the interoperability barriers.

Some of the interoperability measurement approaches are based on maturity models that evaluate the stages an object must follow to achieve a required level of performance for interoperability. Several maturity models have been developed: Spectrum of Interoperability Model (SoIM) [52], The Quantification of Interoperability Methodology (QoIM) [53], Military Communications & Information Systems Interoperability (MCISI) [54], Levels of Information System Interoperability Model (LISI) [55], Interoperability Assessment Model (IAM) [56], Organizational Interoperability Model (OIM) [57], Levels of Conceptual Interoperability Model (LCIM) [58], Enterprise Interoperability Maturity Model (EIMM) [59], ISO 11354-1:2010 [60]. These maturity models may allow assessment of potential interoperability or the ability to interoperate with others. Some of these models are studied and characterized by [61].

Yahia et al. [36] study interoperability in data exchanges only based on conceptual barriers. In this approach, a method called semantic matching plays a specific role particularly for the measurement of compatibilities. According to [15], similar methods are applied in the following areas: ontology integration, message transformation between databases or data directories and more recently in the coordination of web services [15]. As mentioned above, the ultimate goal of matching is to find the differences between two objects. These differences are named in several ways in the literature: dissimilarity [15], mismatch [62] or conflict [19]. Different methods of matching are proposed in the literature [63]. Detailed literature review are carried out in [64]. Concerning technologies or tools, which have the objective of facilitating the matching process, Lu conducted a study on the technologies applied in certain categories of matching, especially on the basis of automation of the similarity measure [65]. Many of these technologies have also been studied in [66,67].

Fig. 3. Framework for process/product performance management including interoperability measurement.
Although matching is widely applied as compatibility measurement approach, the choice of its method, operator and tool depends mainly on the objectives of matching, objects to be studied and their characteristics [84]. In design processes including data modeling approaches such as Model Driven Engineering, the matching is essential for the management of models [68]. In fact, when the models are large and complex (containing a large number of entities), dedicated tools are needed to compare them [86]. To this end, model comparison techniques are developed. Stephan and Cordy [69] discuss model-to-model comparison techniques. They illustrate and evaluate qualitatively comparison techniques, while highlighting the strengths and weaknesses of each the studied techniques. These authors show that EMF Compare, which is applied in this paper, has the best overall performance compared with techniques such as ECL, C-Saw, DSMDiff, and SmoVer. However, EMF Compare, ECL, and SmoVer only work with Eclipse Modeling Framework (EMF), but other techniques can be extended to work with other frameworks and modeling languages [69]. “EMF is a modeling framework and code generation facility for building tools and other applications based on a structured data model” [70]. An example of application of EMF compare can be found in [71], which introduces a new approach for comparing models using this tool. This approach is based on a comprehensive comparison of model components. The author proposes also a system of weights which calculates the magnitude of the difference between two elements of a model [71].

2.2. Interoperability solutions

In interoperability measurement, the actions available that can improve the measured value must be also identified. These action plans can be interoperability solutions embraced by the approaches of interoperability, the third dimension of the framework of Chen presented earlier in the introduction section. Concerning the development and application of these solutions, companies may have different modeling strategies. Indeed, modeling of the applied domains, particularly in the design context, can be based on a unique model for the process or based on the application of several business models (see Fig. 4). These models, standard or based on an agreement within the company, focus on file exchange formats, modeling languages and meta-models (ontology). In the case of using multiple models in collaborations, complementary interoperability solutions such as model transformation is required to resolve the issue of heterogeneous models.

2.2.1. Standard models: data file formats, modeling languages and ontology

This solution focuses on the use of standards models and formats (e.g., STEP) in data exchanges [72–74] conducted surveys on the standards used in the life cycle of products. In [73], it is firstly noted that the exchange of information, between producers and consumers of information, requires a language that convey the content of the information. This language is defined mainly by the type of this content, its structure and the data instances [73].

According to Terzi [74], many standards cover product development, production and use. Currently, several CAD/CAM applications contain a module to read and write data defined by one of the STEP Application Protocols [75]. The lack of flexibility, dynamism and automation of standards and the increasing volume of product data have led to the development of ontology-based approaches [17].

Regarding standard models, ontology and meta-models can also be mentioned. According to [76], a product ontology, obeying the rules of the field, provides the opportunity to express and share product data between: actors, business applications and information systems. Furthermore, in a collaborative environment, ontology play a vital role in the interoperability of different systems in a particular area by emphasizing on the semantic aspect of data integration [77].

In the context of product development, domain ontology usually offers a description of areas: simple, comprehensive, understandable and implementable by human. A recent survey of domain ontology is presented by [17]. This survey, mainly about the use of inference-oriented ontology (reasoning about data), studies the ontology in three dimensions: stages of the life cycle of products, the scale of the proposed models and finally the involved fields (product, process or service) [17]. Ontology can also be used in annotation to overcome semantic interoperability barriers. According to [78], annotation is a factor that enhances the semantic interoperability as it promotes the exchange of data between collaborative actors. In fact, these actors can annotate the exchanged data to enrich the semantics of their entities.

2.2.2. Model Transformation

Using a single standard or reference model as interoperability solution requires an effort by the users to understand the concepts and the methods of implementation. Therefore, the standard must be adapted to all stakeholders in a consensual manner, since in some areas, it is almost impossible. Currently, considering the diversity of actors in the product development, several standards or models are used in the data exchange. In this case, some treatments on models such as the model transformation should be considered in order to achieve interoperability. Model Transformation is defined as the process of converting a system [product] model to another model of the latter [78,79]. Czarnecki et al. [80] propose a taxonomy of classification of different existing approaches for model transformation. This taxonomy is described by a functionality model, which explicitly supports choices of model transformation development. The authors also classify existing approaches as follows: direct manipulation, relational, graph-based, and structure-driven and hybrid. The mechanism of model transformation is more detailed here since it is applied in thecased study of this paper. For this purpose, the architecture illustrated in Fig. 5 is applied which is based on Meta Object Facility [81].

---

**Fig. 4.** Modeling strategies.

**Fig. 5.** Architecture of Model Transformation in MOF [81].
In this architecture, the main objective is to transform a source model to a target model. These models, at M1 level, conform to their own abstractions, the meta-models at M2 level, and these abstractions conform to a single reference model, the meta-meta-model at M3 level. Regarding collaborative processes in design, the source model contains product data that are provided by a design activity and the target model is based on the data required by other activities. These models conform to their own meta-models or, in other words, to their modeling languages. The transformation is based on transformation rules: mapping. Miller defines the latter as the specification of a mechanism for transforming the elements of a model, conforming to a meta-model, particularly into elements of another model, conforming to another meta-model [79]. The transformation also requires a transformation language that implements the transformation rules. This language itself may conform to M3 level in architecture of transformation.

3. Methodology

Considering the global framework presented in the introduction section, the proposed methodology addresses the first step of this framework (Step 1a in Fig. 3). Enterprises, developing products through multiple data exchanges in their processes, can apply such methodology. The methodology is composed of two major phases (illustrated by the IDEF0 diagram in Fig. 6):

Phase 1a.1, process modeling: this phase includes modeling the design process by putting an emphasis on collaborations and on pair of data exchanges between designers. In this phase, the first task is the process model instantiation. In the second task, the process model is completed according to process expert knowledge.

Phase 1a.2, interoperability measurement: this phase consists of a measure of performance based on interoperability. In this phase, according to the information provided by the previous phase (1a.1), the value of an indicator of interoperability, Data Adequacy, is measured. In the first task of this phase, the syntax matching in the pairwise data exchanges are verified. Then the second task is performed in order to quantify the semantic matching by comparing data abstractions.

In the following, the tooling environment is first introduced. Then, the key elements of this methodology are detailed.

3.1. Tooling environment

In order to support the methodology proposed in this paper, it is suggested to use a set of tools provided by the Eclipse Modeling Framework (EMF), chosen as the implementation framework because of its maturity and large tool support.

The different chosen tools and their use in each phase is illustrated in Fig. 6 (as bold italic). These tools are briefly introduced below. Their concrete application in our methodology is detailed in the following sections along with its different steps.

ECORE [82] is a meta-meta-model, based on OMG’s MOF 2.0 [83], which implements its subset Essential-MOF (EMOF). ECORE allows defining domain specific languages (in our case, the process meta-model). All the other EMF tools act on ECORE meta-models. EMF provides a default model editor (in our case, to instantiate a given process model), and offers the possibility to define custom editors, including graphical ones (using for instance GEMF). EMF also offers a set of validation tools, built around the Object Constraint Language (OCL) [83], allowing to further specify and check various properties on created models. Finally, EMF compare is a tool for analyzing the differences between (meta-) models, which in our case is used in the semantic matching phase.

3.2. Meta-model of collaboration

This meta-model is the core of phase 1a.1, which aims to analyze the collaboration performance. The collaboration must be firstly defined by emphasizing on its key elements. Concerning this definition, a limit was initially expressed by [10]: a clear and conventional definition of collaboration is difficult to achieve. Facing these issues, several definitions of collaborations were studied and analyzed. As an example, the following definition can be mentioned which is proposed in the field of collaborative design [34]: “collaboration is defined according to two aspects. The first concerns the freedom of designers to collaborate with each other. The second is a reflection of the design team’s collaborative experience.” The extracted definitions from literature were analyzed using text-mining tools: Text Content Analysis Tool [139] and KH Coder [140]. These analyses led to identification of key elements of collaboration. These elements are modeled as a conceptual model of collaborative process. The conceptual model is also extended based on the elements introduced by the use of interoperability solutions in the design processes. Eventually, this conceptual model is formalized as an ECORE meta-model, represented in Fig. 7 as a class diagram, giving our methodology access to EMF’s model edition and validation tools for process instantiations. A set of OCL constraints further specify the validity of the instantiated processes.

Below is a brief explanation of the main concepts of this meta-model:
"CollaborativeProcess", representing the whole design process, is the root of the meta-model.

"DesignActivity" represents the design activities in the process with a single designer.

"CollaborativeActor" models the designers working on design activities with their software tools.

"ProductDataFile" represents the exchanged files containing the product data. Each file has a status: produced, delivered or required. Indeed, a file can be produced, required by a design activity or it can be delivered to an activity by communication. Each data file has also a file format.

"AbstractionOfDataFile" represents the data structure of its associated data file.

"CommunicationType" represents a particular mode of communication available in the process. It can be a solution, which affects the syntax or semantics of a data file such as model transformations. In fact, a Communication Type is not a design activity parameter but it supports the interoperability. It can be applied in one or several communications.

"Communication" models the unilateral and unidirectional data flow from one activity to downstream one. Each communication has a sequence that designates the chronological combination of Communication Types applied in the communication process.

"CollaborativeWork" is an intermediate class that embraces two design activities associated with a unique communication. In reality, this class is instantiated by the existence of these elements.

3.3. Granularity of the measurement, pairwise collaboration

In the proposed methodology, a collaborative process is studied by focusing on its sub-elements that meet still respect the definition of collaboration. Indeed, a collaborative process can be decomposed into minimal elements which themselves establish collaboration. This minimal unity, called Pairwise Collaboration (PwC), is formed by a communication between two activities. Indeed, a PwC, illustrated in Fig. 6, it-self is a collaboration that mainly contains:

Two actors with their own attributes, activities and objectives, but with a shared objective realized through collaboration,

two activities, one producing structured data and one consuming the produced data,
a communication, single and unilateral data exchange, receiving and delivering produced data.

3.4. Performance measurement of a pairwise collaboration

As explained in Section 3.1, the communication element of collaboration may include other elements called Communication-Types that can modify data before their delivery. Therefore, it is considered that the measurement of the collaboration performance should be performed in two [vertical] levels (Fig. 8):

Potential level: This level indicates the compatibility between produced and required data.

Current level: This level indicates the compatibility between delivered and required files.

In order to measure the Data Adequacy in each level, two matching methods are used. In these matching, formats and abstractions of data files are the measurement variables. Having the value of these variables from the process model, the syntax and semantic matching are performed respectively associated with formats (A, B and C in Fig. 8) and the abstractions (a, b and c in Fig. 8).

Fig. 7. Excerpt of collaboration meta-model.
(A) Syntax matching

For the syntax matching of file formats, a simple similarity function (based on terminological similarity of terms) is used. This function verifies the terminological equivalence of file formats. For example, having two files $F_D$ (delivered) and $F_R$ (required), with respective formats $B$ and $C$ (see Fig. 8):

\[
B = C \rightarrow \text{Similarity function } (F_D, F_R) = \text{True} \land B \neq C \rightarrow \text{Similarity function } (F_D, F_R) = \text{False}
\]

(1)

Based on these similarity functions, for all collaboration the matching rules indicated in Table 1 must be verified. Two mechanisms, manual and automatic, are proposed for this purpose:

Manual syntax matching: The rules of syntax matching can be verified manually. To this end, the process model must first be studied for each collaboration (represented by the Collaborative Work class) in order to identify exchanged files. Then, the formats of these files are compared based on matching rules. The manual syntax matching can be tedious and complex for cases with several collaborations. Therefore, an automatic mechanism is also proposed.

Automated syntax matching: the use of OCL constraints is proposed to increase the automation of matching. These constraints are added to the meta-model of process based on the matching rules. In this meta-model, the class “Collaborative Work” is the context of the OCL constraints.

In order to increase the precision of data compatibility measurement, the matching can be done in a second level, semantic matching, in case of having open formats and availability of the data structures.

(B) Semantic matching

In the methodology, the semantic matching is defined based on the data compatibility at a conceptual (abstract) level. Here, an abstraction $C$ is defined as a three level conceptual structure:

\[
\text{Abstraction } C = \{c_1, \ldots, c_n\}
\]

where, $c_i$ is a class in $C$ and

\[
c_i = \{cp_{i1}, \ldots, cp_{im}\}
\]

where, $cp_{ij}$ is a property of the class $c_i$. The first property of the class is $cp_{i1}$ is unique (the mandatory identifier which is mainly the class title) but other properties are optional; they may change from one concept to another. An optional property can be independent, as an attribute, or related to another class, as an association. Here, the heritage (generalization) association is not considered as an association. Meanwhile, each class having such relation inherits the optional properties of its upper class. The optional properties can also have their own obligatory and optional properties:

\[
cp_{ij} = \{cp_{i1j}, \ldots, cp_{ijm}\}
\]

(4)

These properties are limited based on the formalism of abstraction. For example, in this paper for a UML class diagram or an EMF meta-model, three properties are considered:

- Properties of attributes: type, lower and upper bounds,
- Properties of associations: type (aggregation or composition), min or max cardinalities.

In the proposed process of semantic matching, the abstractions must first be adapted to the above definition of an abstraction. Once this condition is verified, a similarity function based on textual equivalence of the entities of abstractions (concepts) is used for semantic matching. Indeed, it is considered that equivalent terms represent a unique concept for designers. This assumption is due to an issue located mainly outside the perimeter of this paper (it can be studied in linguistics). However, in case of availability of domain ontology in the design process, this assumption is more realistic.

Having two abstractions $X_{source}$ and $X_{target}$ based on the same concrete syntax, the similarity function verifies if the source includes all the concepts identified as required in the target:

\[
\text{Similarity function } (X_{source}, X_{target}) = \text{True} \rightarrow X_{source} \supset X_{target} \left( \forall c \in X_{target} : \exists c' \in X_{source} : c' = c \right)
\]

(5)

The results of the similarity function can show mismatches, missing elements of the target in the source. Based on the conceptual levels of an abstraction a mismatch can be:

- Mismatch between the concepts of the first level (classes): This type of mismatch occurs when a class of $X_{target}$ does not exist in the $X_{source}$.
- Mismatch between the concepts of the second level (properties): This type of mismatch occurs when for a class of $X_{target}$

Table 1 Syntax matching rules.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Matching level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Potential: In collaboration, the “produced data” with file format “A” should have the same file format “C” as the “required data”. → Similarity function: $A \rightarrow C$</td>
<td>OCL constraint 1 (context: CollaborativeWork): self.hasAvProducing.avProduces.format = self.hasAvRequiring.avRequires.format.</td>
</tr>
<tr>
<td>2 Current: In collaboration, the “delivered data” with file format “B” should have the same file format “C” as the “required data”. → Similarity function: $B \rightarrow C$</td>
<td>OCL constraint 2 (context: CollaborativeWork): self.hasAvProducing.avProduces.format = self.hasAvRequiring.avRequires.format.</td>
</tr>
</tbody>
</table>
already exists in \( X_{source} \) but a property; attribute or association, of this class is missing in \( X_{source} \).

Mismatch between the concepts of the third level (properties of properties): This type of mismatch occurs: when a class of \( X_{target} \) and its attribute/association exist in \( X_{source} \), but a property of this attribute/association, for example the lower cardinality is different.

The mismatches can be weighted according to their relative importance. This weighting can be based on the opinion of the expert process or actors of collaboration. In this paper, these weights are respectively set to 1, 0.5 and 0.1.

After the identification of mismatches, the semantic matching value is estimated by this equation:

\[
\text{Semantic Matching Value} = 1 - \frac{\text{WME}}{\text{WE}}
\]  

(6)

where, \( \text{WME} \) and \( \text{WE} \) are respectively the weighted values of the elements of \( X_{target} \) and the weighted value of mismatches (missing element of \( X_{target} \)). The parameters of Eq. (6) are calculated as below:

\[
\text{WE} = \text{NC} \times \text{WC} + \text{NP} \times \text{WP} + \text{NPP} \times \text{WPP}
\]  

(7)

\[
\text{WME} = \text{NMC} \times \text{WC} + \text{NMP} \times \text{WP} + \text{NMP} \times \text{WPP}
\]  

(8)

where,

\( \text{NC} \): Number of concepts of level one (classes) in \( X_{target} \),

\( \text{NP} \): Number of concepts of level two (optional properties) in \( X_{target} \),

\( \text{NPP} \): Number of concepts of level three (optional properties of properties) in \( X_{target} \),

\( \text{WC}, \text{WP} \) and \( \text{WPP} \): weights of the concept in level 1, 2 and 3. These weights must defined by the designer receiving the data based on the relative importance of concepts at each level.

\( \text{NMC}, \text{NMP} \) and \( \text{NPP} \): Number of mismatches between the concepts of level 1, 2 and 3.

The calculated value of semantic compatibility can be compared to a threshold defined based on the opinion of actors of collaboration. This threshold indicates the amount of data considered as sufficient for these actors. Here, the minimum threshold is set to 0.5. For all collaboration, semantic matching rules included in Table 2 must be therefore verified.

The semantic matching as well as the syntax matching is performed at two levels in collaboration and results may indicate a current or a potential value of the performance. The same as syntax matching, semantic matching can be performed manually or in an automatic way:

Manual semantic matching: The rules of matching semantics can be verified manually. To this end, the process model must first be studied for each collaboration (represented by the CollaborativeWork class) in order to identify files exchanged. Then, the abstractions of these files are used in the verification of matching rules. The manual matching can be tedious and complex cases with abstractions with several concepts of in case of having several collaborations in the process. Therefore, an automatic mechanism to achieve this matching is also proposed.

Automated semantic matching: in order to increase automation in the comparison of data models (based on the semantics of the design domain), it is proposed to use EMF Compare. This tool provides automatic matching between two EMF models based on structural and textual comparisons. However since these comparisons may sometimes be inaccurate regarding the semantics of the design domain, the given results should only be considered hints and require further verifications by a design/ process expert.

4. Case study

In order to verify the applicability of the proposed methodology, a case study is performed. The latter is based on a design process of a mechanical coupling between a propeller and a diesel engine. The general information on this case is first presented. Then, the application of the methodology and obtained results are discussed. It should be mentioned that the application of the methodology in this case study is with a higher objective for its use (the process management). This objective is the comparison and analysis of design process and its alternatives in order to select the best one based on the results of interoperability measurement. The case focuses on a part of this process illustrated in Fig. 9.

As illustrated in Fig. 9, a unique Pairwise Collaboration (PwC) is analyzed. The objective of this collaboration is to exchange data on the product components and their energy flows, based on a model called SK2 [84], in order to facilitate the geometrical design. In this collaboration, the first and second activity respectively produces and requires data. Therefore, the focus is on the results of the first activity and the data required by the second one. The required data must be in STEP AP-203 format, which is considered as an exchange format for geometrical data. For the communication part of this collaboration, an alternative process exists in the process.

In Fig. 9, the alternative process is a prototype that has been developed based on a federative approach in order to verify the possibility of replacing design activities by model transformations. Therefore, two consecutive model transformations are applied in the collaboration. The first one transforms a SK2 model into and intermediate model called ASB; the second one transforms the latter into an STEP-AP203 model. More detailed information on these transformation are found in [84]. It should be noted that, after the application of these transformations in the alternative, not only the communication but also some other elements of the initial case require modifications. For example, the format of Data File DF06_D will be based on “STEP XML” instead of “MS Visio Drawing” since the output format of the model transformation is based on the first one.

4.1. Methodology application

Having the necessary information on the design process and as well as the opinion of the expert process, the methodology can be applied.

Instantiation of the process model: For phase 1a.1 of the methodology (process modeling), a specific process model is instantiated from our meta-model of collaboration presented in Section 3.2. The process expert provides the information input. Technically, this is realized using EMF’s model editors, which uses XML as a serialization and exchange format. An excerpt of the tree-view of the generated model for the case initial version is shown in Fig. 10.
Validation of the process model: The instantiated process is validated using EMF’s model validation tool, which checks both ECORE structural constraints and additional OCL constraints.

Identification of the pairwise collaborations (PwCs): The process model (instantiated, validated and completed) is then studied in order to identify the PwCs (see Section 3.3). To this end, a zoom is necessary on the “Collaborative Work” class in the process model. As it is illustrated in Fig. 10, this class has only one instance: CW06-07 in the process model of the case. This shows that one PwC exists in the process.

Extraction of exchanged data formats: For CollaborativeWork CW06-07, studying the process model, the exchanged data files (produced, delivered and required) should be identified in order to extract their formats:
- **DF06_P** which is the produced data file. The format of this file is based on “MS Visio Drawing” and “SK2 XML” respectively in the initial case and the alternative.
- **DF06_D** which is the delivered data file. The format of this file is based on “MS Visio Drawing” and “STEP XML” respectively in the initial case and the alternative.

---

![Fig. 9. Partial design process in case study including initial and alternative processes.](image)

---

![Fig. 10. Excerpt of the tree-view of the process model instantiated in Eclipse for initial case.](image)
**Table 3**  
Results of syntax matching for initial and alternative cases.

<table>
<thead>
<tr>
<th>Input</th>
<th>Rule 1 (potential level)</th>
<th>Rule 2 (current level)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OCL</td>
<td>Manuel</td>
</tr>
<tr>
<td>Collaboration CW06-07 (initial case)</td>
<td>Unsatisfying (MS Visio drawing ≠ STEP XML)</td>
<td>Unsatisfying</td>
</tr>
<tr>
<td>Collaboration CW06-07 (alternative)</td>
<td>Unsatisfying (SK2 XML ≠ STEP XML)</td>
<td>Unsatisfying</td>
</tr>
</tbody>
</table>

**DF06_R** which is the required data file. The format of this file is based on “STEP XML” in the initial case and the alternative.

**Syntax matching of data formats:** The identified formats, in the unique PwC of the case, are used to define the syntax aspect of Data Adequacy (see Section 3.4) in collaboration CW06-07. This is performed via two syntax matching:

- Syntax matching at potential level indicating the compatibility between the formats of data files DF06_P (produced data) and DF06_R (required data).

  - Current level indicating the compatibility between the formats of data files DF06_P (produced data) and DF06_R (required data).

  - Given the size of this case study, both automatic and manual syntax matching approaches (see Section 3.4A) can be applied which provided the same results. These results are summarized in Table 3.

**Extraction of exchanged data abstractions:** After syntax matching, the exchanged files, identified previously for Collaborative Work CW06-07, are once again investigated in order to verify the availability of their abstractions. In the case, the following abstractions were available:

  - ADF06_P, which is the abstraction of produced data, file DF06_P. This abstraction is based on the meta-model of “MS Visio Drawing” and meta-model of “SK2 XML” (XSD) respectively in the initial and alternative cases.

  - ADF06_D, which is the abstraction of delivered data, file DF06_D. This abstraction is based on the meta-model of “MS Visio Drawing” and meta-model of “STEP XML” respectively in the initial and alternative cases.

  - ADF06_R, which is the abstraction of required data, file DF06_R. This abstraction is based on the meta-model of “MS Visio Drawing” and meta-model of “STEP XML” respectively in the initial and alternative cases.

**Semantic matching of data abstractions:** The above identified abstractions (available also as ECORE meta-models) were used in this step of methodology in order to increase the precision of interoperability measurement:

- Semantic matching at potential level indicating the compatibility between the abstractions of data files ADF06_P and ADF06_R, which respectively play the role of source 1 and target.

  - Semantic matching at current level indicating the compatibility between the abstractions of data files ADF06_D and ADF06_R which respectively play the role of source 2 and target.

- Given the size of the above abstractions, the automatic matching process was applied (see Section3.4B). This step starts with the selection of two abstractions, already imported in Eclipse as ECORE meta-models, as the source and target. Then using the EMF Compare tool, the results of the comparison are presented as structural differences between the abstractions.

- The results of semantic matching for the alternative case are included in Table 4. In this table, the designer requiring the data defines the threshold (T) based on the minimum data required for realizing his activity. In the case study, this variable is set to 0.5, which indicates that at least half of the required data should be delivered to the designer after the data exchange.

**4.2. Final results**

The results of syntax and semantic matching for both initial and alternative cases are summarized in Table 5. Considering the results indicated in this table, it is observed that in the initial case, the performance indicator has an unsatisfying value at potential and current levels. In the alternative case, the value of the indicator at the current level is improved. These observations can be detailed. Regarding the alternative case, a pre-developed model transformation is applied based on semantic relations between product models. Such transformation is capable of ensuring the semantic aspect of interoperability. Meanwhile, the transformed models must be based on a common language, mainly XMI or XML, to ensure the syntax aspects.

- In the case study, the alternative affects the potential value of syntax matching. In fact, this alternative affects not only communication, but also the choice of design tools. The model transformation requires changing the MS Visio tool to Eclipse model editor. Such change in a design process might be considered as a radical change. Therefore, the decision-making is important in this case, which requires further studies on the other criteria such as cost or time of implementation of model transformation.

**Table 4**  
Sample results of semantic matching for alternative case.

<table>
<thead>
<tr>
<th>Class property</th>
<th>Variable</th>
<th>Propriety of propriety</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>Attribute</td>
<td>Association</td>
<td></td>
</tr>
<tr>
<td>Elements in target ADF06_R</td>
<td>13</td>
<td>NC</td>
<td>47</td>
</tr>
<tr>
<td>Mismatches according to source 1 ADF06_P</td>
<td>13</td>
<td>NMC1</td>
<td>47</td>
</tr>
<tr>
<td>Mismatches according to source 2 ADF06_D</td>
<td>5</td>
<td>NMC2</td>
<td>14</td>
</tr>
<tr>
<td>Weights</td>
<td>1</td>
<td>WC</td>
<td>0.5</td>
</tr>
<tr>
<td>Threshold (T)=0.5</td>
<td>13 × 1+(47 + 16)×0.5 + 189 × 0.1 = 63.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight of targets elements (WE)</td>
<td>13 × 1+(47 + 16)×0.5 + 189 × 0.1 = 63.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight of mismatches in source 1 (WME1)</td>
<td>5 × 1+(14 + 5)×0.5 + 57 × 0.1 = 20.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight of mismatches in source 2 (WME2)</td>
<td>1 − 44/44 = 0 &lt; T → Unsatisfying</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic matching value (potential level)</td>
<td>1 − 44/44 = 0.68 &gt; T → Satisfying</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5
Summary of case study results.

<table>
<thead>
<tr>
<th>Phase 1a.1: Modeling</th>
<th>One pare-wise collaboration in the process is found, known by the class Collaborative Work CW06-07</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Phase 1a.2: Measurement (for collaboration CW06-07)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Syntax matching</th>
<th>Potential</th>
<th>Initial case</th>
<th>Alternative case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current</td>
<td>Unsatisfying</td>
<td>Satisfying</td>
</tr>
<tr>
<td>Semantic matching</td>
<td>Current</td>
<td>Unsatisfying</td>
<td>Satisfying</td>
</tr>
<tr>
<td>Performance of collaboration</td>
<td>Potential</td>
<td>Unsatisfying</td>
<td>Unsatisfying</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>Unsatisfying</td>
<td>Satisfying</td>
</tr>
</tbody>
</table>

Results are provided from Table 3.

Results are provided from Table 4.

5. Conclusions

In this paper, it is considered that collaboration between designers is realized through the exchange of data files. In such collaborations, a low level of performance can cause problems of data adequacy that can potentially cause failures in design products. Regarding the evaluation of the collaborations’ performance, an approach is still missing that studies such performance in terms of both conceptual and technical barriers of interoperability since analyzing the collaborations as white boxes. Therefore, a tooled methodology is proposed in this paper to fill the gap identified based on the literature review. The innovations and advantages of this methodology, illustrated by this case study, are primarily:

- The adaptability of the proposed methodology to collaborative environments might be its main advantage since its elements are based on the state of the art of the concepts of collaboration and interoperability. In fact, the methodology is based on these concepts to distinguish the collaboration performance from the performance of individual activities and from the performance evaluations based on classical criteria such as cost or time. This reduces the risk of interference of parameters that are outside the collaboration scope.
- For example, in the modeling phase (phase 1a.1), among all process elements, the only elements that affect the collaboration performance are preserved. Moreover, in the measuring phase (phase 1a.2), the indicator is derived from the concept of interoperability.
- The automation is increased in several ways in this methodology.
- For example, in the modeling phase, using the meta-model of collaboration developed in Eclipse, it is possible to generate automatically an initial instance of the process model. The validation of the syntax of this model can also be done automatically. Moreover, in the measuring phase, using OCL and EMF Compare tool, respectively, the defined syntax and semantic matching rules can be verified. On one hand, these examples of automation can reduce the risk of human error or judgments in applying the application of methodology and on the other hand, they can reduce the execution time.

Integrating technical and conceptual aspects of interoperability: Collaboration performance is evaluated based on interoperability at two levels: technical and conceptual. Considering the complementary role of these levels to achieve interoperability in the exchanged data, in the methodology both aspects are integrated in the measurement phase.

Interaction with stakeholders: several interactions between the user of the methodology and stakeholders (designers, process management) are considered in the measurement phase to have a better translation of the results based on objective of the process. For example, the threshold of satisfaction in data adequacy is based on the opinion of the designers. Indeed, the designer defines to what extent the delivered data are adequate for performing its activities.

Regarding the perspectives of this paper, it is primarily suggested to study the applicability of the proposed methodology for its potential users. This methodology requires the availability of technical arrangements such as Eclipse installations and basic knowledge of data modeling. In addition, the intelligibility of tasks and results for users must be studied. This allows defining the level of user autonomy in performing the various tasks. As long as the key feature of this methodology, its approach, is respected, it is possible to evolve or upgrade the current choice of tools or methods. Such modifications can be done based on the case under evaluation. This might also increase the accuracy of results. Currently in the modeling phase (Phase 1a.1), the granularity is limited to pairwise collaborations. On one hand, collaboration between multiple actors with multiple data exchanges must be studied. On the other hand, the impact of performance of each data exchange on the overall value of the process performance must be investigated. Regarding the measuring phase (Phase 1a.2), the similarity functions in matching can be replaced by more advanced functions [67,85]. For semantic matching via EMF Compare tool, an evolution may be also considered by integration of expert opinion or domain ontology in the matching process [86].

Acknowledgment

This research paper is a result of the corresponding author’s PhD thesis, performed at LCFC and LIS demonstrator of Ecole Nationale Superieure d’Arts et Metiers (ENSAM - Arts et Metiers ParisTech) France. This work has been funded by ENSAM.

References

W.I.M.Y.W. operability,'', plans:

Communications 58 (7 (Sep)) (2007) 687–697.


K. Husted, S. Michalova, Dual allegiance and knowledge sharing in Inter-firm R&D collaborations, Organizational Dynamics 39 (1 (Jan)) 2010 37–47.


Amir Pirayesh Neghab is currently a research engineer in the International Virtual Laboratory for Enterprise Interoperability (InterOP-VLab), Talence, France. He obtained his PhD in industrial and mechanical engineering from Ecole Nationale Supérieure d’Arts et Métiers, France in 2014. His research interests include Interoperability Evaluation, Performance Measurement, Risk Management, Data Modeling and Simulation, in product life cycle.

Alain Etienne is an associate professor of computer and industrial engineering in the Laboratory of Design, Manufacturing and Control (LCFC) at Ecole Nationale Supérieure d’Arts et Métiers, France. He obtained his PhD in industrial engineering from University of Nancy, France in 2007. His research interests include artificial intelligence, information system, risk management and knowledge formalization applied to decision-making, design of both products and manufacturing systems.

Mathias Kleiner is an associate professor of computer science in the Laboratory of Science, Information and Systems (LISIS) at Ecole Nationale Supérieure d’Arts et Métiers, France. He obtained his PhD from University of Méditerranée, Marseille, France in 2007. His research interests include artificial intelligence, graphs, constraint programming and metamodeling, and the application of these technics to mechatronic products design.

Lionel Roucoules is currently professor at Arts et Métiers ParisTech Institute (France - www.ensam.eu) since September 2008. He develops his research in the Information and System Science Laboratory (www.lsis.org). The context of his research is integrated design and collaborative IT platform in a global PLM vision. His specific interest is product-process interface. He proposed a DFM-synthesis approach, which is now part of larger DFX by least commitment modeling concepts supported by a Model Driven Engineering platform for Virtual Product Development.