POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'informazione

Dipartimento di Elettronica, Informazione e Bioingegneria

Master of Science in

Computer Science and Engineering

The notion of relevance: Modeling for an exploratory approach

Supervisor:

PROF. PAOLO PAOLINI

Master Graduation Thesis by:

DARIO NESI

Student Id n. 813493

Academic Year 2014-2015

POLITECNICO DI MILANO
Scuola di Ingegneria Industriale e dell'informazione
Dipartimento di Elettronica, Informazione e Bioingegneria
Corso di Laurea Magistrale in
Ingegneria Informatica

La nozione di rilevanza: Modello per un approccio esplorativo

Relatore:
PROF. PAOLO PAOLINI

Tesi di Laurea Magistrale di:
DARIO NESI
Matricola n. 813493

Academic Year 2014-2015

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

The template has been adapted by Emanuele Mason, Andrea Cominola and Daniela Anghileri: *A template for master thesis at DEIB*, June 2015

Final Version as of November 29, 2015 (classicthesis version 1.0).

To my mother Donatella and my father Paolo

RINGRAZIAMENTI

In queste righe voglio porgere i miei ringraziamenti alle persone che hanno reso possibile questo lavoro di tesi, ed intendo farlo in italiano.

Desidero ringraziare il Prof. Paolo Paolini, per l'opportunità offertami di lavorare a questo progetto e per i suoi suggerimenti.

Desidero ringraziare Diletta ed Elena, che talvolta si sono adoperate per darmi i dati o i chiarimenti che mi erano necessari riguardo al progetto EDOC.

Desidero ringraziare lo staff di HOC-Lab, che mi ha fornito uno spazio di lavoro in un contesto a stretto contatto con uno dei casi di studio.

Desidero ringraziare di cuore i miei genitori, che mi hanno permesso di studiare al Politecnico di Milano, lontano da casa, fornendomi supporto morale e materiale durante tutta la magistrale.

Desidero ringraziare Irene, al pari di quanto fatto verso la mia famiglia, in quanto senza ombra di dubbio ha permesso a questo lavoro di concretizzarsi fino a questo testo. Il suo aiuto non è mai mancato, sia nel rivedere le mie riflessioni su questa ricerca, sia nell'affrontare gli alti e bassi del mio periodo a Milano (ma in particolare di questi ultimi giorni).

CONTENTS

Abstract xiii
Preface xv
1 INTRODUCTION 1
1.1 Relevance nuances 3
1.1.1 Historical relevance 3
1.1.2 Geographical relevance 3
1.1.3 Political relevance 3
1.1.4 Emotional relevance 4
1.1.5 Strategic relevance 4
1.1.6 Popularity relevance 4
1.1.7 Qualitative relevance 4
1.1.8 Quantitative relevance 5
2 STATE OF THE ART OF FUZZY CONCEPTUAL MODELING 7
2.1 Introduction 7
2.1.1 Imperfect information 8
2.1.2 Basic concepts of fuzzy data 9
2.2 The road to Fuzzy database conceptual models 11
2.2.1 Level of fuzziness 11
2.2.2 Fuzzy logic concepts in EER 12
2.2.3 Fuzzy object oriented databases 13
2.3 Fuzzy UML conceptual model 14
2.3.1 Fuzzy class 15
2.3.2 Fuzzy generalization 18
2.3.3 Fuzzy aggregation 21
2.3.4 Fuzzy association 22
2.3.5 Fuzzy dependency 23
2.3.6 A complete example 25
2.4 Other notable works 25
3 CPN RELEVANCE MODEL 27
3.1 Introduction 27
3.2 The model 28
3.2.1 Characterizing 28
3.2.2 Prevalence 29
3.2.3 Noteworthy 30
3.3 Formal definition 31
3.3.1 Horizontal, Vertical and Mixed functions 32

4 MUSEUM OF MUSICAL INSTRUMENTS 35
4.1 Introduction 35
4.2 Keyboard Instruments 36
4.3 Facet analysis 37
4.4 CPN Model implementation 38
4.4.1 Data model 38
4.4.2 Function implementations and examples 40
5 EDOC@WORK 45
5.1 Introduction 45
5.2 Educational experiences 46
5.3 Facet analysis 47
5.4 CPN Model Implementation 48
5.4.1 Data model 49
5.4.2 Function implementations and examples 51
6 CONCLUSIONS AND FUTURE WORK 57
6.1 Doubts about this work 59
6.2 Future work 61

BIBLIOGRAPHY 63

i APPENDIX 67
A CHECKLIST TO IMPLEMENT CPN RELEVANCE MODEL 69
A.1 Introduction 69
A.2 The goal 69
A.3 Having the right domain 69
A.4 Relevance representation and value ranges 71
A.5 Homogeneous facets 72
A.6 Keeping relevance information and contextual data close 72
A.7 Another feasible CPN-enhanced context 73
B EDOC@WORK REPOSITORY 75
B.1 Differences 75
C EDOC DATA 77
D MUSEUM DATA 85

LIST OF FIGURES

Figure 2.1 Fuzzy class 17
Figure 2.2 Fuzzy generalization 20
Figure 2.3 Fuzzy aggregation 22
Figure 2.4 Fuzzy association 24
Figure 2.5 Fuzzy dependency 24
Figure 2.6 Fuzzy UML model 25
Figure 3.1 Iron Crown of Lombardy 29
Figure 4.1 Keyboard instruments in the case study 36
Figure 4.2 Example of Museum CPN Data 39
Figure 4.3 Example of Museum horizontal information 42
Figure 4.4 Example of Museum vertical information 43
Figure 5.1 Example of EDOC CPN Data 50
Figure 5.2 Example of EDOC horizontal information 53
Figure 5.3 Example of EDOC vertical information 54
Figure 6.1 Comparison between measures on CPN data model and the boolean one 58
Figure 6.2 Most relevant instruments in a simulation for testing purposes 60
Figure C.1 EDOC repository data (1/8) 78
Figure C.2 EDOC repository data (2/8) 79
Figure C.3 EDOC repository data (3/8) 80
Figure C.4 EDOC repository data (4/8) 81
Figure C.5 EDOC repository data (5-6/8) 82
Figure C.6 EDOC repository data (7-8/8) 83
Figure D.1 Museum repository data (1/6) 86
Figure D.2 Museum repository data (2/6) 87
Figure D.3 Museum repository data (3/6) 88
Figure D.4 Museum repository data (4/6) 89
Figure D.5 Museum repository data (5/6) 90

Figure D.6 Museum repository data (6/6) 91

LIST OF TABLES

Table 6.1 Relevance function types 59
Table 6.2 Comparison CPN vs Boolean 62
Table B.1 Comparison between EDOC Repository version 76

ABSTRACT

This thesis is about the exploration of "rich" datasets, where by "exploration" user sessions where the goal can vary from sense-making to deep understanding (rather than specific item search) and by "rich" datasets items characterized by a multi-faceted semantics are meant. The purpose of this work is to investigate how the notion of "relevance" can support deeper and more effective exploratory sessions.

After a thorough literature review on fuzzy conceptual modeling (the mechanics of which is the closest to this work), different interpretations of the notion of relevance have been identified, leading to the "CPN (Characterization, Prevalence and Noteworthiness) model", composed by data structures and abstract functions that can be instantiated to collect and measure relevance of items and attributes (treated as Fuzzy ones) using different relevance interpretations.

The model has been tested against two real-life case studies: one is the dataset of the Museum of Musical Instruments at Castello Sforzesco, the other is a set of experiences' reports collected within the frame of a national research project (EDOC), where the impact of technology on education has been surveyed. For the Museum case, 75 keyboard instruments have been analyzed, with a focus on the materials in the active and the passive parts of the instruments: over 2000 relevance values have been identified. For the EDOC case, 113 educational experiences have been analyzed, for 21 facet-values in 3 different relevance interpretations. Data from both case studies have been used in a simulated environment to test the capability of the model to support sense-making and understanding.

Test results show how the several types of data readings from CPN-enhanced repositories can provide more accurate insights than the ones from data models where "object X has/doesn't have attribute Y" only holds. This has also been confirmed by field experts that analyzed data provided by the redesigned database.

Future works would focus on the aspect of the implementation of algorithms in exploratory portals and on the user interface, that should provide a non trivial mass of information to start and not disrupt the user dialog.

SOMMARIO

Questa tesi si colloca nell'ambito dell'esplorazione dei "rich datasets", dove con esplorazione si intende una sessione utente in cui l'obiettivo spazia dall'intuire all'approfondire la conoscenza (piuttosto che specificatamente ricercare) e per "rich datasets" si intende una collezione di dati caratterizzati da una semantica multi-sfaccettata. L'obiettivo è indagare come la nozione di "rilevanza" supporti approfondite ed efficaci sessioni esplorative.

Dopo un esteso studio sul Fuzzy Conceptual Modeling (le cui meccaniche assomigliano a quelle usate qui), sono state identificate diverse interpretazioni del concetto di rilevanza, che sono confluite nel modello "CPN (Characterization, Prevalence, Noteworthy)", le cui strutture dati e funzioni astratte possono essere usate per raccogliere e misurare la rilevanza di oggetti e attributi (trattati come Fuzzy) secondo le diverse interpretazioni.

Il modello è stato testato in due casi di studio reali: uno rappresentato dalla collezione del Museo degli Strumenti Musicali del Castello Sforzesco, l'altro dalle esperienze didattiche svolte nel progetto di ricerca nazionale EDOC, che ha analizzato l'impatto della tecnologia sull'educazione. Nel caso del Museo sono stati analizzati 75 strumenti a tastiera, considerando i materiali delle parti attive e passive dello strumento: oltre 2000 valori di rilevanza sono stati identificati. Per il caso EDOC sono state analizzate 113 esperienze, con 21 sfaccettature possibili in 3 differenti interpretazioni di rilevanza. I dati di entrambi i casi sono stati usati in simulazioni per verificare la capacità del modello di supportare l'esplorazione dei dati.

I risultati mostrano come le letture dei dati supportati da CPN possano fornire informazioni più accurate rispetto a quelle ottenute da modelli dove un attributo Y può appartenere ad un oggetto X in maniera booleana. Le conferme arrivano da esperti del settore che hanno osservato i dati provenienti dai database riprogettati.

I lavori futuri si concentreranno sull'implementazione degli algoritmi nei portali esplorativi e sull'interfaccia utente, che deve fornire una notevole quantità di dati per iniziare ed alimentare il dialogo con l'utente.

PREFACE

MODELING is the process of simplifying reality in order to make a representation that covers the key aspects of interest.

Implementing relevance information, as it has been done in this thesis, means dramatically change the paradigm of modeling in order to de-simplify models to extract more knowledge, always keeping the process sound and structured.

In fact, adding information just for the sake of completeness doesn't always lead to better results than what achievable with a well-designed and purpose driven artifact.

Below, an extract from *On Exactitude in Science* by Jorge Luis Borges.

> ...In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography....
>
> *(Suárez Miranda, Viajes de varones prudentes, libro IV, cap. XIV, Lérida, 1658)*

1 INTRODUCTION

As native language speaker, in natural language conversations the concept of relevance of one or more objects or people appears as intuitive, but trying to address the topic from a more analytical perspective issues start arising. The multiple meanings that term "relevant" can assume, in fact, could lead to ambiguous and uncertain interpretations that reflects on how data is treated in human-computer interactions. As example, people unconsciously use several relevance criteria in ordinary statements and evaluations like: historical relevance, geographical relevance, political relevance, emotional relevance, strategic relevance, popularity relevance, quality relevance or economical relevance. In a context where a data repository aims to be accessed from human beings, not treating relevance information can lead to heavy disadvantages.

Without treating relevance criteria, in fact, a request to a large data set could return too much information, causing an overload and data dispersion such that potentially useful results are going to be missed by the user, who therefore, as an instinctive reply, can try to enforce stricter filters losing the overall vision. Examples of such context are cultural heritage repositories, touristic databases or educational databases, where too much rough simplifications can make information inaccessible and data is for its nature greatly nuanced.

Let's take as first example an interrogation to a repository for archaeological tourism, where it's asked to retrieve all the objects classified as Etruscan. On the territory, museums and historical collections contains objects of almost all epochs and periods and some of them can be related to the Etruscan civilization for several aspects. Having the repository capable of handling relevance information enables the system to return results that have certain a degree of pertinence to the request (in this case the degree of pertinence represents a measure of being Etruscan), and a justification of why objects are returned, like the finding site or manufacturing techniques. The systems can then answer to the user with a more human-understandable information, which can be seen from a multitude of different perspectives and be the base for a human-machine dialog.

Another example comes from the world of educational resources collections and archives, where educational experiences are collected and classified forming a rich and variegated data set. Experience profiles contain data about the impact of technologies and methodologies on the daily learning of students from primary school to high school. Experiences can then be characterized by one or more feature or results, and they can be rated by considering their performance or quality. It can be stated that for a given experience, the inclusive factor observed is more relevant than the technology used, or that the pedagogical structure of the experience has played a role so important that can be seen as a characterizing aspect. Without the database considering relevance information, but only Boolean values about presence of the various experience's features, a lot of data assume the same level of significance and the user can't have a clear vision of the spectrum of possibilities and their impacts (key information in educational scenarios).

This work does not address a specific context of relevance, though two case studies are used to demonstrate the validity of the thesis. It aims to establish and provide a classification and a framework to deal with relevance information and relevance-capable databases. It's not among the aims of this project to establish why and how objects can assume a certain degree of relevance, even if to get to the reported conclusions several different kind of contexts has been analyzed. That, in fact, would require deep knowledge of the specific fields, so data for testing purposes is given by experts in the form of museum catalogue and appraisals portfolio.

This thesis is composed as follows: in Chapter 2 an introduction about Fuzzy Conceptual Modeling is given, with the current state of the art of the field and the relevant works that led to the current technologies; in Chapter 3 a description of the theoretical modeling of relevance information is given, with formalization and description of data structures and abstract functions; Chapter 4 and Chapter 5 are two chapters containing the case studies, one for the Museum of Musical Instrument of Castello Sforzesco in Milan and one for the EDOC@Work project, providing examples about the implementation of abstract relevance information defined and its impact on the usability of data; finally in Chapter 6 results and future improvements are reported.

This chapter continues with condensed descriptions of the relevance criteria listed above, in order to make the reader to appreciate

the various, often implicit, meaning that the statement *This object is relevant* can assume.

1.1 RELEVANCE NUANCES

Studying two different cases and brainstorming around a lot of the possible statements that can be expressed concerning domain data, I've identified different declinations of the concept of relevance. In the following subsections all the interpretations are explained in detail with the characterizing aspect of each category and clarifying examples. It appears correct that more than one interpretations of relevance can hold at the same time for the same object in the same context.

1.1.1 *Historical relevance*

An object can have historical relevance when it is involved in a relevant historical event as trigger, participant or direct and live representation.

For example, the weapon used to stab the Empress Elisabeth "Sisi" of Austria (exposed at the Hofburg Palace in Vienna), or the Tomb of Pharaoh Tutankhamun (discovered in 1922), have a (different) degree of historical relevance.

1.1.2 *Geographical relevance*

An object can have a geographical relevance when its relative position on earth is used as a landmark.

For example, the Greenwich Meridian in the United Kingdom or the Pole star show geographical relevance.

1.1.3 *Political relevance*

An object can have a political relevance when it plays a crucial role in the political scenario of one or more country. An object is not necessarily an inanimate object, also a human being or an event can have political relevance.

For example, the Russian Revolution in 1917 or the anti-apartheid revolutionary Nelson Rolihlahla Mandela have political relevance in certain contexts.

1.1.4 *Emotional relevance*

An object can have an emotional degree of relevance when for one or more human being feels an emotive attachment of some sort to it.

For example, for a toddler a Teddy Bear or a blanket can have emotional relevance, when for adults can be a motorcycle or some photographs.

1.1.5 *Strategic relevance*

An object can have strategic relevance when it represents a leverage point in the accomplishing of a purpose.

For example, in the game of Monopoly the property of Park lane, when already owning the Mayfair property space, can have a strategic relevance in order to win the game.

1.1.6 *Popularity relevance*

When a large majority of people agrees on the same opinion about a certain aspect of an object, this can assume a popularity relevance degree (that can even be completely misleading).

For example, in the recipe for the green peppercorn fillet the presence of the pepper is perceived as characterizing even with a low (absolute and relative) presence in the dish and independently from the quality of the peppercorns. Another example is the Iron Crown of Lombardy, that ironically has no iron in it but it has always been known with that name and with that is addressed.

1.1.7 *Qualitative relevance*

An object's qualitative relevance degree can be determined by the degree to which a set of its inherent characteristics fulfills requirements.

For example, an object can have a high quality degree depending of its conservation state, a room can be rated according to its acoustic, or a movie's quality can be determined by film criticism. Example

of qualitative relevance value are "good", or "excellent" or "poor". The quality then can determine the qualitative relevance degree. It's perhaps the most intuitive interpretation of relevance, since it can support hypothesis as "A is better than B", but it should not be mistaken with the other kind of relevance ("The event E is better than the event X" has not the same meaning of "The event E is more historically relevant than the event X").

1.1.8 *Quantitative relevance*

Quantitative relevance is a kind of relevance strictly number-driven. Measurable properties of objects can assign them a quantitative relevance degree. The most common properties of objects that arose during the development of this thesis are:

- Economical value
- Size
- Weight
- Multiplicity
- Percentage

These four kinds can further be decomposed in more specific properties semantic, like the Economical that can be intended as

- Production/Transformation cost
- Trade value

It's common knowledge that an object can be very cheaply produced but highly costly recycled, so for it the transformation cost assume a different relevance rather than the production cost or the trade value. Also the size can be expanded, with for examples the usually considered dimensions like:

- Height
- Width
- Depth

For example, considering the height of several building as the interpretation of relevance, the Empire State Building has a high relevance

degree, while the same can't be said for Palazzo Lombardia in Milan (the regional headquarter building) that high only 161 meters (this is true if higher importance is given to taller building, note that can be true the exact opposite).

The multiplicity relevance is, instead, to be intended as the number of occurrences in a certain context of the object of which the relevance is considered.

For example, the number of screws used to build the Eiffel Tower in Paris makes the screw object (assuming there is no distinction between the single pieces) have a certain degree of relevance with respect to the Tower itself or another context. It's not specified that with the growing of the multiplicity also the relevance degree increases: an object can be highly relevant because its multiplicity is very low (this is also seen as rarity of an object, or uniqueness).

With a similar meaning to the multiplicity relevance, the percentage relevance indicates that the relevance degree of an object is based on a percentage value that relates to the object itself. For example, in a leather suitcase the leather can be a relevant material because present in high percentage.

An object can have a quantitative relevance not only when the considered measures is plainly the highest or the lowest but also in case the measure differs from an average value for a given set. In generally a function can be applied to a measure to derive another value that express the relevance degree of the object.

2

STATE OF THE ART OF FUZZY CONCEPTUAL MODELING

2.1 INTRODUCTION

Fuzzy is a keyword associated with the mathematical domain, and later with the information technology, since 1965 when Lotfi A. Zadeh worked on a new theory regarding partial membership of an element to a given set[25]. Since then, a lot of consensus has grown around the peculiar idea of using Fuzzy theory to handle imprecise information[19] and hence to bring models closer to the real world. Modeling the real world means defining way to structure information and way to act in certain real context, defining behaviors and rules (for the machines) to follow in order to accomplish designing goals.

With conceptual data modeling are intended models with a level of abstraction that enables to be independent from the system implementation[1]. As known, from conceptual models lower level models like logical models can be derived automatically, so it appears to be essential to engineer solutions on the more flexible scale provided by conceptual models to produce complete and sound representations of the real world.

The introduction of Fuzzy logic in modeling has brought innovation also in conceptual modeling frameworks during time, always accordingly with the current most advanced data representation techniques. First the Entity-Relationship (ER) model[5] was revisited by Zvieli and Chen in [27], then its extended version (Extended ER) by Ma et al. in [17] and the Object Oriented data model in [2, 16]. At the moment of writing, the most advanced and complete framework for Fuzzy conceptual modeling is Fuzzy UML by Ma et al. [14], which combines the rich expressiveness of UML with Fuzzy logic. Besides from the main works, a notable paper can be found in literature from Bouaziz, Rafik et al. [4] focused on introducing Fuzzy logic in the Unified Semantic Model (a "competitor" of ER/EER) to develop a Fuzzy Semantic Model (FSM). The key concepts of this last one are reported in section 2.4, apart from the chronological coverage of the topic.

A distinction should be introduced between *Fuzzy data*, to which this work relates, and *Fuzzy behavior*. These two terms are used to group two major set, found from a literature recognition. It's in fact possible to find usages of "Fuzzy" term with a different semantic: the one concerning data representation (Fuzzy data) and the one regarding techniques of artificial intelligence (Fuzzy behaviour). Therefore, let assume that under the term Fuzzy data are involved all the techniques and theories concerning the structuring and handling of information with the purpose of storing or manipulating knowledge. Fuzzy behavior, instead, is the collection of all the strategies and ideas implemented into devices in order to make them able to interact with the real world and its tolerances and imprecision. Fuzzy behavior is then the part of processing input data and react considering a degree of uncertainty or imprecision. Hence there is a sharp division between the concepts of fuzzy data, which deals with representations of information, and fuzzy behavior, which concerns how to (automatically) use imprecise data.

In the rest of the chapter, several ways to represent concepts and real world objects as Fuzzy data are presented. First of all a description of the *imperfect information* concept is given below, then *Fuzzy sets* are introduced. The latter are an enhanced version of the mathematical sets, where partial membership of elements is enabled[25]. Following, the progressively injection of fuzzy in the database modeling techniques is described chronologically. The current best tool to model real world contexts is represented, in my opinion, by Fuzzy UML[14], which is explained in detail to conclude the chapter.

2.1.1 *Imperfect information*

Bosc and Prade [3] in a work of 1993 classified various possible kinds of *imperfect information*. They listed 5 possible type of information, each one with a clear and distinctive meaning:

A. *inconsistent information* groups data that cannot be reconciled because several contradictory occurrences exist at the same time (i.e. George's age is known to be 34 and 37 simultaneously); it's often caused by processes of information integration [6]

B. *imprecise information* represents information whose imprecision is related to the value of an attribute which is not a single value

but an interval or a set; this forces to make a choice when processing data (i.e. George's age is 34, 35, 36, 37)

c. *vague information* related too to the value of an attribute, cannot be represented as a proper set or interval but that can be a linguistic value [26] (i.e. George's height is *tall*)

d. *uncertain information* represents information that should be considered with a specific degree of truth (John's age is 35 with a degree of truth of 85%)

e. *ambiguous information* indicates that some element of the model is completely free of semantics and can raise several possible interpretations (i.e. it may not be clear if one person's salaries stated are per week or per month).

It's possible and likely that in the same model several kinds of imperfect information coexists together. An example of imprecise and uncertain information is John's age, that is 30 years old with a confidence of 0.9, 31 years old with 0.85 and 32 years old with 0.7. Ambiguous or inconsistent information is often avoided by methodical engineering and designing processes, instead to treat imprecise, vague and uncertain information Fuzzy sets looks to be the perfect tool to quantify imprecision and uncertainty.

2.1.2 *Basic concepts of fuzzy data*

Referring to a Fuzzy set means referring to an upgraded version of the classic concept of mathematical set, with the capability of including partially its elements in such a measure defined as *degree of membership*, that ranges in real numbers from 0 to 1 (included).

Let assume that an individual named John is 30 years old. Trying to fit John in a group of people labeled as "Young" raises the question of how much John can be considered young. If people are commonly said to be progressively less young reaching their 40s, it follows that John is not exactly young but he can be rated as 0.75 young. So his degree of membership to the group of young people is 0.75 (it's enough to have one but also other in the group can be not fully young members and have their degree of membership < 1), here having a non complete membership brings to a fuzzy set: a more tailored representation of the reality.

Tackling the issue of imprecise and uncertain information, another point of view of the example above can be drawn. Assuming that John's age is only vaguely known, a Fuzzy set can be used to describe such information. Several integers representing John's age are associated with a certain degree of confidence as i.e. he is 30 years old with a confidence of 0.9, 31 years old with 0.85 and 32 years old with 0.7. In this way the information about the age is imprecise, although the vagueness is described exactly and henceforth can be later processed.

In a more structured and formal way, provided a *Universe of discourse*[1] with the powerful concept of Fuzzy set, it's possible to introduce the idea of *Fuzzy Value*, which is a fuzzy set F in the universe just mentioned with a defined membership function $\mu F : U \to [0, 1]$ where $\mu F(u)$ for $u \in U$ denotes the degree of membership in the fuzzy set $F = \mu F(u1)/u1$.

The concept of fuzzy value easily fits for the example of John's vague age above.

When the membership function $\mu F(u)$ is considered as to be a measure of the possibility that a variable X has the value u in this approach, where X takes values in U, a fuzzy value can be described by a possibility distribution π_X[20]

$$\pi_X = \pi_X(u_1)/u_1$$

where $\pi_X(u_i), u_i \in U$, denotes the possibility that u_i is true. Possibility distribution is, instead, more appropriate to support the example of John belonging to young people group described above.

A Fuzzy set can support the same operations of an ordinary set and for simplicity, an element with a degree of membership equals to zero just doesn't belong to the set[25].

Since a Fuzzy set is a generalization of an ordinary set, imposing that the membership degree of each element can be only 0 or 1 brings the context back to the ordinary theory of being in/being out in a classic boolean fashion.

1 The totality of facts, things, or ideas that are implied or assumed in a given discussion, argument, or discourse (Collins American English dictionary)

2.2 THE ROAD TO FUZZY DATABASE CONCEPTUAL MODELS

The need of representing and manipulating fuzzy information in real world models has led to the introduction of fuzzy logic in database models [2, 16, 17, 27].

A database modeling process can produce two kinds of models [13]: conceptual data models, as entity-relationship (ER) or UML models, and logical data models like relational databases (RDBs) or object-oriented databases (OODBs). It is common knowledge that a conceptual model can be automatically transformed through mapping techniques in its related logical model that can be later refined for performances improvements or other low level adjustments. Conceptual modeling, therefore, can provide the best support to generate the most complete and least expensive (if compared to other design tools) specification from the real world since its high level of abstraction and robust and consolidated framework.

2.2.1 *Level of fuzziness*

Since the early '80s many works focused on enhancing the current state of the art for data modeling providing fuzzy mechanism to store and handle data for a large class of application. The first concepts to be revisited by Zvieli and Chen[27] were basic ER concepts, in which fuzzy entity sets, fuzzy relationship sets and fuzzy attribute sets were introduced, in addition to fuzziness in entity and relationship occurrences and attribute values. In that improved model three different levels of fuzziness can be identified:

- a first level, where entities, relations and attributes can partially belong to the ER model
- a second level, where instances of entities and relationships can belong to their (respectively) entity and relationship with a membership degree
- a third level, that considers only attributes (of entities and relationship) which can assume fuzzy values, like an attribute "Age" can assume value "young"

From the point of view of the artifact that follows the design, the first level is ultimately shrunk down to a classic boolean fashion since

an entity or a relationship has or has not to appear in the implementation. Anyway the first level plays a crucial role in the design phase as well as in the maintaining one.

Formally, let E, R, and A be the fuzzy entity type set, fuzzy relationship type set, and fuzzy attribute set of the fuzzy ER model, respectively μ_E, μ_R and μ_A be their membership functions. Then

- for an entity type, say E_i, we have $\mu_E(E_i)/E_i$, where $\mu_E(E_i)$ is the degree of E_i belonging to E and $0 \leqslant \mu_E(E_i) \leqslant 1$,
- for a relationship type, say R_i, we have $\mu_R(R_i)/R_i$, where $\mu_R(R_i)$ is the degree of R_i belonging to R and $0 \leqslant \mu_R(R_i) \leqslant 1$, and
- for an attribute of entity type or relationship type, say A_i, we have $\mu_A(A_i)/A_i$, where $\mu_A(A_i)$ is the degree of A_i belonging to A and $0 \leqslant \mu_A(A_i) \leqslant 1$

2.2.2 *Fuzzy logic concepts in EER*

After ER has been extended into EER for supporting specialization/-generalization, category and aggregation, Ma et al.[17] worked to introduce a conceptual framework in this new context, always referring to the three levels of Zvieli and Chen[27]. The key point of the new fuzzy model capabilities is to keep imperfect and complex real world objects as close as possible to their representation. New constructs has been introduced:

FUZZY GENERALIZATION (SPECIALIZATION) that have meaning of grouping entity types with common features (or defining subclasss from an entity type according to a certain feature) but with the following constraint

$$\forall e \in U \quad \mu_{specialized\_entity}(e) \leqslant \mu_{generic\_entity}.$$

So, considering a fuzzy superclass E and its fuzzy subclasses $S1, S2, ..., Sn$ with memberships functions $\mu_E, \mu_{S1}, \mu_{S2}, ..., \mu_{Sn}$, follows that

$$\forall e \in E, \forall S \in \{S1, S2, ...n, Sn\}, \quad \mu_S(e) \leqslant \mu_E(e).$$

To make $S1, S2, .., Sn$ be a fuzzy total specialization of E it's enough that $\exists S \in \{S1, S2, .., Sn\}$ such that the previous expression is true. To have, instead, $\{S1, S2, .., Sn\}$ be a fuzzy partial

specialization of E should exists an element $e \in E$ such that for all $S \in \{S1, S2, .., Sn\}$ it's true that $\mu_E(0) > 0 \wedge \mu_S(e) = 0$

FUZZY CATEGORY that is a subclass of the union of the superclass with different entity types; let $\{S1, S2, .., Sn\}$ and E be fuzzy entity set with membership functions $\mu_{S1}, \mu_{S2}, ..., \mu_{Sn}$ and μ_E respectively, E is a fuzzy category if

$$\forall e \in E, \exists S \in S1, S2, .., Sn, \quad \mu_E(e) > 0 \wedge \mu_S(e) \geqslant \mu_E(e).$$

The main difference between a fuzzy category and a fuzzy subclass is that for the latter one the formula above must hold for all the superclasses and not just one (being S_i the set of superclasses and E not a category but a subclass).

AGGREGATION where a number of entity types, say $S1, S2, ..., Sn$, are aggregated to form an entity type, say E; if $\mu_{S1}, \mu_{S2}, ..., \mu_{Sn}$ and μ_E are the corresponding membership functions, then the following holds:

$$\exists e \in E, e_1 \in S1, e_2 \in S2, ...e_n \in Sn \quad \text{s.t.} \quad \mu_E(e) = \prod_{1 \leqslant i \leqslant n} \mu_{Si}(e_i)$$

2.2.3 *Fuzzy object oriented databases*

When Object-Oriented Databases started to play a more central role on the scene of academic research, another trail opened with papers concerning how to deal with fuzzy data in the new object-oriented approach. Even though Bordogna, Pasi et al. in [2] have specifically designed a logical database model (FOOD), from their work a conceptual framework can be extracted. The authors introduce an extension of the graph-based object model[12] in order to manage both crisp and imperfect information by using fuzzy set theory and possibility theory. The imperfection in data representation is modeled in FOOD in order to define the following aspects:

- Vague attribute values, that refer to vague information in the Bosc and Prade classification. They are linguistic terms drawn from the term set of a linguistic variable, which extends the domain of the attribute.
- Uncertain property and uncertain link relationships, that concerns about uncertain information [27] but with the degree of

truth replaced by the degree of uncertainty which represents the likelihood that the actual attribute value is different from the one specified (the opposite of the degree of truth).

- Strengthened property and strengthened link relationships, that don't express the uncertainty of the information but add more data to the specific property which is associated to (and can be deduced by the context).

  For example, the strengthened property "participate" don't indicate that a "project" may have a "staff" different form that specified, but that each specified member of the "staff" has a specific involvement degree in the "project".

- Fuzzy classes and fuzzy class hierarchies, intended only as fuzzy occurrences of classes

Even if their work was conceived as an extension of an object-oriented database management systems and their prototype lacked some fundamental features of complete applications, they achieved to make the first step toward a good fuzzy object-oriented database.

In [16], Ma et al. take the ideas and results from previous works in the field and clarify concepts that haven't still well defined as fuzzy object specialization and fuzzy object relationship introducing a generic model for object-oriented database. In the paper metrics and basic properties are defined, as well as key concepts which an object-oriented model must support and that will be elaborated in future works from the same authors to establish the current state of the art for conceptual modeling.

2.3 FUZZY UML CONCEPTUAL MODEL

Actually, the most advanced solution to obtain clear high level design specification and deal with imperfect information is represented by UML conceptual models, introduced by Ma et al in [14]. In their work, all previous concepts introduced by themselves and others are used to extend UML capabilities of modeling and representation. Fuzzy sets and possibility distribution introduce different levels of fuzziness into the class of the UML model, enabling the UML class diagrams to model fuzzy information. Furthermore, based on the corresponding graphical representation of the fuzzy UML data model, the paper describes the formal mapping of the fuzzy UML data model into the fuzzy object-oriented database model.

Before introducing Fuzzy logic into UML modeling, a brief formal description of a UML class diagram is given.

A simple UML class diagram is a tuple $D = (C, A, R, O)$, where

$C = \{c_1, c_2, ..., c_k\}$ is a finite set of classes,

$A = \{a_1, a_2, ..., a_l\}$ is a finite set of attributes,

$R = \{r_1, r_2, ..., r_m\}$ is a finite set of relationships, and

$O = \{o_1, o_2, ..., o_n\}$ is a finite set of objects.

Moreover:

- $R \subseteq C \times C$ is a binary relation that represents the generalization, aggregation, association or dependency (described in [22]).
- For $c_i \in C(1 \leqslant i \leqslant k)$, $A(c_i)$ represents a set of attributes of c_i. Clearly $A(c_i) \subseteq a_1, a_2, ..., a_l$, i.e. $A(c_i) \subseteq A$. For $a_j \in A(1 \leqslant j \leqslant l)$, $a_j(c_i)$ denotes the attribute a_j of class c_i. In the context of the given c_i, a_j is used instead of $a_j(c_i)$.
- For $c_i \in C(1 \leqslant i \leqslant k)$, $O(c_i)$ means a set of objects that c_i contains. $O(c_i) \subseteq o_1, o_2, ..., o_n$, i.e. $O(c_i) \subseteq O$. For $o_p \in O(1 \leqslant p \leqslant n)$, and $a_j \in A(1 \leqslant j \leqslant l)$, $o_p(c_i)$ denotes the object o_p of c_i, and $o_p(a_j(c_i))$ denotes the value of object o_p on attribute a_j. In the context of the given c_i, o_p is used instead of $o_p(c_i)$ and $o_p(a_j)$ is used instead of $o_p(a_j(c_i))$.

In the context of Fuzzy UML model, a class diagram is referred as a tuple $\tilde{D} = \{\tilde{C}, \tilde{A}, \tilde{R}, \tilde{O}\}$ where each element is a fuzzy set with the same semantic as for the simple UML class diagram above.

2.3.1 *Fuzzy class*

As the authors explain in [14],from a theoretical point of view a class can be considered as:

A. a collection of objects with similar features (extensional point of view, it basically concerns objects that are already "living")

B. a set of attributes with their own value domains (intensional point of view)

When a class is defined from the extensional point of view, if it's the case that objects have different degree of membership to the group, the class results to be a Fuzzy class. From the intensional point of view, instead, it's the domain of a class attribute that makes or not the class a fuzzy one. When defining a subclass, the new class is designed from the intensional perspective (it's a special case of point 2) and if the superclass is a Fuzzy class then also the specialized one is fuzzy.

Refferring to the three levels of fuzziness:

A. The first level evaluates the extent to which the class belongs to the data model as well as the fuzziness of the content in terms of class' attributes. At this point each class c_i has a degree of membership to the fuzzy set of classes $\tilde{C}$ defined by a membership function i.e. $\mu_{\tilde{C}}(c_i)$. The same reasoning holds for class attributes, that can partially belong to a class by a membership degree i.e. $\mu_{\tilde{A}}(a_i)$.

B. The second level of fuzziness evaluates the degree of membership of "living" objects to their class. An object $o_p(c_i)$ is fuzzy if contains at least one fuzzy attribute, and its degree of membership to the class c_i can be established as $\mu_{\tilde{O}}(o_p(c_i))$

C. The third level of fuzziness is related to the value of class attributes. An attribute in a class defines a value domain, when this domain is a fuzzy subset or a set of a fuzzy subset, the value of an attribute on an object, say $o_p(a_j(c_i))(1 \leqslant p \leqslant n)$, is a fuzzy one represented by a possibility distribution, say $\{(v_1, \pi(v_1)), (v_2, \pi(v_2)), ..., (v_q, \pi(v_q))\}$. Here, $\pi(v_s)(1 \leqslant q)$ denotes the possibility that $o_p(a_j(c_i))$ has value v_s.

The traditional UML class diagram is simply a special case of the Fuzzy UML class diagram, as shown below.

- For the first level of fuzziness, if a class c_i belongs to the diagram (and so to the set of classes C) its membership function shall be $\mu_{\tilde{C}}(c_i) = 1$. If a class doesn't appear in the diagram, its membership function is $\mu_{\tilde{C}}(c_i) = 0$ but since such a class wouldn't be in C either, it's pointless to specify a membership function. The same reasoning holds for class attributes, the membership function of attribute a_j shall be $\mu_{\tilde{A}}(a_j(c_i)) = 1$ if class c_i has an attribute a_j, otherwise it's just missing from the diagram.

- For the second level of fuzziness, an object o_p is instance of c_i and has $\mu_{\tilde{C}}(o_p(c_i)) = 1$ or $\mu_{\tilde{C}}(o_p(c_i)) = 0$
- For the third level of fuzziness, the attribute value of object o_p, $o_p(a_j(c_i))$, is a possibility distribution $\{\pi(v_1)/v_1, \pi(v_2)/v_2, ..., \pi(v_q)/v_q\}$, which means that the attribute has the value $v_s (1 \leqslant s \leqslant q)$ with probability $\pi(v_s)$, but without imprecision or uncertainty either $\pi(v_s) = 0$ or $\pi(v_s) = 1$. This leads the possibility distribution to a reduced crisp set and $o_p(a_j(c_i))$ (the attribute value) can be only a crisp value.

To model the first level of fuzziness, i.e., an attribute or a class with a membership degree, the attribute or class name should be followed by a pair of word WITH mem DEGREE, where $0 \leqslant mem \leqslant 1$, which is used to indicate the degree to which he attribute belongs to the classs or to which teh class belongs to the data model [23, 18]. For example "Employee WITH 0.6 DEGREE" and "Office Number WITH 0.8 DEGREE" are a lass and an attribute, respectivley, with the first level of fuzziness. Generally, an attribute or class will not be declared when its membership degree is 0. In addition, "WITH 1.0 DEGREE" can be omitted when the case. Note that attribute values may be fuzzy. To model the third level of fuzziness the keywork FUZZY is introduced and placed in front of the attribute name. For the second level of fuzziness, the membership degree to which an instance of the class belongs to the class must be indicated. For this purpose, an additional attribute named μ is introduced into the class to represent the membership degree of the instance to the class, with domain value $[0, 1]$.

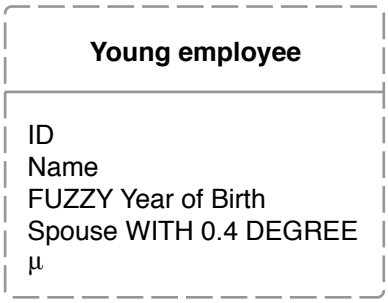


Figure 2.1: Fuzzy class

Graphically, a class with the second level of fuzziness is denoted by a rectangle with a dashed outline. Figure 2.1 shows the fuzzy class *Young Employee*. Here, the attribute *Year of birth* takes on fuzzy values; namely, its domain is fuzzy. It is unclear whether the class *Young em-*

ployee has the attribute *Spouse*, but it is known that youung employees have spouses with intermediate possibility, say 0.5 so the attribute is defined as uncertain. This class has fuzziness at the first level, and its membership degree is 0.9, in addition it's not possible to determine whether an object is an instance of the class because the class is fuzzy. Therefore, an additional attribute μ is introduced into the class for this purpose.

2.3.2 *Fuzzy generalization*

The inheritance mechanisms makes possible to define subclasses of a given class, letting the subclass obtain attributes and methods from its superclass. Because a subclass is a specialization of its superclass, every object belonging to the subclass must belong to the superclass either.

In the Fuzzy UML data model, classes may be fuzzy. A class produced from a fuzzy class by means of inheritance may itself be fuzzy. Therefore the relationship between two classes (general and specialized) can be fuzzy.

Since object instances of fuzzy class have a membership degree to their class, it's possible to leverage that information to determine the subclass/superclass relationship (after having established a threshold β).

Hence, the following characterizes the subclass(c_S)/superclass(c_G) relationship <u>from an extensional point of view</u>:

A. for any fuzzy object $o_p \in \tilde{O}$, $\mu_{\tilde{O}}(o_p(c_S)) \geqslant \beta$ (the object o_p belongs to the subclass c_S for more than the threshold) and

B. $\mu_{\tilde{O}}(o_p(c_S)) \leqslant \mu_{\tilde{O}}(o_p(c_G))$ (the object o_p belongs to the subclass less or equally than how belongs to the superclass)

The subclass is then a subclass of the superclass with a membership degree that is the minimum of the membership degrees to which these objects belong to the subclass.

Formally, let c_G and c_S be fuzzy classes, and β be a given threshold, then c_S is a subclass of c_G if

$$\forall o, \quad \beta \leqslant \mu_{c_S}(o) \leqslant \mu_{c_S}(o)$$

where $o \in \tilde{O}$ is an object instance of c_G and c_S and $\mu_{c_G}(o)$ and $\mu_{c_S}(o)$ are membership degrees of o to c_G and c_S respectively.

The membership degree of the subclass/superclass relationship is $\min_{\mu_{c_S}(o) \geqslant \beta}(\mu_{c_S}(o))$

Above is the case that c_G and c_S only have the second level of fuzziness. Therefore for classes with the second and the first level of fuzziness, from an extensional point of view, the subclass(c_S)/ superclass(c_G) relationship is characterized by:

$$\forall o, \quad (\beta \leqslant \mu_{c_S}(o) \leqslant \mu_{c_S}(o)) \wedge (\beta \leqslant mem_{C_S} \leqslant mem_{C_G})$$

where the classes are defined as follows:

- c_G WITH mem_{C_G} DEGREE
- c_S WITH mem_{C_S} DEGREE

When classes are defined from an intensional point of view, there is no object available, so to determine the subclass/superclass relationship the notion of *inclusion degree* must [15] be used. Ma et. al extended the inclusion degree to evaluate the membership degree of an object to a class and to further develop the relationship between fuzzy subclass and superclss.

Formally, let c_G and c_S be fuzzy classes and the degree to which c_S is the subclass of c_G be denoted by $\mu(c_G, c_S)$. For a given threshold β c_S is a subclass of c_G if

$$\mu(c_G, c_S) \geqslant \beta$$

where $\mu(c_G, c_S)$ evaluates the relationship between specialized class and general class according to the inclusion of the attribute domains of c_S with respect to the attribute domains of c_G as well as the weights of the attributes.

When c_G and c_S are classes with the first level of fuzziness and their membership degree is, respectively, mem_{c_G} and mem_{c_S} the formula above gets tuned to:

$$(\mu(c_G, c_S) \geqslant \beta) \wedge (\beta \leqslant mem_{C_S} \leqslant mem_{C_G})$$

(note the similarity with the extensional defined class case).

A critical issue in subclass/superclass hierarchies is the multiple inheritance of a class. Let c_S, c_{G1}, c_{G2} be fuzzy classes with the first and the second level of fuzziness and β be a given threshold. The fuzzy generalization relationship with multiple inheritances of class can be interpreted according to a conjunctive or disjunctive interpre-

tation. The two cases follows distinctly to highlight the key semantic points of the definitions, grouped by the tool used to characterize the subclass/superclass relationship.

When using the degree of membership of an object to a class, the conjunctive interpretation of multiple inheritance leads to say that c_S is a subclass of c_{G1} and c_{G2} if:

$$\begin{aligned} \forall o, \forall c, \quad & c \in \{c_{G1}, c_{G2}\} \\ & \wedge \beta \leqslant \mu_{c_S}(o) \leqslant \mu_c(o) \\ & \wedge \beta \leqslant mem_{C_S} \leqslant \min(mem_{C_{G1}}, mem_{C_{G2}}) \end{aligned}$$

for the disjunctive interpretation, instead, must hold:

$$\begin{aligned} \forall o, \exists c, \quad & c \in \{c_{G1}, c_{G2}\} \\ & \wedge \beta \leqslant \mu_{c_S}(o) \leqslant \mu_c(o) \\ & \wedge \beta \leqslant mem_{C_S} \leqslant \min(mem_{C_{G1}}, mem_{C_{G2}}) \end{aligned}$$

If using the inclusion degree to determine the relationship between subclass and superclasses, in the conjunctive interpretation c_S is a subclass of c_{G1} and c_{G2} if:

$$\begin{aligned} (\mu(c_{G1}, c_S) \geqslant \beta) \wedge (\mu(c_{G2}, c_S) \geqslant \beta) \\ \wedge (\beta \leqslant mem_{C_S} \leqslant mem_{C_{G1}}) \\ \wedge (\beta \leqslant mem_{C_S} \leqslant mem_{C_{G2}}) \end{aligned}$$

for the disjunctive interpretation, instead, must hold:

$$\begin{aligned} (\mu(c_{G1}, c_S) \geqslant \beta \wedge \beta \leqslant mem_{C_S} \leqslant mem_{C_{G1}}) \\ \vee (\mu(c_{G2}, c_S) \geqslant \beta \wedge \beta \leqslant mem_{C_S} \leqslant mem_{C_{G2}}) \end{aligned}$$

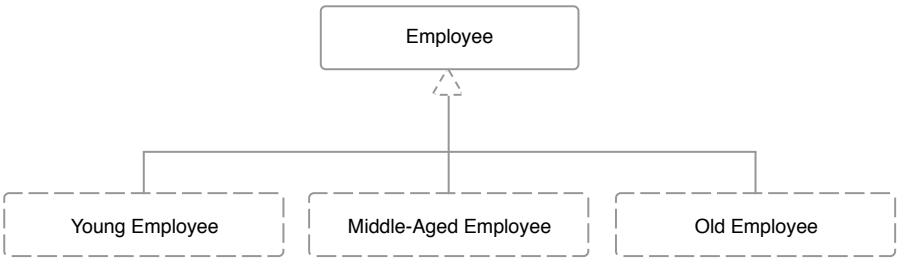


Figure 2.2: Fuzzy generalization

Graphically, a dashed triangular arrowhead is used to represent a fuzzy generalization. In Figure 2.2 the class *Young Employee*, *Middle-Aged Employee* and *Old Employee* all exhibit the second level of fuzzi-

ness, meaning that the classes may have some instances that belong to the classes with membership degrees. The two classes can be generalized into the class *Employee*.

2.3.3 *Fuzzy aggregation*

The aggregation concept express the whole-part relationship between an aggregate (the whole) and several constituent parts (the part), where the constituent part can exist independently[22]. Each instance of an aggregate can be projected into a set of instances of constituent parts. Formally, let c' be an aggregation of constituent parts c''_1, c''_2, ..., c''_n. For $o \in c'$, the projection of o to c''_i is denoted by $o \downarrow_{c''i}$ which represents an instance of c''_i. A class aggregated from fuzzy constituent parts may be fuzzy. If the former is still called an aggregate, the the aggregation is a fuzzy aggregation. Therefore each part belongs to the aggregation for a specific membership degree. The following method can be used to determine the fuzzy aggregation relationship, from an extensional point of view:

- For any fuzzy object, the membership degree to which it belongs to the aggregate is greater than or equal to a given threshold, and
- the membership degree to which it belongs to the aggregate is less than or equal to the membership degree to which its projection to each constituent part belong to the corresponding constituent part.

Then, the following holds

$$\forall o, o \in c' \wedge \beta \leqslant \mu_{c'}(o) \leqslant \min(\mu_{c''1}(o \downarrow_{c''1}), ..., \mu_{c''n}(o \downarrow_{c''n}))$$

when the fuzzy class c' is the aggregate of any group of fuzzy classes $c''_1, c''_2, ..., c''_c$. Here o is an object instance of c' and β a given threshold. The membership degree of c' when the class is an aggregation of the set $\{c''_1, c''_2, ..., c''_n\}$ is $\min_{\mu_{c''i}(e\downarrow_{c''i})} \mu_{c''i}(o \downarrow_{c''i})(1 \leqslant i \leqslant n)$. The idea behind the method above is that each object (instance of a fuzzy class) belongs primarily to its class rather than to a possible aggregation, this leads the membership degree of the object to the aggregation to be less or equal to the membership degree of the object to its class.

When classes $c', c''_1, c''_2, \ldots$ are classes with also the first level of fuzziness, and $mem_{c'}, mem_{c''_1}, mem_{c''_2}, \ldots$ are their membership degrees, the formula becomes:

$$\forall o, o \in c' \wedge \beta \leqslant \mu_{c'}(o) \leqslant \min(\mu_{c''1}(o \downarrow_{c''1}), \ldots, \mu_{c''n}(o \downarrow_{c''n})) \\ \wedge mem_{c'} \leqslant \min(mem_{c''_1}, mem_{c''_2}, \ldots)$$

From the intensional point of view, the inclusion degree must be used to determine the aggregation relationship between classes with first and second level of fuzziness:

$$\min(\mu(c''_1, c' \downarrow_{c''1}), \ldots, \mu(c''_n, c' \downarrow_{c''n})) \wedge mem_{c'} \leqslant \min(mem_{c''_1}, \ldots)$$

here $\mu(c''_i, c' \downarrow_{c''i})$ indicates the degree to which c''_i semantically includes $c' \downarrow_{c''n}$.

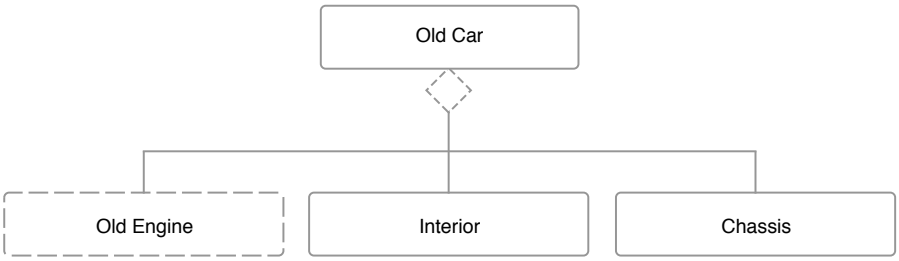


Figure 2.3: Fuzzy aggregation

Graphically, a dashed open diamonds is used to denote a fuzzy aggregate relationship. In Figure 2.3 a car is aggregated from engine, interior and chassis. In 2.3, the engine is old, so the fuzzy class *Old Engine* is a class with the second level of fuzziness. The class *Old Car* is aggregated from the classes *interior* and *chassis* and the fuzzy class *old engine*, making *Car* a fuzzy class with the second level of fuzziness.

2.3.4 *Fuzzy association*

An association relationship with an association name is defined as a binary structural relationship between two classes that associates them, which specifies that the objects of one class are connected to the objects of another class. The association relationship is defined on the basis of classes, not on the basis of the objects of classes, anyway each pair of objects in the corresponding classes has the same association relationship. When fuzzy classes are associated they constitute the fuzzy association relationship. For fuzzy classes with the second

level of fuzziness, the class instances belong to the given classes with membership degrees.

Formally, let c' and c'' be two classes with the second level of fuzziness. The instance o' of c' is an object with the membership degree $\mu_{c'}(o')$, and the instance o'' of c'' is an object with the membership degree $\mu_{c''}(o'')$. The association relationship $assc(o', o'')$ between objects has membership degree as

$$\mu_{(assc(o',o'')} = min(\mu_{c'}(o'), \mu_{c''}(o''))$$

If c' and c'' only have the first level of fuzziness, with membership degree $mem_{c'}$ and $mem_{c''}$ respectively, the association relationship between classes is $ass(c', c'')$, and can be computed as:

$$\mu_{(assc(c',c'')} = min(mem_{c'}, mem_{c''})$$

When the classes have first and second level of fuzziness, the two types membership degrees are

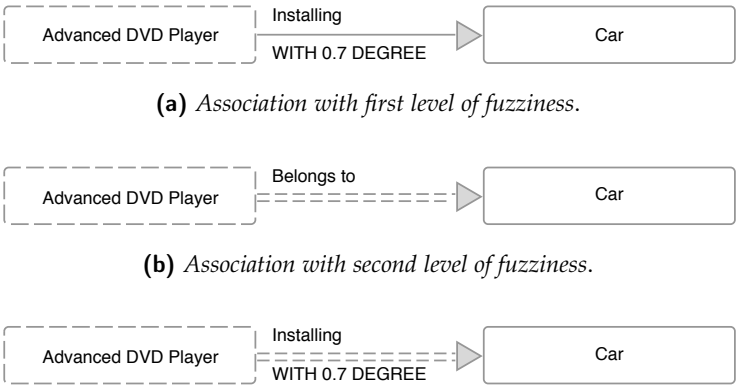
$$\mu_{(assc(o',o'')} = min(\mu_{c'}(o'), \mu_{c''}(o''), mem_{c'}, mem_{c''}) \tag{2.1}$$

$$\mu_{(assc(c',c'')} = min(mem_{c'}, mem_{c''}) \tag{2.2}$$

The pair of words WITH mem DEGREE ($0 \leqslant mem \leqslant 1$) after the association name of an association relationship represents the association relationship with membership degree. Graphically, a double line with an arrowhead is used to denote the association relationship by which the class instances are associated. In Figure 2.4(a) is shown the case where the fuzziness on the relationship is only defined at class level; in (b) it's certain that the relationship belongs to the context but it's unclear the level of participation of the instances of *Advanced DVD Player*; in (c) both previous case hold simultaneously.

2.3.5 *Fuzzy dependency*

A dependency signifies a supplier/client relationship between model elements, where the modification of the supplier may impact the client model elements. The client is then non complete without the supplier. A dependency relationship is only unidirectional. As the association relationship, the dependency relationship is established on the basis of classes, not on the basis of the objects of classes. In fact, the dependency in only related to the classes themselves, and



(a) *Association with first level of fuzziness.*

(b) *Association with second level of fuzziness.*

(c) *Association with first and second level of fuzziness.*

Figure 2.4: Fuzzy association

doesn't require a set of instances to give it meaning. Anyway each pair of objects, that are instances of classes involved in a dependency relationship, is in the relationship itself.

The same reasoning holds also for fuzzy classes with the second level of fuzziness, where the dependency relationship is not affected by such classes because of its class-to-class nature. Still, objects have their degree of membership to their class. When the supplier class in the relationship is fuzzy, also the client is fuzzy (it's not true the other way around), and the degree of membership of the latter is determined (1-to-1) by the one of the former. Formally, taken two classes involved in a dependency relationship, $c_Supplier$ the supplier class with membership degree $mem_{c_Supplier}$ and c_Client be the client class in the relationship, the degree of membership of c_Client is

$$mem_{c_{Client}} = mem_{c_{Supplier}}$$

Figure 2.5: Fuzzy dependency

Graphically a dashed line with an arrowhead is used to denote the fuzziness in the dependency relation. In Figure 2.5 it's shown a dependency case between fuzzy classes *Employee Dependent* and *Employee*. The client one obtains its degree of membership to the model from the supplier.

2.3.6 *A complete example*

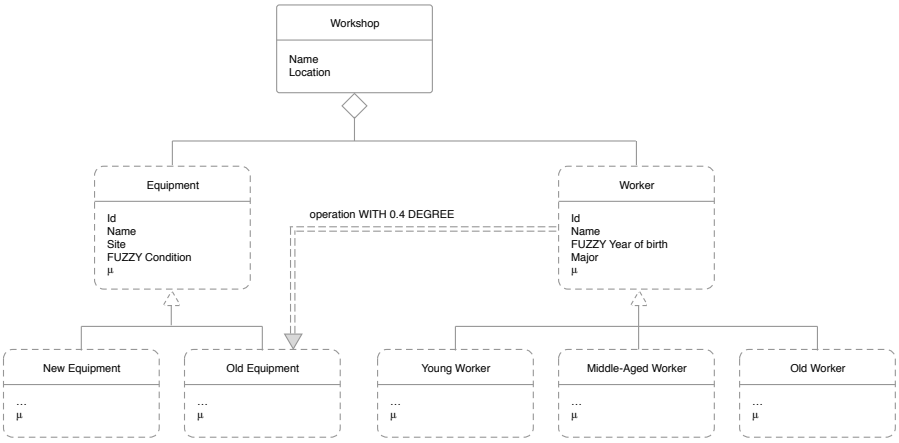


Figure 2.6: Fuzzy UML model

In Figure 2.6, a complete example of Fuzzy UML conceptual model is given. In this example, a *Workshop* is simply aggregated by two classes, *Equipment* and *Worker*. The class *Equipment* is a superclass with *New Equipment* and *Old Equipment* as its two fuzzy subclasses in that they may have fuzzy instances. Similarly, the class *Worker* has three fuzzy subclasses: *Young Worker*, *Middle-Aged Worker* and *Old Worker*. The classes *Worker* and *Old Equipment* have the fuzzy association relationship operation, which exhibits fuzziness at both the first and second levels. The class *Worker* has four attributes, with *Id*, *Name* and *Major* taking on crisp values, and the attribute *Year of Birth* taking on a fuzzy value. Also, the class *Equipment* has four attributes, with *Id*, *Name* and *Site* taking on crisp values and the attribute *Condition* taking on a fuzzy value.

2.4 OTHER NOTABLE WORKS

In literature few other works can be found, here are reported two of the most interesting, the first one notable for the different approach to the fuzzy introduction in data modeling and the last one for having opened the branch of Fuzzy XML modeling.

FUZZY SEMANTIC MODEL Apart from the track started da Zvievi and Chen e followed by Ma et al, a notable work can be found in [4] by Bouaziz, Rafik et al. In the paper the authors propose an extension to the Unified Semantic Model (USM) in order to handle fuzzy information named Fuzzy Semantic Model (FSM).

They claim that USM is better than ER/EER since USM represents more coherently some typical constructs of semantic models and other than abstract concepts of specification/generalization, aggregation and association USM defines constructs to represents constraints among class relationships and distinguishes composite classes from aggregate classes. It can be seen as a derivate of the work by Ma et al with a key difference, the model produced is closer to the context in analysis (to the occurence of the context). For FSM the authors provide modeling directives and graphical representation rules as well as metrics to obtain various degree of memberships and few specification for query handling.

FUZZY XML MODELING Yan Li in [24] shifts back the focus to a more deeper level of fuzzy modeling applied to XML databases. In his work he shows how to extend the attribute value domain with fuzzy sets in order to handle fuzzy information, he proposes three data types: primitive fuzzy, fuzzy collections and user-defined fuzzy data type, which can be restrictions of primitive types or completely new structures with fuzzy attributes (or plain ones). This is the first paper that tackles together Fuzzy data representation and XML modeling.

3

CPN RELEVANCE MODEL

In this chapter, the relevance information model is presented.

3.1 INTRODUCTION

The CPN Relevance Model is a meta-model that enables database designers to handle and represent relevance information about repository content in database schemas.

The CPN acronym stands for Characterizing, Prevalence and Noteworthy. These three are elements that come from a generalization of all the interpretation of relevance met in the research (a glimpse of them is provided in the introduction) and form the basic information that is henceforth expected to be treated.

Let's take an object that can be said to have a relevant attribute, with a certain degree of relevance. The relevance can then be intended in different ways, as following.

- The attribute is *characterizing* the object, that means when the object is described the presence of the attribute has an impact of a certain proportion.
- The attribute is *prevalent within the object*, that means the object shows several aspect and the attribute is more present (or less) that the others
- The attribute has a certain degree of *noteworthiness*, that means that a comparison can be made among several instances of the attribute to establish a ranking based on some criteria.

Each relevance interpretation does not need to be fitted in a context, like a specific interpretation of relevance (an economical relevance criterion, in fact, can be totally inappropriate in an historical context). The three relevance interpretations above highlight a common denominator among all the interpretations, and they model the meaning that human would assume if those were used in natural language conversation.

The triplet of CPN information can then be used to store data about the relevance of an attribute and enhance the quantity and the quality of information that can be extracted from a repository, since a bare "attribute present/not present" can represent a too simplistic modeling.

In this chapter a detailed description of the model components and how to properly intend their meaning is given, in order to clarify the matter for data gathering design purposes and model implementation. Follows a formal description of the model, with definitions of properties and available functions. Those can be seen implemented in the two case studies, later on in this thesis.

3.2 THE MODEL

The three components of the CPN Relevance Model can henceforth be addressed as C-relevance, P-relevance or N-relevance.

Must be kept in mind that the CPN Relevance Model user considers the abstract schema of a repository and not a running instance.

Also, when considering the relevance value of an attribute, with the subjective evaluation of the attribute an end-user (an user for which a CPN enhanced system is designed) can strongly disagree. The key point is that, during the information gathering phase, other than choosing the right scale to represent values (see the issue reported in Characterizing relevance subsection), the CPN Relevance Model user must be sure that information comes from a trusted and reliable source. This source can be impersonated by committee or experts (of the field considered), as for this thesis, who are involved in the most appropriate ways and time.

3.2.1 *Characterizing*

An attribute can be considered characterizing for a specific object obj in a certain degree $C(obj)$. That means that when a natural language speaker has to describe the object obj, he or she will consider the usage of the term associated with the attribute, and therefore the expressive power of the attribute, as degree $C(obj)$ in a range defined a-priori.

For example, given a musical instrument, saying that "the instrument is made of ebony" implies that ebony has a high characterizing power. This has nothing to do with the percentage presence of ebony

in the instrument. Ebony wood can or can not be the prevalent material used to build the instrument. As a clarifying example of this take the the Iron Crown of Lombardy (shown in Figure 3.1), which is highly characterized by iron even if there is no iron at all.

A simple question to verify if an attribute can assume a characterizing role is the following:

> *Can this information be used to describe a more complex object, and can be established how much this information characterize the object it's referring to?*

Or rephrased:

"Can be said what's the weight of the attribute in the object's description?"

Once the question is answered positively, a value range can be defined to support the C-relevance. A wide range allows a finer accuracy when processing the information, but it can be overly complicated to establish the meaning gap between two adjacent values. On the contrary, a short range can be more intuitive and comfortable to use when the relevance information is gathered, but it can result to be too expressively poor. The extreme case is a Boolean value, where the attribute can result to be characterizing or not, and the difference between two possibly characterizing attributes is completely absent.

Figure 3.1: The Iron Crown of Lombardy shows that an attribute can play a strongly characterizing role even if it absent (has a null presence).

3.2.2 *Prevalence*

The prevalence relevance information is perhaps the most intuitive interpretation one can think of, since it expresses the fractional presence

of the attribute with respect to the object. An attribute can represent more than a half of the object is referring to, in this case its P-relevance value could be 60% or 0.6 or 60, for example. The P-relevance nature leads values to be constrained to form the complete, no less and no more, object when all the attributes all considered. To rephrase, the sum of P-relevance values must be a defined and fixed number for all the objects of the repository (60% plus some other attribute P-relevance values must be equal to 100%).

To understand if an attribute can have a P-relevance value, a question follows.

"Can this information indicate the percentage presence of an attribute with respect to the object it refers to?"

Or rephrased:

"Can the attribute partially belong to the object, in different measure that the rest of the attributes?"

For example, in a vanilla-chocolate ice cream, if the vanilla is three fourth of the ice cream, its P-relevance can be expressed as 75%, and this leads the chocolate P-relevance value to be 25%.

For the P-relevance value ranges there is no actual recommendation, since the value can mostly be intended as a percentage value and so treated as a fraction or a percentage. Also, the issue of representing too poorly P-relevance values cannot be encountered without a semantic contradiction. Let's take the extreme case of C-relevance value ranges, if a P-relevance value could be only 0 or 100%, in case for an object two attributes are present at the same time, the total P-relevance would be 200%, which is not permitted.

3.2.3 *Noteworthy*

The trickiest relevance information to understand is the one obtained by the noteworthiness of an attribute. This information is referring to instances of attributes, that can be subject to a quality evaluation for some defined criteria. Another aspect to keep in mind is that the attribute must be considered as part of an object that is part of a Universe of discourse: given that an attribute instance must be comparable with other instances for other objects. A noteworthiness comparator can be intended as provider for N-relevance information.

For example, in a series of shoots for a driver licence photo, each picture can be judge as better than some others, or worse, establishing

so a ranking based on the quality and the appreciation of the single image. The evaluation of the picture is the N-relevance value.

Again, a question to understand if an attribute can represent a valid N-relevance candidate follows.

"Can this attribute itself be rated with some criteria?"

The same indications about the C-relevance value range hold also for N-relevance.

3.3 FORMAL DEFINITION

This section gives a formal definition of the CPN Relevance Model. A CPN Relevance enhanced database is a tuple $T = \langle \rho, F, V, D, \Sigma \rangle$, in which ρ is a finite set of relevance interpretations, F is a finite set of facets, V is a finite set of facet-values, D is a finite set of domains and Σ is a finite set of items.

In this work it is assumed the available relevance interpretations $\Sigma = \{\sigma_1, \sigma_2, ..., \sigma_k\}$ are the ones described above in the chapter, so it is correct to consider $\rho_1 = C, \rho_2 = P, \rho_3 = N$. Actual systems can implement only partially the set of relevance interpretations proposed.

Moreover:

- $F = \{f_1, f_2, ..., f_k\}$ is the set of facets, a set of mutually exclusive and jointly exhaustive information categories, each made by isolating one perspective on the items (a facet), that combine to completely describe all the objects in the repository.

- For $f_i \in F(1 \leqslant i \leqslant k), V(f_i)$ represents the set of values that the facet can assume. Clearly $V(f_i) \subseteq \{v_1, v_2, ..., v_l\}$, i.e. $V(f_i) \subseteq V$. When f_i is given by the context, simply v_j or V is used instead of $v_j(f_i)$ or $V(f_i)$.

- $D \subseteq \rho \times \mathbb{R} \times V$ is a ternary relation that specifies which numeric values the relevance interpretations can assume for each facet-value. Note that integer numbers are also included as known subset of real numbers.

- $\Sigma = \{\sigma_1, \sigma_2, ..., \sigma_k\}$ is the set of items that the database contains and that are classified according to the defined facets.

Then $x(\sigma_i, f_j, v_u, \rho_t) \in D_{\rho_t, v_u}$ represents the value of the t-th relevance interpretation of the i-th object evaluated for the j-th facet and its u-th facet-value and has value according to the allowed domain. It's the basic unit of information the model expects to treat.

3.3.1 *Horizontal, Vertical and Mixed functions*

In addition to the defined model, a series of functions are proposed in order to make use of the relevance information contained in a CPN Enhanced database. Below they're given in their abstract form, since each specific applicative domain would have its own need and requirements, making impossible to report an exhaustive list. Instances of the different kind of functions are presented for the two case studies, in order to show how they can be implemented and what are the possible information that the model can enable to extract from data.

Relevance functions are classified according to the level of user interaction that they can enable within the system implementing the model object of this thesis.

On a first level of interaction, the horizontal functions enable to determine the relevance of the various items with respect to the entire dataset. They take the following form

$$\alpha(\sigma, \overline{F}, \overline{V}, \overline{\rho}) = y \in \mathbb{R}$$

where σ is the object the relevance value y is being computed for, $\overline{F}$ is a subset of the available facets, $\overline{V}$ is a subset of their relative facet-values, and $\overline{\rho}$ is a subset of the available interpretations.

On the second level of interaction, the vertical functions can determine relevance information of facet-values among the repository, denoted vertical information. It's a reading of the data that completely differs from the one made by horizontal functions. In vertical functions the object the relevance is computed for is a facet-value in V not an item in Σ. Item's relevance data is aggregated and results are relative to the parent facet or the entire dataset, in case more than one facets share the same facet-value with the same semantic. Then a vertical function takes the form of:

$$\beta(v, \overline{F}, \overline{\rho}) = w \in \mathbb{R}$$

where v is the facet-value the relevance w is computed for, $\overline{F}$ is a subset of the available facets and $\overline{\rho}$ is a subset of the available relevance interpretations. An example of a beta function can be a function to average P relevance values of the items in order to get the average prevalence of a facet-value with the respect to its facet.

Mixed forms of functions can be made, in order to obtain horizontal information results (regarding item relevance) using vertical informa-

tion. In this case mixed-γ functions (or simply γ function) take the form of:

$$\gamma(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}) = y \in \mathbb{R}$$

where the first four arguments have the same meaning as for horizontal functions, σ is the item the relevance value y is being computed for, $\overline{F}$ is a subset of the available facets, $\overline{V}$ is a subset of the available facet values that holds for the facets in $\overline{F}$, ρ is a subset of the available relevance interpretations, and $\overline{\beta}$ represents a set of vertical functions that computed vertical information to be used by the gamma function.

Also, the dual can be defined, that is a mixed-η function (or simply η function) computing vertical information about a facet-value relevance using horizontal information in the process. This takes the form of:

$$\eta(v, \overline{F}, \overline{\rho}, \overline{\alpha}) \to \mathbb{R}$$

where v is the facet-value the relevance information is being computed for, $\overline{F}$ is a subset of the available facets, $\overline{\rho}$ is a subset of the available relevance interpretations and $\overline{\alpha}$ is a set of horizontal functions to be used by the η function.

For all the kinds of functions defined and described above, an extended version exists that can be applied on a restriction of the dataset. Functions treating a subset of the available items are denoted by a star symbol, then the following ones exist:

- $\alpha^*(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\Sigma}) \to \mathbb{R}$
- $\beta^*(v, \overline{F}, \overline{\rho}, \overline{\Sigma}) \to \mathbb{R}$
- $\gamma^*(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}, \overline{\Sigma}) \to \mathbb{R}$
- $\eta^*(v, \overline{F}, \overline{\rho}, \overline{\alpha}, \overline{\Sigma}) \to \mathbb{R}$

They all have an argument more than their more restrictive version described above, $\overline{\Sigma}$, that represents the subset of items that have to be involved computing the function value. Clearly when $\overline{\Sigma}$ is equal to Σ, the starred version acts as the plain one.

4

MUSEUM OF MUSICAL INSTRUMENTS

In this chapter, the first of two case study is presented.

4.1 INTRODUCTION

The Civic Museum of Musical Instruments in Milan came into being in 1958, thanks to the far-seeing interest of the city government, which purchased the collection of instruments formed by Maestro Natale Gallini in the course of a search lasting around forty years. The main core of the museum's holdings is thus not of historical origin, but is, rather, the fruit of private collecting: a frequent phenomenon in the story of musical life in Milan[9].

In this chapter it's reported the study made on the keyboard instruments section of the museum, and how the CPN model has been implemented on a repository designed ad-hoc from scratch to extract non-trivial information about instruments and the classification method itself.

Relevance information is essential in classifying properly every item given the continuous restoration made on the elements in the collection. In fact, each wood type that did characterize an instrument in a century could have been almost entirely replaced, shifting its classification even at the other end of the spectrum. All the terms and qualifications given and attributed to each instrument must be supported and justified by an interpretation that resolves every ambiguity that can arise.

It's not in the aims of these thesis give an exhaustive description of how the various types of instruments differs from each other, as well as how the single parts of each instrument relates to historical periods or geographical location. An overall description of the subject topic is provided, to understand what kind of data has been analyzed to design a relevance-enabled data model, then a formal implementation of the abstract functions from Chapter 3 is given.

4.2 KEYBOARD INSTRUMENTS

Among the various kind of musical instruments, the ones played using a keyboard are classified as Keyboard Instruments. Under this umbrella there are, regarding this thesis, Harpsichords, Clavichords, Virginals, Spinets, Organs and Pianos. Each one of these can be classified according to [21] that refers to the Hornbostel-Sachs classification[10] as shown in Figure 4.1.

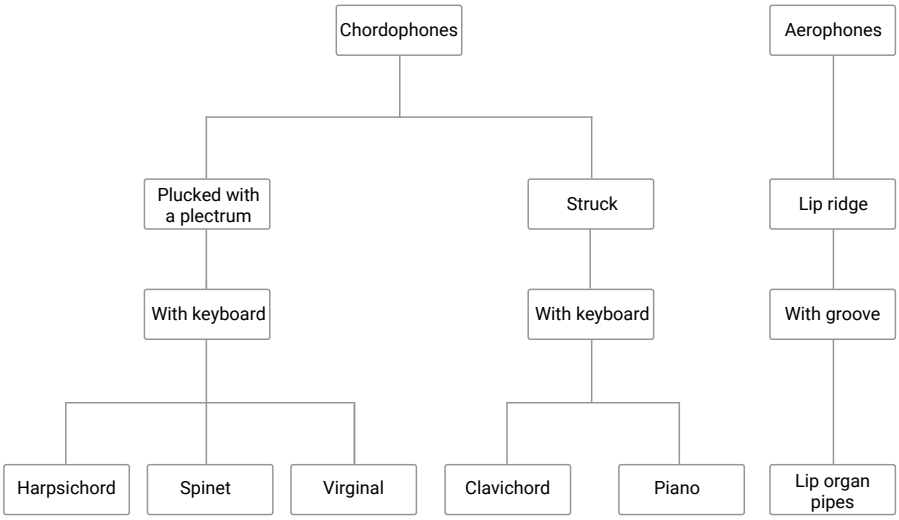


Figure 4.1: Classification of the keyboard instruments analyzed

A keyboard instrument is composed of several parts, depending on the specific instrument type.

Instrument parts can be considered as unique objects, even if compositions of other constituent parts, or as simply aggregations of single and specific objects and so treated as collections. An example is the keyboard, which is a collection of keys but it's almost always considered as an atomic entity, or the jacks in a harpsichord that are collectively referred to as a collection.

Every component is made of one or more wood type and has hallmarks left by authors or by restorers during the centuries-long life of the instrument and its evolution. In addition, every component can be attributed to one or more epoch or geographical area, for its constructive or mechanical features and can represent a distinctive characteristic of an instrument with respect to another.

A deep, precise analysis by field experts has been conducted and each instrument has been profiled and recorded into the catalogue of the Museum of Musical Instruments[9]. The descriptive profiles of the instruments report the peculiarity of each aspect in analysis as

well as considerations about the conservation state or author attribution, whenever it can be ambiguous. In total, the catalogue reports 75 instruments.

The catalogue moreover provides hints of relevance information that have been used to generate the data set for this project.

The set of materials that can overall found reported in the catalogue can be grouped in three subsets:

WOOD TYPE ebony, maple, mahogany, cypress, walnut, beech, rosewood, poplar, linden, boxwood, chestnut, durmast, oak, fruit, fir, cherry, elm, birch

METALS tin, copper, steel, lead, brass, iron

OTHER bone, felt, leather, ivory

4.3 FACET ANALYSIS

For this case study, faceted analysis is conducted as followings.

A. Keyboard materials: this facet permit to identify instrument by the presence of a specific material in active components that are responsible to produce sound (keyboard, mechanics, jacks).

B. Soundboard material: this facet describes the composition of the passive parts of the instrument in charge of amplifying sound or enclosing the instrument in a furniture shaped case.

C. Place of birth: this facet indicates the city where the instrument has been built.

D. Time of birth: this facet indicates the year when the instrument has been built.

E. Conservation state: through the centuries, instruments have been kept with the due regards or just used and mistreated; this facet indicates and evaluation of the conservation state of the instrument as we can look at it nowadays.

F. Artistic merits: instruments could have been built for Kings, Queens or Emperors or more modest owner and hence have been embellished with different kinds of ornaments, some most elegant or precious than others; this facet indicates the finesse degree of the instrument in its entirety.

4.4 CPN MODEL IMPLEMENTATION

As already mention, the study of the instruments has been conducted by experts working along with the Museum of Musical Instruments under the supervision of Castello Sforzesco directory.

Each instrument profile has then been analyzed and relevance information has been retrieved following the CPN Relevance Model. Two of the facets have been enhanced to handle relevance information: *Keyboard material* and *Soundboard material*. Figure 4.2 shows how data has been stored and what's the semantic for the example, as the general idea is following.

Facet values are material types for both facets, and the implementation of CPN Relevance Model can be intended as following:

- Characterizing Relevance: a material can be strongly representative in the profile of the object, and the estimation of its weight is given in a range from 1 to 10 where 1 represents that the facet-value (the material) has really no expressive power when used to describe the object and 10 represents that its presence is the most prominent and absolutely dominant characteristic for the object profile.

- Prevalence Relevance: more than one material can compose an object (Keyboard or Soundboard), then the percentage of presence of a material, with respect to the object, represent the relevance in Prevalence criteria. I.e. in a keyboard made for 50% of walnut, the walnut P-relevance value is 0.5.

- Noteworthy Relevance: each material in the instrument have a different quality degree, and can be noteworthy in a certain measure according to the judgment of an expert like a restorer or a field researcher. The N-relevance value indicates this noteworthiness degree.

A formal definition of the CPN Relevance model implementation is now provided.

4.4.1 *Data model*

The CPN components of the tuple $T = \langle \rho, F, V, D, \Sigma \rangle$ are defined as follows

- The set of facets F is $F = \{\mathtt{Keyboard}, \mathtt{Soundboard}\}$

KEYBOARD MATERIALS

	ebony			fir			maple		
	C	P	N	C	P	N	C	P	N
447. Clavicembalo				3	35	4	2	5	2
448. Clavicembalo	5	10	5				1	10	6
449. Clavicembalo	8	5	5	1	35	6			
450. Clavicembalo	9	10	7	6	5	5			
451. Clavicembalo	2	10	7	8	10	8			
452. Clavicembalo in miniatura				7	90	3			
453. Clavicembalo	7	5	7	3	35	6			
454. Clavicembalo convertito in pianoforte	10	5	7	7	15	8			
455. Clavicembalo	7	10	6	1	90	4			

Figure 4.2: Detail of the CPN data for the museum repository. Where data is missing means that the material has not been found in the instrument. It can be seen that *ebony wood* can be considered highly characterizing for the keyboard of instrument 449,450,454, while is almost irrelevant for instrument 451, even if they are estimated on the same level of noteworthiness. For instrument 452, *fir wood* is the prevalent material in its keyboard, and can be considered characterizing while not so noteworthy. For some instrument, as for 455, also can happen that the most prevalent material is neither noteworthy or characterizing.

- The set of facet-values is composed by all the materials that can be found in the catalogue instrument descriptions, $V = \{ebony, boxwood, maple, mahogany, cypress, walnut, beech, cherry, rosewood, poplar, linden, boxwood, tin, brass, bone, felt, leather, ivory\}$
- Σ is the set of instruments contained in the catalogue
- $\rho = \{\rho_1 = C, \rho_2 = P, \rho_3 = N\}$, since for this case study data for all the relevance interpretations have been collected
- The relevance value domains referred by D are described above as $[1; 10] \in \mathbb{N}$ for ρ_1 and ρ_3 and $[1; 100] \in \mathbb{N}$ for ρ_2

4.4.2 *Function implementations and examples*

From the model definition in Chapter 3, four types of abstract functions come those can be implemented given the specific level of interaction that has to be allowed to the end user:

- α functions, that accounts only information about one item in order to establish its relevance value considering one or more interpretations of one or more facet-values and facets (providing horizontal information).
  These take the form of: $\alpha(\sigma, \overline{F}, \overline{V}, \overline{\rho}) \to \mathbb{R}$
- β functions, that are meant to retrieve vertical information about relevance of facet-values.
  These take the form of: $\beta(\nu, \overline{F}, \overline{\rho}) \to \mathbb{R}$
- γ functions, that computes horizontal information using items data and vertical information together.
  These take the form of: $\gamma(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}) \to \mathbb{R}$
- η functions, duals of the γ ones, that computes vertical information using aggregated items' data and horizontal information together.
  These take form of: $\eta(\nu, \overline{F}, \overline{\rho}, \overline{\alpha}) \to \mathbb{R}$

For all the abstract functions reported, the following holds: σ or ν are respectively the item or the facet-value for which the relevance value is being computed for, $\overline{F}$ is a subset of the available facets, $\overline{V}$ is a subset of the available facet-values, $\overline{\rho}$ is a subset of the available

relevance interpretations, $\overline{\alpha}$ is a set of horizontal functions and $\overline{\beta}$ is a set of vertical functions.

Let's recall that $x(\sigma, f, v, \rho)$ is the relevance value with interpretation ρ of the facet-value v of facet f in object σ.

Horizontal functions enable the system to answer user questions as, for example:

- *What's the relevance of this instrument if considering only the quality of the wooden materials of its keyboard?*
  In this case, the researched value is

  $$\mathrm{SumN}(i, \{\mathrm{keyboard}\}, \{\mathrm{ebony}, \mathrm{fir}, ...\}, \{\rho_3\}).$$

  Assuming the function name as SumN and the $\overline{F}$ and $\overline{\rho}$ subset as a single parameter to lighten the notation, function SumN can be implemented as

  $$\mathrm{SumN}(\sigma, f, \overline{V}, \rho_3) = \sum_{m \in \overline{V}} x(\sigma, f, m, \rho_3)$$

- *What are the instruments where there is at least one really characterizing and noteworthy material?*
  In this case it has to be computed

  $$\arg\max_{\sigma \in \Sigma} \{\mathrm{MaxCN}(\sigma, F, V, \{\rho_1, \rho_3\})\}$$

  The horizontal function in this case can be implemented as:

  $$\mathrm{MaxCN}(\sigma, F, V, \{\rho_1, \rho_3\}) = \max_{f \in F, v \in V} \{x(\sigma, f, v, \rho_1) \cdot x(\sigma, f, v, \rho_3)\}$$

Examples of information that vertical functions can enable to obtain are:

- *The average presence of ebony in instruments in the collection*
  The following value has to be computed:

  $$\mathrm{AverageP}(\mathrm{ebony}, F, \rho_3)$$

  and the average vertical function can be implemented as:

  $$\mathrm{AverageP}(v, \overline{F}, \rho_2) = \frac{1}{\|\overline{F}\|} \cdot \sum_{f \in \overline{F}} \left(\frac{1}{\|\Sigma\|} \sum_{\sigma \in \Sigma} x(\sigma, f, v, \rho_2) \right)$$

KEYBOARD MATERIALS SCORE

	FN1	FNMax	%N > average	Average N	Sum PN values	Sum PN values when N = Max N	Total	Ranking
447. Clavicembalo	4,80	3,84	4,00	3,60	3,55	0,30	28,39	69
448. Clavicembalo	7,71	6,69	7,14	6,43	6,45	0,90	50,51	12
449. Clavicembalo	6,28	5,70	4,44	5,56	5,54	0,35	37,11	55
450. Clavicembalo	6,31	5,76	3,75	5,63	5,15	1,40	36,90	56
451. Clavicembalo	6,83	5,90	5,00	5,67	4,60	0,80	39,60	44
452. Clavicembalo in miniatura	5,25	4,65	5,00	4,50	3,30	0,60	33,90	63
453. Clavicembalo	7,29	6,71	8,57	6,57	6,70	2,40	57,79	4
454. Clavicembalo convertito in pianoforte	6,92	6,05	5,00	5,83	5,15	1,20	41,35	41
455. Clavicembalo	5,50	5,10	0,00	5,00	4,20	0,60	21,00	74
456. Clavicembalo	5,75	5,55	2,50	5,50	5,65	3,90	37,75	53

Figure 4.3: Horizontal functions produce horizontal information. They process instrument's data to obtain a relevance measure. In the snippet above, several functions run: *FN1* and *FNMax* account *Max N* values and *Average N* values among keyboard's materials with different weights; *%N > average* is the percentage of instrument that is composed by material rated as above average for their noteworthiness (in this case *%N > average* is a γ-function, which concern of item relevance); *Sum PN* and *Sum PN* functions are two different way to combine the noteworthiness of the keyboard's materials with their quantitative presence; *Total* is the weighted sum of the previous measures; *Ranking* shows the positioning of the instrument in a list ordered by relevance of its keyboard from most relevant to least relevant. In the case study, also Soundboards are evaluated, so the score from keyboard gets combined with the score from the soundboard to obtain the instrument relevance.

- *At maximum how much the walnut characterizes soundboards*
  The following value has to be computed:

  $$\mathrm{MaxC}(walnut, \mathrm{Soundboard}, \rho_1)$$

  and the vertical function takes the form:

  $$\mathrm{MaxC}(v, f, \rho_1) = \max_{\sigma \in \Sigma} x(\sigma, f, v, \rho_1)$$

- *How many instruments have a noteworthy part made of fir wood*
  The researched values is:

  $$\mathrm{CountN}(fir, F, \rho_3)$$

  and the vertical function is:

  $$\mathrm{CountN}(v, \overline{F}, \rho_3) = \left\| \bigcup_{f \in \overline{F}} \{\sigma \in \Sigma | x(\sigma, f, v, \rho_3) \geqslant \lambda\} \right\|$$

SOUNDBOARD

	Average P	Max N	Average N	Average CN	Average CP	Average CP * Average N
ebony	1,52	7	5,50	5,04	0,09	0,49
fir	32,47	9	5,99	34,23	2,06	12,31
maple	6,00	9	4,96	8,31	0,34	1,70
mahogany	6,87	7	5,63	5,33	0,27	1,52
cypress	9,53	8	5,70	7,99	0,47	2,70
walnut	11,00	8	5,79	14,13	0,60	3,49

Figure 4.4: Vertical functions produce vertical information. They process aggregated instrument data to obtain a relevance measure for facet-values. They provide insights about property of the dataset. In the snippet above, several functions run: *Average P* shows the average presence of various materials in the collection; *Max N* shows the noteworthiness of the best instance of an element of a specific material in the dataset; *Average N* computes the average noteworthiness of a material; *Average CN* tends to answer to a question such as *On average, how much does the noteworthiness of a material characterize an instrument?*; while *Average CP* computes data to answer to *On average, how much does a material participate in characterizing an instrument? And what's its weight?*; in the right-most column, *Average CP · Average N* relates the characterizing strength of a material with its presence and the average noteworthiness to express a value about how good is a material in relation to how much it's present.

Example of information that a γ function can enable to extract is *How many materials the instrument sigma has rated as noteworthy above their average noteworthy rate?*
Given a vertical function $\beta_1(v, f, \rho_3)$ that computes the average noteworthiness of a given material v for the facet f, the researched value is:

$$\mathrm{AboveAvgN}(\sigma, F, V, \rho_3, \beta_1)$$

that can be calculated as

$$\mathrm{AboveAvgN}(\sigma, F, V, \rho_3, \beta_1) = \left\| \bigcup_{f \in \overline{F}} \{v \in V | x(\sigma, f, v, \rho_3) > \beta_1(v, f, \rho_3)\} \right\|.$$

An example of information that a η function can enable to extract is: *In how many instruments ebony wood represents the prevalent wood*
Given a function $\alpha_1(\sigma, \overline{F}, \overline{V}, \rho_2)$ that computes the maximum value of prevalence among the materials of the instrument sigma, the researched value is

$$\mathrm{MostPresent}(\mathrm{ebony}, F, \rho_2)$$

that can be computed as

$$\mathrm{MostPresent}(v, \overline{F}, \rho_2) = \left\| \bigcup_{f \in \overline{F}} \{\sigma \in \Sigma | x(\sigma, f, v, \rho_2) = \alpha_1(v, f, \rho_2)\} \right\|.$$

5

EDOC@WORK

In this chapter, the second of two case study is presented.

5.1 INTRODUCTION

EDOC@Work 3.0, Education and WORK On Cloud, is a national research project developed in the frame of Smart Cities program and Communities Regioni Convergenza, promoted by the Minister of Education, University and Research.

During the academic year 2014-15, 120 teachers have been interviewed about the "Policultura Digital Storytelling" project of Politecnico di Milano, in order to verify how the introduction of modern technology could impact on educational activities. The survey was intended to identify how learning processes have been designed and which methodologies and tips have been used to create each technology-enhanced experiences. Those experiences were all aimed to produce a digital story, composed used the tool 1001 Storia[7].

The disruptive impact of technology evolution on traditional educational methodologies has been analyzed with the Dynamic TPACK model[8], an evolution of the TPACK model[11] that allows to track the growth of interactions between content, pedagogical and technological knowledge of involved subjects.

One of the ultimate goal of Edoc@Work 3.0 is achieving what's called Smart Education, which is a cornerstone for Horizon 2020 and European Digital Agenda. Some of aims are:

- Educational innovation through modern digital instrumentation
- Paradigm switch from "School as content provider" to "School for skills and aptitude development"
- Improvement of quality and efficiency of information leveraging new learning models and digital technologies
- Double the count of graduated students by 2020

ICT and technological innovation enables school system to enhance research activities and create new figures to share and spread knowl-

edge. Just using modern devices, as tablets, notebooks, e-readers, video-walls and multimedia interactive blackboards is not enough to fully express the Smart Education vision. The final phase of the Edoc@work 3.0 project is the individuation of best methodologies and solutions to introduce technologies and dramatically renew didactics. A survey of educational experiences has been conducted to understand if and how technology supported the achieving of pedagogical benefits and to estimated the real convenience.

The repository of data for this case study is already up and running, but does not implement any relevance model at all. A database of educational experiences already exists and a web portal provides basic exploratory functionalities. This has been used as a starting point to consider relevance information incorporation. A description of such database can be found in Appendix C.

In this chapter, at first an overview is given of what and how educational experiences have been collected and what are their main characteristics. Then, it's described how the CPN Relevance Model has been implemented to enhance the quality of exploration for this repository and following a formal implementation of the abstract functions from Chapter 3 is given.

5.2 EDUCATIONAL EXPERIENCES

Educational experiences data has been collected and profiled interviewing one-to-one teachers who participated at Edoc@Work 3.0, investigating the main aspects of educational experience where ICT technologies played a crucial role. In particular, analysis focused on three areas:

- Identification of specific context
- Experience's stand out characteristics
- Achieved results

Within the specific context are collected information to precisely picture the experience scenario, like: school typology, geographical location, social-cultural-economical level, characteristics about students involved as problematic elements or excellences, motivational level and participation, and about the teacher: what's her/his idea of him/herself, what's her/his favorite pedagogical approach, how's her/his relationship with technologies also related with didactics. To better understand specific characteristics, other than information about how

hardware and software tools have been used, also the implementation of the experience has been investigated. Questions asked by researchers to participants were about what was the organization, in which phases the work has been split, if roles were assigned to students and how much time did the experience take.

For the purposes of understanding the growth of teachers and students, the Dynamic TPACK model has been used[8]. The acronym stands for Technology, Pedagogy And Content Knowledge. Through the TPACK model, authors described that three kind of knowledge are needed by teachers to efficiently and effectively transmit knowledge: content knowledge, pedagogical knowledge and technological knowledge. These three domains, in fact, interacts with each other generating new knowledge. It's not enough that a teacher is competent for his subject or highly skilled in the use of computers or digital devices, a teacher should know how leverage each aspect to maximize knowledge transmission. The Dynamic component express the change of these knowledge over time. Researchers then have asked about how was the level of knowledge for students and teachers before the experience and what's the feeling at the end, to obtain information about changes and progress.

Among the achieved results there are information about the output of the experience and, more interesting, the benefits obtained by students participating to the experience. It has been asked if communicative skills have been improved, or team-working skills or motivation level has changed. It has been evaluated the content knowledge gained and if the students have developed transversal skills.

In total, 113 educational experiences have been analyzed.

5.3 FACET ANALYSIS

Educational experiences can be classified according to several aspects like year, school level, subject and so on. Among all these information, some have been selected as candidate to be treated with the CPN Relevance Model. The chosen facets included in this case study are listed below.

A. Pedagogical organization: students can face an exercise grouped together or alone.

B. Site: students can work at home, at school or on a guided visit in some non-scholastic environment. This facet indicates which kind of organizational choices are included in an experience.

C. Benefits: this facet indicates which benefits have been acquired by students with the experience. The values this facet can assume (and so the benefits students could have gained) are: Content comprehension, Curiosity, Critical thinking, Communication skills, Collaboration skills, Media literacy, ICT literacy, Initiative, Design capability, Leadership and Motivation.

D. Teacher TPACK: this facet indicates how the technological, pedagogical or content knowledge have increased for the teacher with the experience.

E. Student TPACK: same as Teacher TPACK, the facet indicates the growing of student technological capabilities and content knowledge.

F. School level: the facet indicates at what level of education the experience took place

Each one represents a different perspective to look at the repository of experiences.

5.4 CPN MODEL IMPLEMENTATION

Relevance information has been gathered for *Pedagogical organization*, *Site*, *Benefits*, *Teacher TPACK* and *Student TPACK*. Following the CPN Relevance Model, experts have analyzed 113 experiences profiles; they provided relevance value in these ranges:

- for C-relevance: integers from 0 to 10;
- for P-relevance: integers from 0 to 100 (representing the after-point digits of a percentage value);
- for N-relevance: integers from 0 to 10.

Figure 5.1 shows how data has been stored and what's the semantic for the example. To better clarify how relevance has been intended during this case study, the meanings of each component for the various facets follow.

- Pedagogical organization information has been evaluated considering (for each experience) if the fact students were working alone or grouped had a characterizing impact for the experience (C-relevance), in which measure they worked (P-relevance) and how the organization choice resulted a winning choice (N-relevance). Typical questions experts have been called to answer were: "How characterizing is for this experience that 200 students were grouped and that they worked together to write and record dialogues?" or "Students have worked most of the time alone, and they produced high level material that have been uploaded by the teacher. How much is it noteworthy?"

- Site has been subjected to a similar evaluation as the Pedagogical organization, here the workplace of students for an experience has been taken into account. In cases teachers led their students on a guided tour, this have been considered more characterizing than just playing or hanging outside the school area or interviewing parents and relatives. The measure they worked at home or at school or remotely have been used to establish the P-relevance. The importance of the place with respect to the experience has been used as noteworthiness degree.

- For each benefit, it has been evaluated how the impact of the benefit characterizes the experience, what's the relative impact for the various benefits acquired, then transformed in percentages for the experience, and how strong has been the impact (for the N-relevance). The same reasoning holds for Teacher TPACK and Student TPACK.

A formal definition of the CPN Relevance model implementation is now provided.

5.4.1 *Data model*

The CPN components of the tuple $T = \langle \rho, F, V, D, \Sigma \rangle$ are defined as follows

- The set of facets F is $F = \{\mathtt{TeacherTPACK}, \mathtt{StudentTPACK}, \mathtt{Site}, \mathtt{Pedagogical\ Organization}, \mathtt{Benefits}\}$

- The set of facet-values is composed by the values that facets can assume

	TPACK														
	Delta Teacher Tk			Delt Teacher Pk			Delta Teacher Ck			Delta Student Tk			Delta Student Ck		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_302 (Sozzi)				7	100	10				2	100	6			
EXP#14_306 (Artale)	4	60	5	3	20	1	0	20	0	1	50	4	7	50	3
EXP#14_314 (Bossi)	3	60	5				2	40	3	3	40	5	2	60	6
EXP#14_315 (Cappelli)	2	50	7				2	50	3	2	60	7	2	40	6
EXP#14_318 (Castellani)	5	30	8	5	40	5	5	30	6	3	80	8	5	20	6
EXP#14_311 (Di Camillo)	4	50	7	6	40	4	3	10	3	4	70	6	0	30	5
EXP#14_317 (Di Ronza)	0	5	1	2	10	3	0	85	1	0	50	6	0	50	6
EXP#14_309 (Gavazzi)	5	40	6	5	40	6	5	20	2	5	50	6	5	50	6
EXP#14_308 (Mascape)	2	90	7	1	5	1	2	5	1	2	30	4	10	70	8

Figure 5.1: Detail of the CPN data for the EDOC repository, TPACK Data. Where data is missing means that the specific aspect has been missing in the interview with participants to the EDOC project. It can be seen that for experience 14_318, the teacher's technology knowledge grew quite a lot but the most prevalent growth has been in pedagogical knowledge. Students, in experience 14_308, grew strongly in content knowledge and this is considered as extremely characterizing. In the same experience, the teacher didn't increase his knowledge of content or pedagogy but strengthened his technology skills. In experience 14_302, students only acquired technology skills but in a significantly lower measure than their teacher. For the sake of clarity, the P-relevance for the Teacher TPACK, gives the percentage value of how the growth can be split among technology, content and pegaogical knowledge. The same applies for Student TPACK.

  - $V(\text{TeacherTPACK}) = \{\text{Tk}, \text{Pk}, \text{Ck}\}$
  - $V(\text{StudentTPACK}) = \{\text{Tk}, \text{Ck}\}$
  - $V(\text{Site}) = \{\text{At home}, \text{At school}, \text{External}\}$
  - $V(\text{Pedagogical organization}) = \{\text{Teamwork}, \text{Individual work}\}$
  - $V(\text{Benefits}) = \{\text{Content comprehension}, \text{Curiosity}, \text{Critical thinking}, \text{Communication skills}, \text{Collaboration skills}, \text{Media literacy}, \text{ICT literacy}, \text{Initiative}, \text{Design capability}, \text{Leadership}, \text{Motivation}\}$

- Σ is the set of educational experiences surveyed
- $\rho = \{\rho_1 = C, \rho_2 = P, \rho_3 = N\}$, since for this case study data for all the relevance interpretations have been collected
- The relevance value domains referred by D are described above as $[1; 10] \in \mathbb{N}$ for ρ_1 and ρ_3 and $[1; 100] \in \mathbb{N}$ for ρ_2

5.4.2 *Function implementations and examples*

Also from the model definition in Chapter 3, four types of abstract functions come those can be implemented given the specific level of interaction that has to be allowed to the end user:

- α functions, that accounts only information about one item in order to establish its relevance value considering one or more interpretations of one or more facet-values and facets (providig horizontal information).
  These take the form of: $\alpha(\sigma, \overline{F}, \overline{V}, \overline{\rho}) \to \mathbb{R}$
- β functions, that are meant to retrieve vertical information about relevance of facet-values.
  These take the form of: $\beta(v, \overline{F}, \overline{\rho}) \to \mathbb{R}$
- γ functions, that computes horizontal information using items data and vertical information together.
  These take the form of: $\gamma(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}) \to \mathbb{R}$
- η functions, duals of the γ ones, that computes vertical information using aggregated items' data and horizontal information together.
  These take form of: $\eta(v, \overline{F}, \overline{\rho}, \overline{\alpha}) \to \mathbb{R}$

For all the abstract functions reported, the following holds: σ or v are respectively the item or the facet-value for which the relevance value is being computed for, $\overline{F}$ is a subset of the available facets, $\overline{V}$ is a subset of the available facet-values, $\overline{\rho}$ is a subset of the available relevance interpretations, $\overline{\alpha}$ is a set of horizontal functions and $\overline{\beta}$ is a set of vertical functions.

Let's recall that $x(\sigma, f, v, \rho)$ is the relevance value with interpretation ρ of the facet-value v of facet f in object σ.

Horizontal functions enable the system to answer user questions as, for example:

- *What's the relevance of educational experience σ if I account only for the prevalence of collaborative work?*
  In this case, the researched value can be computed with a PRel horizontal function as:

$$\begin{aligned}&\mathsf{PRel}(\sigma, \mathsf{Organization}, \mathsf{CollaborativeWork}, \rho_2)\\ &= x(\sigma, \mathsf{Organization}, \mathsf{CollaborativeWork}, \rho_2)\end{aligned}$$

- *What are the most relevant experiences where teacher had a higher growth in technology knowledge and students worked mainly at home?*
  It has to be computed

$$\arg\max_{\sigma \in \Sigma} \{\mathsf{VLKUPIF}(\sigma, \{\mathsf{TeacherTPACK}, \mathsf{Site}\}, \{\mathsf{Tk}, \mathsf{Athome}\}, \{\rho_2, \rho_3\})\}$$

  with the horizontal function:

$$\begin{aligned}&\mathsf{VLKUPIF}(\sigma, \{p1, p2\}, \{v1, v2\}, \{\rho_2, \rho_3\})\\ &= \begin{cases} x(i, p1, v1, \rho_3) & \text{if } x(i, p2, v2, \rho_2) \geqslant \lambda \\ 0 & \text{otherwise} \end{cases}.\end{aligned}$$

  that selects a relevance value for the facet-value of an item when another facet-value shows a relevance value over a defined λ

Examples of information that vertical functions can enable to extract are:

- *On average how much of the experience is teamwork*
  The following value has to be computed:

$$\mathsf{AverageP}(\mathsf{teamwork}, \mathsf{Organization}, \rho_2)$$

	Teacher TPACK		Student TPACK		Organiz.	Benefits		Overall	
	Average C values	Sum PN values	Average C values	Sum PN values	Average C values	Average N values	% N > average	Total	Ranking
EXP#14_302 (Sozzi)	7,00	10,00	2,00	6,00	10,00	2,50	0,17	82,83	68
EXP#14_306 (Artale)	2,33	3,20	4,00	3,50	3,00	4,17	0,33	56,10	107
EXP#14_314 (Bossi)	2,50	4,20	2,50	5,60	5,00	7,20	0,60	85,60	53
EXP#14_315 (Cappelli)	2,00	5,00	2,00	6,60	7,00	6,25	0,25	84,55	61
EXP#14_318 (Castellani)	5,00	6,20	4,00	7,60	2,00	6,67	0,33	94,07	21
EXP#14_311 (Di Camillo)	5,00	5,40	2,00	5,70	2,00	6,00	1,00	83,30	66
EXP#14_317 (Di Ronza)	0,67	1,20	0,00	6,00	3,00	6,00	0,25	62,52	103
EXP#14_309 (Gavazzi)	5,00	5,20	5,00	6,00	7,00	6,17	0,50	90,10	34
EXP#14_308 (Mascape)	1,67	6,40	6,00	6,80	0,00	8,00	0,00	95,27	18

Figure 5.2: Horizontal functions produce horizontal information. They process instrument's data to obtain a relevance measure. In the snippet above, several functions run: *Average C values* computes the average characterizing power of the various facet-values; *Sum PN* combines the noteworthiness of an aspect with its percentage presence; *Average N values* computes the average noteworthiness of the various facet-values; *% N > average* is the percentage of the experience's aspects that is rated as more noteworthy than the average (in this case *% N > average is a γ-function, which concern of item relevance)*; *Total* is the weighted sum of the previous measures; *Ranking* shows the positioning of the experience in a list ordered by relevance from most relevant to least relevant.

that can be computed with the vertical function

$$\mathrm{AverageP}(v, \overline{F}, \rho_2) = \frac{1}{\|\overline{F}\|} \cdot \sum_{f \in \overline{F}} \left(\frac{1}{\|\Sigma\|} \sum_{\sigma \in \Sigma} x(\sigma, f, v, \rho_2) \right)$$

- *At maximum how much the student content knowledge had grown*
  The researched value is

  $$\mathrm{MaxN}(\mathrm{Ck}, \mathrm{StudentTPACK}, \rho_3)$$

  and the vertical function takes the form of

  $$\mathrm{MaxN}(v, f, \rho_3) = \max_{\sigma \in \Sigma} x(\sigma, f, v, \rho_3)$$

- *How many experiences led to develop leadership skills among students*
  The researched value is

  $$\mathrm{CountN}(\mathrm{leadership}, \mathrm{Benefits}, \rho_3)$$

where the vertical function is:

$$\mathrm{CountN}(v, \overline{F}, \rho_3) = \left\| \bigcup_{f \in \overline{F}} \{\sigma \in \Sigma | x(\sigma, f, v, \rho_3) \geqslant \lambda\} \right\|$$

	Characterizing		Prevalence		Noteworthy	
	C Average	Max C	P Average	Max P	N Average	Max N
Delta Teacher Tk	2,97	8	47,30	100	4,58	9
Delta Student Tk	3,86	8	45,82	100	5,56	10
Delta Teacher Ck	2,76	7	12,92	100	3,93	8
Delta Student Ck	4,66	9	54,18	100	6,07	9
Delta Teacher Pk	3,63	10	39,78	100	4,57	9
Team work	4,72	10	72,26	100	5,71	10
Individual work	4,30	7	27,74	100	5,28	7

Figure 5.3: Vertical functions produce vertical information. They process aggregated experience data to obtain a relevance measure for facet-values. In the snippet above, several functions run: C, P and N *Average* compute respectively the average of the characterizing power, the prevalence and the noteworthiness of the various facet-values; *Max* does the same computing the maximum value, this is an insight on what is the best value that can be achieved in at least one experience. The *TPACK delta-s* are reordered in order to show how vertical functions can provide immediately interesting results.

Example of information that a γ function can enable to extract is: *How many benefits the experience sigma are rated as the most noteworthy for the entire dataset?*
Given a function $\beta_1(v, f, \rho_3)$ that computes the highest noteworthiness value of a given benefit v among the benefits facet f, the researched value is:

$$\mathrm{BestCount}(\sigma, \mathrm{Benefits}, V, \rho_3)$$

that can be calculated as

$$\mathrm{BestCount}(\sigma, F, V, \rho_3, \beta_1) = \left\| \bigcup_{f \in \overline{F}} \{v \in V | x(\sigma, f, v, \rho_3) = \beta_1(v, f, \rho_3)\} \right\|$$

An example of information that a η function can enable to extract is: *The average impact on motivation among experiences highly characterized by a noteworthy teamwork*

Given the function $\alpha_1(\sigma, \overline{F}, \overline{V}, \{\rho_1, \rho_3\})$ that expresses with a 1 value if the experience sigma is characterized by a noteworthy teamwork (or a 0 otherwise), the researched value is

$$AvgImpact(motivation, Benefit, \{\rho_2, \rho_3\}, \alpha_1)$$

that can be computed as

$$AvgImpact(v, \overline{F}, \overline{\rho}) = \frac{1}{\|\overline{\Sigma}\|} \cdot \sum_{\sigma \in \Sigma, f \in \overline{F}, r \in \overline{\rho}} \alpha_1 \left(\sigma, \overline{F}, \overline{V}, \{\rho_1, \rho_3\}\right) \cdot x(\sigma, f, v, r)$$

6 CONCLUSIONS AND FUTURE WORK

This thesis presents a model to handle relevance information, in order to lay the foundations to build future Exploratory Systems. Such a kind of systems aims to give to the user an experience of data exploration, that means the user discovers and observes data from several point of view and establishing a dialogue with the system.

Research began studying in deep the concept of relevance and how the natural language speaker treats relevance interpretations when communicating. In fact, it feels natural to switch among several criteria, without explicitly addressing one, when rating the relevance of an object with respect to a context.

Then two case studies have been faced, in order to both collect data and deeply understand how the relevance could have been intended for two different contexts and domains. What resulted was a set of shared concepts between the Museum case study and the EDOC case study, that led to define the CPN triplet that holds relevance data.

While working on data structures, possible usage and combinations of information have been surveyed as well, to understand how could have been extracted meaningful knowledge from data. Several kind of abstract functions have been defined and, for each one, a possible example implementation has been realized. The abstract functions area reported in Table 6.1 and a brief description is following.

HORIZONTAL FUNCTIONS that compute results for items in the repository (a.k.a. horizontal information), accounting relevance of the single item aspect such as *percentage of ebony* or *noteworthiness of off-school work*

VERTICAL FUNCTIONS that are designed to provide results about the relevance of the aspect in the specific instance of the repository (a.k.a. vertical information), considering item's aggregated data and filtering

MIXED FUNCTIONS that mix the two kind above, providing relevance value for either single items or facet-values, with the support of functions that punctually address horizontal or vertical information

In a simulated environment, (preliminary) test have been conducted to highlight the differences between a fuzzy, nuanced data model and a more classical and rigid one when dealing with knowledge extraction.

In this, the CPN Relevance Model has proven itself with the two case studies as a reliable and scalable framework. The information that can be extracted, compared with a one that only allows attributes to belong or not to objects (in a boolean fashion), is greatly refined. This has been confirmed by field experts that analyzed data provided by the redesigned database. In context where a precise and multi-perspective view of the context is essential in order to properly extract knowledge and answer to non-trivial questions, the simple "x has the attribute y" or "x doesn't have the attribute y" is not enough and can lead to heavy misunderstandings in data readings.

A first example of this is shown in Figure 6.1, whose values are reported in Table 6.2.

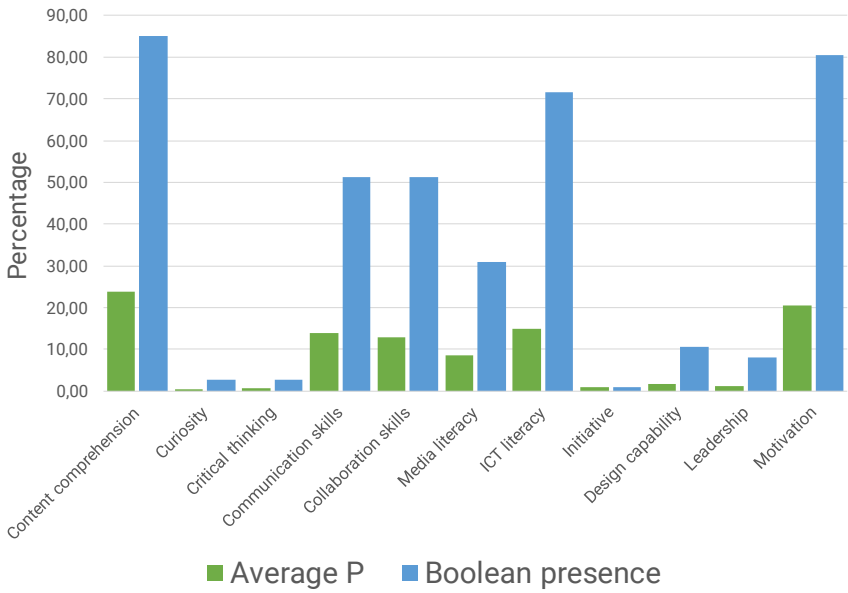


Figure 6.1: Comparison between measures on CPN data model and the boolean one

There, the percentage of the benefit impact is measured for all the experiences with the support of the CPN Model (in green) and the simple boolean one (in blue). It pops up that the impact of almost all *Educational benefits* is greatly overestimated if compared with the finer and more realistic analysis that so far could have been done.

Moreover should be noted that in the example, the relevance has not be considered for the experiences but for an aspect that classify

TYPE	VALUE	FORM
Horizontal	item's relevance	$\alpha(\sigma, \overline{F}, \overline{V}, \overline{\rho}) \to \mathbb{R}$ $\alpha^*(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\Sigma}) \to \mathbb{R}$
Vertical	facet-value's relevance	$\beta(v, \overline{F}, \overline{\rho}) \to \mathbb{R}$ $\beta^*(v, \overline{F}, \overline{\rho}, \overline{\Sigma}) \to \mathbb{R}$
Mixed-γ	item's relevance	$\gamma(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}) \to \mathbb{R}$ $\gamma^*(\sigma, \overline{F}, \overline{V}, \overline{\rho}, \overline{\beta}, \overline{\Sigma}) \to \mathbb{R}$
Mixed-η	facet-value's relevance	$\eta^*(v, \overline{F}, \overline{\rho}, \overline{\alpha}) \to \mathbb{R}$ $\eta^*(v, \overline{F}, \overline{\rho}, \overline{\alpha}, \overline{\Sigma}) \to \mathbb{R}$

Table 6.1: Relevance function types

items in the dataset. In other words, relevance is not only a measure that can be computed for items, but also for the items' attributes in a vision that covers the entire dataset.

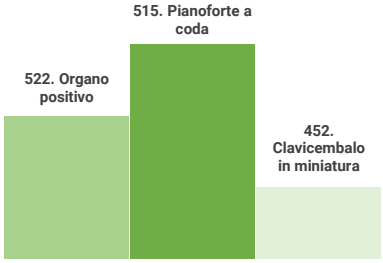
Clearly, also item relevance can be computed. Figure 6.2 reports the "winning" instrument of an hypothetical competition based on the measures implemented in the Museum case study. Metaphorically, each relevance function is a "judge" and its vote has a weight among all the ones of the "jury". Their relevance is then computed with a weighted sum of several horizontal and mixed-γ functions.

The model proves then its flexibility and its capability to enable a more interesting data exploration for systems with a high degree of semantically ambiguous classification. Everything pushes towards a refinement for the support that's given to answering user's exploratory sessions.

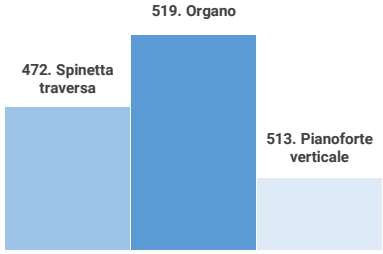
6.1 DOUBTS ABOUT THIS WORK

This work can raise legitimate doubts, such as: what's the cost of data gathering? Or are the relevance values arbitrary? Does it pay to implement the relevance model? Is the simulated environment too complex to obtain meaningful results? And what's the relation between the kind of analysis this relevance model requires and the classical analysis for data hypercubes? Answers are following.

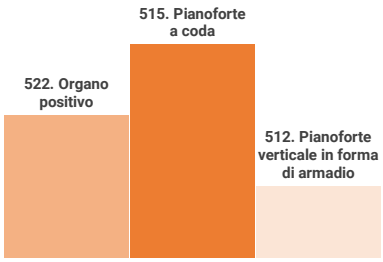
The cost of data gathering is high, especially if done to readapt a repository to handle relevance information rather than start fresh. It has to be clear that also a partial implementation of the model can have a huge impact on the final result. In fact, some can be the cases where collecting data for characterization and noteworthiness



(a) *Podium for Keyboard attributes.*



(b) *Podium for Soundboard attributes.*



(c) *Podium for all attributes.*

Figure 6.2: Most relevant instruments in a simulation for testing purposes

is considered too expensive in terms of time or resources, and just two of three values of the CPN triplet are used (i.e. C and P or P and N). During the development of this thesis, the idea that P-relevance values can be considered the easiest to retrieve arisen, since they're the closest concept to a boolean "there is/there is not" way of treating object attributes among the concepts introduced here. Particular care should be paid on choosing the right value range for the relevance values. In fact, in case of an overestimated value range, data gathering costs can increase easily, given the too many options available that make establishing the relevance value more complicated.

Relevance values can be considered arbitrary, but even an arbitrary values in a nuanced scale is more meaningful than a sharp yes/no.

Does it pay to implement the relevance model? Yes, because it appears, also to field experts, that there is a reading of data that allows to better understand the various phenomena.

Is the simulated environment too complex to obtain meaningful results? It is a complex environment indeed, and for sure it's not intended for the casual user, but all the research has been done to lay down the foundations for the future development of exploratory portals.

Between the CPN model and the classical data hypercube model there is a sort of resemblance, in the sense that for both models data

is seen from multiple point of view. The main difference is that in the CPN model, the facet-values set is not ordered while in hypercubes, dimensions are ordered sets. So there is no meaning in positioning an object according to coordinates established from the facet-values, since facet-values can be casually reordered making the hypercube inconsistent. A more pertinent resemblance can be found comparing a classical hypercube with the CPN triplet, since relevance data, for its nature, is numeric and constitute an ordered set.

6.2 FUTURE WORK

The test here conducted took place in a simulated environment, *id est* functions have been manually built, with a lot of effort, to extract information from data in the most meaningful way. It's desirable that such a job would be extremely simplified, in order to let users create their own relevance measures to address their specific needs.

Seems pretty clear that the CPN Relevance Model has to be implemented by a smart system, that hide the complexity of mixing facet-values, facet, items and all the mathematical artifacts used in the case studies so far, behind an intuitive and adaptable interface. The most suitable environment that could both provide user friendly data exploration, with a rich communicative capability, and a strong mathematical framework appears to be a web portal.

In a more long-term vision, work should focus on the communicative aspect of the portal, since it should convey all the information to the user together with providing clear and accessible tools that foster the user dialogue.

The user's vision of the repository is in a constant evolution, as it would happen if he was interacting with another human who has a larger knowledge available.

All this would be also supported by a notion of relevance embedded in the system, flexible enough to provide the most punctual data when needed, that can be provided implementing the CPN Relevance Model.

BENEFIT	CPN	BOOLEAN MODEL	GAP
Content comprehension	23,87	84,96	-61,09
Curiosity	0,53	2,65	-2,12
Critical thinking	0,75	2,65	-1,90
Communication skills	13,91	51,33	-37,42
Collaboration skills	12,97	51,33	-38,35
Media literacy	8,72	30,97	-22,26
ICT literacy	15,00	71,68	-56,68
Initiative	0,88	0,88	0,00
Design capability	1,73	10,62	-8,89
Leadership	1,19	7,96	-6,77
Motivation	20,44	80,53	-60,09

Table 6.2: Difference between information extracted with CPN Relevance Model and a "boolean" model

BIBLIOGRAPHY

[1] Carlo Batini, Stefano Ceri, and S Navathe. *Entity Relationship Approach*. Elsevier Science Publishers BV (North Holland), 1989 (cit. on p. 7).

[2] Gloria Bordogna, Gabriella Pasi, and Dario Lucarella. "A fuzzy object-oriented data model for managing vague and uncertain information." In: *International Journal of Intelligent Systems* 14.7 (1999), pp. 623–651 (cit. on pp. 7, 11, 13).

[3] Patrick Bosc and Henri Prade. "An introduction to fuzzy set and possibility theory based approaches to the treatment of uncertainty and imprecision in database management systems." In: *Proceedings of the second workshop on uncertainty management in information systems: from needs to solutions*. 1993 (cit. on p. 8).

[4] Rafik Bouaziz et al. "Database design and querying within the fuzzy semantic model." In: *Information Sciences* 177.21 (2007), pp. 4598–4620 (cit. on pp. 7, 25).

[5] Peter Pin-Shan Chen. "The entity-relationship model—toward a unified view of data." In: *ACM Transactions on Database Systems (TODS)* 1.1 (1976), pp. 9–36 (cit. on p. 7).

[6] Linda G DeMichiel. "Resolving database incompatibility: An approach to performing relational operations over mismatched domains." In: *Knowledge and Data Engineering, IEEE Transactions on* 1.4 (1989), pp. 485–493 (cit. on p. 8).

[7] Nicoletta Di Blas, Luisa Marini, and Bianca Boretti. "Comunicare cultura con le nuove tecnologie: 1001Storia per PoliCultura." In: () (cit. on p. 45).

[8] Nicoletta Di Blas et al. "Distributed TPACK: going beyond knowledge in the head." In: *Society for Information Technology & Teacher Education International Conference*. Vol. 2014. 1. 2014, pp. 2464–2472 (cit. on pp. 45, 47).

[9] Andrea Gatti. *Museo degli Strumenti Musicali*. Electa, Milano, 1997 (cit. on pp. 35, 36).

[10] E.M. von Hornbostel and C. Sachs. *Classification of Musical Instruments*. 1914. URL: https://books.google.it/books?id=TsswAQAAMAAJ (cit. on p. 36).

[11] Matthew J Koehler and Punya Mishra. "What happens when teachers design educational technology? The development of technological pedagogical content knowledge." In: *Journal of educational computing research* 32.2 (2005), pp. 131–152 (cit. on p. 45).

[12] Dario Lucarella and Antonella Zanzi. "A graph-oriented data model." In: *Database and Expert Systems Applications*. Springer. 1996, pp. 197–206 (cit. on p. 13).

[13] ZM Ma. "A Literature Overview of Fuzzy Database Modeling." In: *Intelligent Databases: Technologies and Applications* (2007), p. 167 (cit. on p. 11).

[14] ZM Ma, Li Yan, and Fu Zhang. "Modeling fuzzy information in UML class diagrams and object-oriented database models." In: *Fuzzy Sets and Systems* 186.1 (2012), pp. 26–46 (cit. on pp. 7, 8, 14, 15).

[15] ZM Ma, Wen-Jun Zhang, and WY Ma. "Assessment of data redundancy in fuzzy relational databases based on semantic inclusion degree." In: *Information Processing Letters* 72.1 (1999), pp. 25–29 (cit. on p. 19).

[16] ZM Ma, Wen-Jun Zhang, and WY Ma. "Extending object-oriented databases for fuzzy information modeling." In: *Information Systems* 29.5 (2004), pp. 421–435 (cit. on pp. 7, 11, 14).

[17] ZM Ma et al. "Conceptual design of fuzzy object-oriented databases using extended entity-relationship model." In: *International Journal of Intelligent Systems* 16.6 (2001), pp. 697–711 (cit. on pp. 7, 11, 12).

[18] N Marın, O Pons, and MA Vila. "Fuzzy types: A new concept of type for managing vague structures." In: *International Journal of Intelligent Systems* 15 (2000), pp. 1061–1085 (cit. on p. 17).

[19] Juan Miguel Medina, Olga Pons, and Maria Amparo Vila. "GEFRED: a generalized model of fuzzy relational databases." In: *Information sciences* 76.1 (1994), pp. 87–109 (cit. on p. 7).

[20] C Negoita, L Zadeh, and H Zimmermann. "Fuzzy sets as a basis for a theory of possibility." In: *Fuzzy sets and systems* 1 (1978), pp. 3–28 (cit. on p. 10).

[21] S. Paganelli. *Musical Instruments from the Renaissance to the 19th Century*. Cameo (Hamlyn Publishing Group). Fratelli Fabbri, 1970. ISBN: 9780600359203. URL: https://books.google.it/books?id=k%5C_7VPwAACAAJ (cit. on p. 36).

[22] James Rumbaugh, Ivar Jacobson, and Grady Booch. *Unified Modeling Language Reference Manual, The*. Pearson Higher Education, 2004 (cit. on pp. 15, 21).

[23] Nancy Van Gyseghem and Rita De Caluwe. "Imprecision and uncertainty in the UFO database model." In: *Journal of the American Society for Information Science (1986-1998)* 49.3 (1998), p. 236 (cit. on p. 17).

[24] Li Yan. "Modeling Fuzzy Data with Fuzzy Data Types in Fuzzy Database and XML Models." In: *Int. Arab J. Inf. Technol.* 10.6 (2013), pp. 610–615 (cit. on p. 26).

[25] Lotfi A Zadeh. "Fuzzy sets." In: *Information and control* 8.3 (1965), pp. 338–353 (cit. on pp. 7, 8, 10).

[26] Lotfi A Zadeh. *The concept of a linguistic variable and its application to approximate reasoning*. Springer, 1974 (cit. on p. 9).

[27] Arie Zvieli and Peter P Chen. "Entity-relationship modeling and fuzzy databases." In: *Proceedings of the Second International Conference on Data Engineering*. IEEE Computer Society. 1986, pp. 320–327 (cit. on pp. 7, 11–13).

Part I

APPENDIX

A

CHECKLIST TO IMPLEMENT CPN RELEVANCE MODEL

A.1 INTRODUCTION

In this chapter, I provide a checklist that should be considered while building from scratch (or, in case a repository already exists, migrating to) a CPN-enhanced repository. Each suggestion comes from field experience gained during this research, experience that can be source of useful insights to better tackle the "repository upgrade process".

The running example, for the sake of clarity, will be a culinary database (instead of the more sophisticated case-studies dealt with in the thesis).

A.2 THE GOAL

Implementing the CPN Relevance Model, brings the repository system to a level with a higher degree of "explorability", useful if the system user is a human being and not another machine. The CPN is designed keeping in mind the human communicative process, with its strength and flaws, and leverages intuition and doubts typical of a human mind. This should be clear before even considering the adoption of CPN Relevance Model.

A.3 HAVING THE RIGHT DOMAIN

Implementing the relevance model is not feasible for all kind of information. What is primarily needed is data that, when retrieved with a user request (or query), can have a different degree of relevance with respect to the intended request. A query for recipes with parsley brings a lot of results back, and for every recipe it can be said without hesitation that the ingredient can fit in various degrees: well, perfectly or it's just marginal. On the contrary, asking to a basic address book to list all the contacts named "Anna" leaves almost no space to misunderstanding, and that is a symptom of an inappropriate context for relevance information.

If there are domain concepts that can belong to many categories with several degree of membership, they represent a potential candidate to be CPN-enhanced. Domain knowledge of these concepts should be reconsidered, trying to answer to the following questions for each candidate concept.

- Can this information be used to describe a more complex object, and can be established how much this information characterize the object it's referring (what's its weight)? Taking for example the recipe of a "Cinnamon and apple crumble pie", the fact the ingredient Apple appears in the name of the recipe gives it a high degree of characterization with respect to the recipe. Instead for a "Fruit salad" the Apple is not a strictly characterizing aspect.

- Can this information indicate the percentage presence of an object with respect to a context? Quantities for ingredients are a crucial information for the success or failure of a dish. Taking the "Basil Pesto sauce", it can be said that basil is a prevalent ingredient in the mixture, in the order of 60% of the total preparation. This information involves an aspect of data that shouldn't be confused with the previous question: a cake can be mainly made of chocolate and pears, and just 1 or 2 grams or vanilla, but when addressed it would be as a "Vanilla chocolate and pear cake". See that the percentage of vanilla is relatively low with the respect to the total, but its characterizing power is not.

- Can the object this information is about be rated? In case objects can have different or various degrees of finesse, it appears that the quality of an object can influence quality of the object the former is part of. For a "Filet mignon", the quality of the meat cut can be noteworthy within a range of meaning nuances. Meat from the best animals, also treated with the most accurate techniques, will have a high noteworthiness degree, less valuable meat will be less noteworthy.

If no domain concept can be used to answer positively to the questions above, the CPN Relevance model cannot be applied.

Instead, in case the answer to the three questions above is at least once positive, the CPN can theoretically be applied. The process of how the relevance information is gathered can be later tackled and designed. For sure, the designer trying to answer these questions, reconsiders the current database schema from a different perspective.

The three questions above involve orthogonal dimensions that only apparently are mutual dependent. It can be argued that a prevalent ingredient for a recipe can be also a strongly characterizing aspect, but it's not for sure as the "Cinnamon Apple crumble pie" shows (Cinnamon dosage is in the order of grams, with respect to the total pie weight in the order of hundreds of grams or even kilos). It can be argued that the quality of a meat cut, or a finesse of a sauce, or the freshness of a vegetable can strongly influence a dish, but for sure do not influence what the recipe is characterized or mostly composed by. The fact that an abstract recipe can't have an evaluation of an actual ingredient is a pertinent observation, but as premise has been established that the example repository is a log of instances of recipes, not a cooking book. There is no way and/or no meaning to normalize the database separating recipes and dishes since each dish is a "unicum".

A.4 RELEVANCE REPRESENTATION AND VALUE RANGES

Relevance information, in order to be processed by a computer, has to be represented numerically. An object can be labeled as "Highly relevant" or "Lowly relevant" but in order for the model to be effective, these label should be mapped one-to-one to a partially ordered set, for example the set of Natural Numbers. The C-relevance or the N-relevance can hence freely assume values as "Exceptional" or "Poor" or "Very low" or "Average", with the only remarks that two labeled objects must result to be comparable and must be possible to establish which one is the mostly relevant or if they're equally relevant. For example, in a practical case can be defined that three labels are available, "Good", "Average", "Poor" and that the correspondent numerical values are 3, 2, 1. The relevance value is then freely attributable, with the exception of the Prevalence Relevance. That has to satisfy a constraint: the sum of all P-relevance for a facet has to represent the entirety of the object which the facet is referring to. The suggestion in this case is to chose an appropriate and handy range for C and N relevance values, such as 1 to 10, 1 to 5 or 0.1 to 1, and a percentage value for the P relevance value, stored in the system as the only percentage component or a fixed point value (24Must be clear that a large range increases the expressive power of the relevance information, at a higher cost for the data gathering process. This is suitable only for contexts that allow a wide range relevance nuances. On the contrary, having a coarse grained relevance information is handier, but may

cause a too rough evaluation of relevance when data is collected. In fact, the extreme case is a range of only two possible values, which would make the relevance model working in a Boolean fashion with the risk of oversimplifying reality.

A.5 HOMOGENEOUS FACETS

When upgrading from a standard repository to a CPN-enhanced repository, facet-values grouping could need to be reconsidered. Can happen that a facet mixes values concerning different subjects to keep classification simpler to handle. This can potentially break the P-relevance constraint, letting the sum of the parts overcome the unity measure. For example, in case of "Filled Dumplings", an ingredient can result be part of both the filling and the wrapping. If there is only a single facet "Ingredients", when computing the sum of the ingredients for the wrapping can result a value over 100%. This indicates that the subject some facet-values are referring to is different inside the facet itself. To avoid this inconsistency the CPN-ized facet-values has to be grouped homogeneously. That means a facet should always refer to a single subject. To recall the example, for the case of "filled dumplings", the most correct way to work should be having a facet for the filling and a facet for the wrapping, not a common one.

Checking this property is crucial to maintain the correct CPN modeled semantic.

A.6 KEEPING RELEVANCE INFORMATION AND CONTEXTUAL DATA CLOSE

Once found some domain concepts that can be well suited for treating relevance information, the issue of storing relevance data arises. My suggestions are following.

In a ER schema, the CPN-ized attributes should be represented by a weak entity composed by a field for the label (interpretation) and an integer or a floating point value as relevance value. They also should have external keys to the object to which relevance values are referring.

In an Object-Oriented schema, dictionaries or maps can be used to link CPN data to actual value, or classes can replace plain attributes declarations as members the CPN values.

A.7 ANOTHER FEASIBLE CPN-ENHANCED CONTEXT

A scientific literature repository can represent a nice mind exercise to self-assess the level of understanding of the CPN Relevance Model, and can be used to warm up the database designer to make practice elaborating on the concepts at the base of this thesis. During this research, the mentioned repository (described below), has been marginally considered when analyzing relevance information and its structure and it didn't result in a detailed use case, even if it fulfills all the requirements for the CPN Relevance Model.

Let us consider a literature repository, with scientific papers collected and classified. Each paper is composed by a several sections, each one can be resourceful and/or extremely well-written, or could be just very long bunches of text with low meaning. Each article is usually about a topic, but can cover more than one subject, and can face deeply one with respect to another. Each work can be object of great innovation or just a report of the current state of the art for field (a survey).

The chances the targeted reader of this work already know about scientific literature are pretty high, so let assume the context is already clear enough with the brief description above.

I found stimulating trying to consider how to apply a CPN Relevance model to the above context and how that would make literature browsing more accurate and accessible.

B

EDOC@WORK REPOSITORY

In this chapter is described the actual running implementation of the EDOC@Work repository.

B.1 DIFFERENCES

The actual implementation of the EDOC@Work repository does not take into account relevance information at all. Nor provides highly detailed measures as the ones enabled by the model, shown in the case study in Chapter 5.

Since CPN Relevance Model is a generalization of a data model that doesn't implement relevance, it's allowed to compare the two data models (the one in the case study and the running instance) in order to highlight differences and improvements.

In the comparison in Table B.1, the aspects of the running repository are limited to the ones concerning this thesis.

ENTITY	CASE STUDY	ACTUAL EDOC
F	Teacher TPACK, Student TPACK, Pedagogical organization, Site, Benefits	TPACK, Pedagogical organization, Benefits
V	• V(Teacher TPACK) = Tk, Pk, Ck • V(Student TPACK) = Tk, Pk • V(Pedagogical org.) = Teamwork, Individual work • V(Site) = At home, At school, External • V(Benefits) = Content comprehension, Curiosity, Critical thinking, Communication skills, Collaboration skills, Media literacy, ICT literacy, Initiative, Design capability, Leadership and Motivation	• V(TPACK) = Tk, Pk, Ck • V(Pedagogical org.) = Teamwork, Individual work, At home, At school, External • V(Benefits) = Content comprehension, Curiosity, Critical thinking, Communication skills, Collaboration skills, Media literacy, ICT literacy, Initiative, Design capability, Leadership and Motivation
Σ	The collection of educational experiences	
ρ	C-relevance, N-relevance, P-relevance	P-relevance
D	$v \in V, \begin{cases} D(v,r) = [1;10] & r = C, N \\ D(v,r) = [1;100] & r = P \end{cases}$	$v \in V, D(v, \text{P-relevance}) = \{1;0\}$

Table B.1: Data structure implementation in case study and in running repository

C

EDOC DATA

This chapter reports data collected in the EDOC@Work case study.

	TPACK															Organization					
	Delta Teacher Tk			Delt Teacher Pk			Delta Teacher Ck			Delta Student Tk			Delta Student Ck			Team work			Individual work		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_302 (Sozzi)				7	100	10				2	100	6				10	100	10			
EXP#14_306 (Artale)	4	60	5	3	20	1	0	20	0	1	50	4	7	50	3	3	100	2			
EXP#14_314 (Bossi)	3	60	5				2	40	3	3	40	5	2	60	6	5	100	7			
EXP#14_315 (Cappelli)	2	50	7				2	50	3	2	60	7	2	40	6	7	50	8	7	50	6
EXP#14_318 (Castellani)	5	30	8	5	40	5	5	30	6	3	80	8	5	20	6	2	50	5	2	50	5
EXP#14_311 (Di Camillo)	4	50	7	6	40	4		10	3	4	70	6	0	30	5	2	100	6			
EXP#14_317 (Di Ronza)	0	5	1	2	10	3	0	85	1	0	50	6	0	50	6	3	100	6			
EXP#14_309 (Gavazzi)	5	40	6	5	40	6	5	20	2	5	50	6	5	50	6	7	100	7			
EXP#14_308 (Mascape)	2	90	7	1	5	1	2	5	1	2	30	4	10	70	8	0	100	0			
EXP#14_310 (Massa)				1	50	2	1	50	6	0	100	5				1	100	4			
EXP#14_320 (Nuovo)	5	30	6	5	70	6				5	10	6	8	90	7	1	10	7	1	90	7
EXP#14_313 (Pandiani)	6	45	7	8	50	8	1	5	2	1	3	5	8	97	8	1	80	7	1	20	6
EXP#14_319 (Tria)	2	45	6	2	45	6	2	10	3	7	80	8	4	20	3	1	50	5	1	50	5
EXP#14_207 (Andriani)	3	40	6	3	20	5	3	40	8	5	50	8	5	50	8	3	100	7			
EXP#14_211 (Baccaro)	2	60	6	2	10	3	3	30	6	3	60	7	2	40	7	9	100	8			
EXP#14_217 (Bertilaccio)	1	40	6	2	30	5	1	30	5	3	50	8	3	50	6	5	50	7	5	50	6
EXP#14_216 (Bleve)	1	70	6	0	20	0	0	10	1	0	90	4	0	10	2	2	100	6			
EXP#14_201 (Calisi)	6	40	9	1	20	5	6	40	7	6	50	9	6	50	7	2	100	6			
EXP#14_215 (Casavecchia)	0	30	1	1	40	5	1	30	3	6	60	6	5	40	6	4	100	6			
EXP#14_205 (Danisi)	2	30	3	2	40	4	3	30	4	6	40	6	7	60	8	5	100	7			
EXP#14_213 (De Giorgi)	4	40	6	4	40	7	2	20	2	3	50	8	3	50	5	3	80	5	3	20	5
EXP#14_218 (De Iaco)	3	20	6	5	40	7	4	40	6	3	40	6	5	60	7	4	30	5	6	70	5
EXP#14_202 (De Nicolo)	4	40	7	4	40	4	1	20	3	3	40	6	4	60	6	3	20	3	2	80	7
EXP#14_210 (Del Prete)	5	70	6	3	15	3	1	15	1	5	90	6	2	10	1	1	20	2	1	80	4
EXP#14_204 (Di Nunzio)	3	30	5	3	40	6	3	30	6	4	50	6	4	50	6	3	100	6			
EXP#14_203 (Gatti)	5	40	6	4	40	6	1	20	2	4	50	6	4	50	6	8	100	10			
EXP#14_208 (Landi)	5	30	6	4	40	6	4	30	5	4	50	6	5	50	5	6	70	6	5	30	6
EXP#14_219 (Licchelli)	1	40	4	7	30	6	4	30	2	3	5	3	4	95	7	0	100	0			
EXP#14_214 (Martina)	2	50	3	5	50	6				5	50	6	6	50	6	3	80	5	3	20	5
EXP#14_220 (Palmieri)	2	40	6	2	30	3	2	30	4	5	50	7	5	50	7	7	100	8			
EXP#14_206 (Parisi)	3	20	5	5	40	6	6	40	7	8	50	8	8	50	9	5	60	7	5	40	6
EXP#14_209 (Pezzuto)	4	40	8	2	20	4	2	40	6	4	50	6	4	50	7	4	40	5	6	60	7

	Site									Benefits											
	At home			At school			External			Content comprehension			Curiosity			Critical thinking			Communication skills		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_302 (Sozzi)				2	60	5	2	40	4	0	20	0							6	20	2
EXP#14_306 (Artale)				3	100	2				5	10	5							0	10	0
EXP#14_314 (Bossi)				2	100	3				2	20	9	5	20	5						
EXP#14_315 (Cappelli)				0	100	0							2	30	6				2	30	6
EXP#14_318 (Castellani)				0	100	0															
EXP#14_311 (Di Camillo)	2	80	6	1	20	6															
EXP#14_317 (Di Ronza)				0	100	0				5	25	6				5	25	6	5		6
EXP#14_309 (Gavazzi)	2	60	8	2	40	7				5	10	6							6	20	5
EXP#14_308 (Mascape)				1	100	6															
EXP#14_310 (Massa)				0	100	0				1	100	6									
EXP#14_320 (Nuovo)				1	50	7	2	50	6	8	40	8									
EXP#14_313 (Pandiani)				1	100	6				8	40	8				7	40	6			
EXP#14_319 (Tria)				0	100	0				1	10	5							2	30	6
EXP#14_207 (Andriani)				5	100	7				7	25	8									
EXP#14_211 (Baccaro)	2	30	6	3	30	6	10	40	7	5	10	7									
EXP#14_217 (Bertilaccio)	4	30	6	4	70	5				4	10	6									
EXP#14_216 (Bleve)				5	50	7	6	50	7	2	20	4									
EXP#14_201 (Calisi)				2	100	6				5	15	7							5	15	9
EXP#14_215 (Casavecchia)				1	10	2	7	90	8	3	5	6							5	20	7
EXP#14_205 (Danisi)	3	30	5	3	20	5	5	50	7	5	20	6									
EXP#14_213 (De Giorgi)	4	30	7	4	30	5	6	40	6	5	10	6									
EXP#14_218 (De Iaco)	4	50	6	4	50	6				4	20	6									
EXP#14_202 (De Nicolo)	4	50	7				7	50	8	2	10	6									
EXP#14_210 (Del Prete)	2	70	6	2	30	6				1	5	3									
EXP#14_204 (Di Nunzio)				5	100	6				4	20	5				6	20	6			
EXP#14_203 (Gatti)				4	100	6				4	20	6									
EXP#14_208 (Landi)				3	100	7				5	20	7									
EXP#14_219 (Licchelli)				3	50	6	6	50	7	5	10	7	5	10	8					10	6
EXP#14_214 (Martina)	4	30	7	4	30	5	6	40	6	6	25	6									
EXP#14_220 (Palmieri)				3	100	6				4	10	5									
EXP#14_206 (Parisi)	5	40	5	5	60	6				5	20	8									
EXP#14_209 (Pezzuto)	4	50	6	4	50	6				6	30	8									

	Benefits																				
	Collaboration skills			Media literacy			ICT literacy			Initiative			Design capability			Leadership			Motivation		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_302 (Sozzi)	0	3	5	0	5	0	0	2	0										9	50	8
EXP#14_306 (Artale)	5	0	3	0	40	7	2	10	2										8	30	8
EXP#14_314 (Bossi)	5	20	8				5	20	5										5	20	9
EXP#14_315 (Cappelli)							2	30	7										2	10	6
EXP#14_318 (Castellani)	5	50	6				5	45	5										5	5	9
EXP#14_311 (Di Camillo)				7	100	6															
EXP#14_317 (Di Ronza)		25		5	25	6															
EXP#14_309 (Gavazzi)	7	20	5	5	30	6	5	10	6										2	10	9
EXP#14_308 (Mascape)										8	100	8									
EXP#14_310 (Massa)																					
EXP#14_320 (Nuovo)	6	5	6	5	10	6	5	5	6										7	40	9
EXP#14_313 (Pandiani)							2	20	7												
EXP#14_319 (Tria)	2	30	6				2	30	6												8
EXP#14_207 (Andriani)	9	25	7	3	10	7	3	15	8				3	20	6				3	5	7
EXP#14_211 (Baccaro)	6	30	7	5	15	7	5	15	7										6	30	8
EXP#14_217 (Bertilaccio)	7	50	9	5	10	6	5	10	6										6	20	7
EXP#14_216 (Bleve)	5	20	7				5	20	6				6	20	7				6	20	6
EXP#14_201 (Calisi)	6	25	7				5	15	7							3	15	6	6	15	8
EXP#14_215 (Casavecchia)	5	20	6				6	30	6				6	20	8				3	5	7
EXP#14_205 (Danisi)	4	30	7	3	10	6	2	10	7										2	30	8
EXP#14_213 (De Giorgi)	6	30	8	5	20	6	3	10	6										5	30	8
EXP#14_218 (De Iaco)	4	20	6	5	20	6	3	20	6										6	20	6
EXP#14_202 (De Nicolo)	4	30	7	3	30	6													6	30	7
EXP#14_210 (Del Prete)	1	5	5	2	45	6	2	45	6												
EXP#14_204 (Di Nunzio)				4	20	3	4	10	6				5	10	6				6	20	7
EXP#14_203 (Gatti)	6	20	6	5	20	5							6	20	6				6	20	7
EXP#14_208 (Landi)	4	20	6	4	20	6	4	20	6										7	20	8
EXP#14_219 (Licchelli)	6	40	8																5	30	8
EXP#14_214 (Martina)	6	25	7				5	25	6										6	25	7
EXP#14_220 (Palmieri)	7	30	6	5	30	6	3	10	6										4	20	6
EXP#14_206 (Parisi)	4	20	6	4	20	6	3	20	7										2	20	6
EXP#14_209 (Pezzuto)	5	10	6	4	20	6	5	10	6										4	30	8

Figure C.1: EDOC repository data (1/8)

	TPACK															Organization						Site								
	Delta Teacher Tk			Delt Teacher Pk			Delta Teacher Ck			Delta Student Tk			Delta Student Ck			Team work			Individual work			At home			At school			External		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_212 (Pico)	3	70	5	4	20	6	4	10	2				3	100	7				5	100	6				5	50	5	5	50	5
EXP#14_175 (Andreoni)	2	30	3	2	40	3	2	30	3	2	40	5	2	60	6	3	40	4	3	60	5	6	30	6	3	40	5	8	30	6
EXP#14_172 (Arosio)				1	100	3							5	100	6				2	100	5				2	90	6	3	10	5
EXP#14_116 (Baima)	2	40	5	5	60	6							3	100	6	5	100	5							5	90	5	3	10	5
EXP#14_173 (Bernardi)	0	100	0							1	50	4	1	50	3	1	20	1	7	80	1	1	80	2	1	20	2			
EXP#14_101 (Bidoia)				2	100	4				2	50	5	3	50	5	4	100	5							2	100	5			
EXP#14_118 (Bosoni)	2	50	4	2	50	3				4	50	6	4	50	6	3	100	3				4	40	6	4	40	6	4	20	4
EXP#14_156 (Buzzi)	2	20	3	4	80	5				1	20	2	7	80	7	0	100	0							0	100	0			
EXP#14_106 (Campanale)	3	40	6	3	20	3	4	40	6	3	60	6	3	40	5	3	50	5	3	50	5	6	50	7	3	50	5			
EXP#14_157 (Capalbo)							4	100	6				4	100	7	4	50	5	4	50	4	3	30	5	3	30	5	3	40	5
EXP#14_115 (Carnevale)	2	30	4	3	40	3	2	30	6	1	10	1	4	90	6	7	80	8	4	20	6	6	30	4	7	40	5	6	30	5
EXP#14_153 (Carta)	0	100	0							4	30	5	6	70	7	2	100	5							2	50	4	4	50	3
EXP#14_122 (Casero)	2	50	5	2	50	2				3	30	3	4	70	6	4	100	5							4	100	5			
EXP#14_160 (Casettari)	1	50	2	1	50	2				4	40	4	5	60	6	8	100	6							7	70	6	4	30	7
EXP#14_135 (Celotto)	1	30	1	1	40	1	1	30	1	3	50	5	2	50	5	5	50	5	5	50	5	4	50	6	4	50	6			
EXP#14_129 (Cerchi)	4	60	7	4	40	6							7	100	9	7	90	8	7	10	7				4	100	7			
EXP#14_131 (Chirico)	4	50	6	8	50	7				3	50	6	6	50	6	4	100	6				8	30	6	5	70	6			
EXP#14_137 (Cimadomo)	4	50	4	6	50	6							4	100	8	3	50	5	3	50	5				4	100	5			
EXP#14_158 (Concina)	0	100	0							9	50	9	9	50	9	5	80	8	4	20	4				4	80	7	7	20	8
EXP#14_166 (D'aria)	0	100	0							4	100	6				4	50	4	4	50	4	4	50	4	4	50	4			
EXP#14_174 (D'Armi)	3	50	6	3	50	2				3	50	6	4	50	6	7	100	7							4	100	5			
EXP#14_109 (D'Oria)	4	40	5	6	60	6				4	50	5	2	50	5	4	100	5							6	70	6	4	30	5
EXP#14_120 (De Bellis)	0	100	0										0	100	5	4	50	5	4	50	5	4	50	4	4	50	4			
EXP#14_108 (De Luca)	2	50	6	3	50	6							3	100	6				1	100	5	3	20	4	3	40	4	3	40	5
EXP#14_136 (Del Vecchio)	3	50	5	3	50	5				3	50	6	5	50	6				2	100	4	1	30	3	1	70	3			
EXP#14_142 (Delalle)	1	50	6	2	50	6				1	70	6	3	30	5	9	50	4	5	50	3	8	90	5	2	10	3			
EXP#14_112 (D'Eletto)	1	30	6	7	40	5	7	30	5	2	50	6	2	50	5	6	80	6	4	20	4				4	30	4	8	70	8
EXP#14_119 (Faroldi)	4	40	6	4	60	8				4	50	6	3	50	6	3	50	5	3	50	5	3	50	3	3	50	3			
EXP#14_105 (Fedi)							4	100	6	4	50	7	4	50	7	5	50	7	7	50	6				7	50	8	8	50	6
EXP#14_154 (Fois)	3	50	6	3	50	6				3	50	6	3	50	6	4	100	4							4	100	4			
EXP#14_163 (Garofano)	3	70	7	3	30	6				8	50	8	5	50	6	6	100	6							6	50	6	8	50	9
EXP#14_126 (Gatto)				3	100	5				5	50	6	5	50	5				6	100	6	3	50	4	3	50	4			

	Benefits																																
	Content comprehension			Curiosity			Critical thinking			Communication skills			Collaboration skills			Media literacy			ICT literacy			Initiative			Design capability			Leadership			Motivation		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_212 (Pico)	4	50	8																												7	50	8
EXP#14_175 (Andreoni)	4	20	6							3	20	6				3	20	6	3	10	6							4	10	6	6	20	6
EXP#14_172 (Arosio)	3	30	6													2	30	5													5	40	6
EXP#14_116 (Baima)																1	100	4															
EXP#14_173 (Bernardi)													1	25	4				1	25	4										1	50	5
EXP#14_101 (Bidoia)	3	15	5										5	15	7				3	5	5							4	25	6	3	40	7
EXP#14_118 (Bosoni)		30	7							2	10	5	4	10	6	3	10	5	6	30	7										7	10	6
EXP#14_156 (Buzzi)	2	80	8																												2	20	7
EXP#14_106 (Campanale)	5	20	7							7	50	6							5	30	8												
EXP#14_157 (Capalbo)	5	30	6							5	30	6	4	20	5																6	20	6
EXP#14_115 (Carnevale)	5	20	6							6	40	7							2	10	3										2	30	5
EXP#14_153 (Carta)	6	30	6							6	20	7							6	20	7										10	30	9
EXP#14_122 (Casero)			6							5	40	6							3	20	4										5	40	5
EXP#14_160 (Casettari)										4	50	6																			3	50	6
EXP#14_135 (Celotto)													6	50	6				3	50	4												
EXP#14_129 (Cerchi)	8	50	8							7	50	7																					
EXP#14_131 (Chirico)										4	30	8	5	30	6				4	30	6										3	10	6
EXP#14_137 (Cimadomo)	5	50	4													4	30	5													5	20	6
EXP#14_158 (Concina)	6	25	7													6	25	7	5	25	6										5	25	7
EXP#14_166 (D'aria)													4	25	6	4	25	6	4	25	6										3	25	6
EXP#14_174 (D'Armi)	5	40	8													4	40	6	3	20	6												
EXP#14_109 (D'Oria)	6	30	7								30	6							3	20	3										4	20	6
EXP#14_120 (De Bellis)	1	100	4																														
EXP#14_108 (De Luca)	4	100	4																														
EXP#14_136 (Del Vecchio)	5	25	6							3	25	6							3	25	6										6	25	7
EXP#14_142 (Delalle)	4	25	6							5	25	6	7	25	7				2	25	5												
EXP#14_112 (D'Eletto)	2	10	6										4	30	6	3	60	8															
EXP#14_119 (Faroldi)	6	30	6							4	10	6							5	30	6										4	30	6
EXP#14_105 (Fedi)	6	30	8										6	20	6				3	20	5				4	10	6	3	10	6	3	10	7
EXP#14_154 (Fois)	5	30	6							7	40	8							7	20	7										3	10	7
EXP#14_163 (Garofano)	8	35	7										10	35	7	3	10	6		10	6											10	
EXP#14_126 (Gatto)	3	40	6																3	30	5										3	30	4

Figure C.2: EDOC repository data (2/8)

TPACK	Delta Teacher Tk			Delt Teacher Pk			Delta Teacher Ck			Delta Student Tk			Delta Student Ck		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_107 (Ghezzi)				2	100	2				3	100	6			
EXP#14_114 (Grimaldi)	2	70	6	2	30	5				5	70	6	3	30	5
EXP#14_143 (Gualdandri)				2	100	5				3	40	5	4	60	8
EXP#14_150 (Lambertini)	1	30	2	1	35	2	1	35	2	5	50	6	5	50	6
EXP#14_165 (Lattanzio)	1	50	1	1	50	1				5	50	6	4	50	6
EXP#14_124 (Laudano)				1	100	1				0	100	0			
EXP#14_102 (Lo Monaco)	4	100	6							6	40	7	8	60	8
EXP#14_148 (Lorini)	1	40	1	1	30	1	1	30	1	7	100	7			
EXP#14_164 (Maione)	4	100	6										4	100	6
EXP#14_167 (Mancin)	0	100	0							5	50	7	5	50	7
EXP#14_146 (Mancuso Malaspina)							6	100	7	5	30	8	7	70	7
EXP#14_117 (Mangano)	6	100	6							6	50	6	5	50	6
EXP#14_179 (Mariani)	0	100	0							0	100	0			
EXP#14_141 (Miracoli)	5	50	6	5	50	6				3	50	6	5	50	6
EXP#14_170 (Montepiani)	0	100	0							0	100	0			
EXP#14_171 (Morandi)	6	100	7										3	100	6
EXP#14_176 (Pagano)	1	10	1	5	90	6				6	60	8	4	40	6
EXP#14_178 (Pagliarulo)	4	50	5	5	50	2				7	50	7	6	50	6
EXP#14_110 (Palmieri)	7	50	8	8	50	8				5	30	6	6	70	6
EXP#14_127 (Passoni)	6	60	6	3	40	2				6	50	7	6	50	6
EXP#14_123 (Pavan)				1	100	6				2	10	1	8	90	7
EXP#14_162 (Pegorer)	1	100	2							4	50	7	4	50	6
EXP#14_139 (Perrone)	8	60	6	4	40	5				6	50	7	6	50	7
EXP#14_159 (Piras)	3	30	3	3	30	2	4	40	6	5	50	5	7	50	6
EXP#14_152 (Pistolesi)	2	50	2	3	50	3				7	40	7	8	60	8
EXP#14_177 (Polizzotto)	5	50	7	7	50	7				7	60	7	5	40	6
EXP#14_147 (Porrello)	3	50	6	5	50	6				2	20	1	4	80	6
EXP#14_130 (Puccio)	3	50	3	4	50	4				2	20	3	6	80	7
EXP#14_169 (Pulcini)	4	50	6	4	50	5				4	50	6	6	50	6
EXP#14_138 (Romanelli)	2	50	2	4	50	5				2	50	5	6	50	7
EXP#14_161 (Rosso)	4	50	6	6	50	6				4	50	6	6	50	6
EXP#14_128 (Rudello)	6	60	6	6	40	6							6	100	6

	Organization: Team work			Organization: Individual work			Site: At home			Site: At school			Site: External		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_107 (Ghezzi)	4	70	8	3	30	3				7	50	7	7	50	6
EXP#14_114 (Grimaldi)	7	100	8							0	100	0			
EXP#14_143 (Gualdandri)	8	80	7	3	20	5				0	100	0			
EXP#14_150 (Lambertini)	5	50	4	5	50	4	6	50	4	4	50	4			
EXP#14_165 (Lattanzio)	5	100	6							6	50	6	5	50	6
EXP#14_124 (Laudano)	5	100	5							5	100	5			
EXP#14_102 (Lo Monaco)	7	100	7				7	40	7	5	20	6	8	40	8
EXP#14_148 (Lorini)				4	100	6				6	70	6	7	30	7
EXP#14_164 (Maione)	5	60	6	5	40	6				3	100	5			
EXP#14_167 (Mancin)	7	100	7							0	100	0			
EXP#14_146 (Mancuso Malaspina)	5	50	6	5	50	6				5	100	6			
EXP#14_117 (Mangano)	6	100	6				8	80	8	4	20	5			
EXP#14_179 (Mariani)	5	100	5				5	70	5	5	30	5			
EXP#14_141 (Miracoli)	6	100	8							5	70	6	7	30	6
EXP#14_170 (Montepiani)	4	100	6							4	100	5			
EXP#14_171 (Morandi)				3	100	5				2	100	6			
EXP#14_176 (Pagano)	5	50	6	5	50	6	6	30	7	6	40	6	7	30	7
EXP#14_178 (Pagliarulo)	5	95	5	5	5	6				4	100	6			
EXP#14_110 (Palmieri)	4	40	7	6	60	7				4	50	6	9	50	8
EXP#14_127 (Passoni)	6	50	6	6	50	6							9	100	7
EXP#14_123 (Pavan)	5	90	6	6	10	6	5	10	5	5	90	6			
EXP#14_162 (Pegorer)	6	30	4	6	70	6				6	100	6			
EXP#14_139 (Perrone)	7	100	7							7	100	7			
EXP#14_159 (Piras)	4	40	6	5	60	6	5	20	6	6	80	6			
EXP#14_152 (Pistolesi)	6	100	6							3	100	6			
EXP#14_177 (Polizzotto)	6	100	7							5	100	6			
EXP#14_147 (Porrello)	5	100	6							6	100	5			
EXP#14_130 (Puccio)	3	100	4							4	50	6	4	50	6
EXP#14_169 (Pulcini)	6	100	6							5	50	5	6	50	6
EXP#14_138 (Romanelli)	5	50	6	6	50	3				5	100	6			
EXP#14_161 (Rosso)	4	50	6	4	50	6	5	10	5	5	90	5			
EXP#14_128 (Rudello)	8	100	8							5	100	5			

Benefits	Content comprehension			Curiosity			Critical thinking			Communication skills			Collaboration skills		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_107 (Ghezzi)	5	15	6							8	40	6			
EXP#14_114 (Grimaldi)	4	25	8										6	25	8
EXP#14_143 (Gualdandri)	3	25	6							5	25	6			
EXP#14_150 (Lambertini)	2	20	6							4	20	5	6	20	6
EXP#14_165 (Lattanzio)	4	30	6												
EXP#14_124 (Laudano)															
EXP#14_102 (Lo Monaco)	6	25	7							3	15	5	7	15	7
EXP#14_148 (Lorini)	7	60	7												
EXP#14_164 (Maione)	5	30	6							7	40	6			
EXP#14_167 (Mancin)	6	30	8							8	30	8			
EXP#14_146 (Mancuso Malaspina)	6	30	7							5	30	6			
EXP#14_117 (Mangano)	5	20	6							6	20	6	8	20	6
EXP#14_179 (Mariani)	6	30	5												
EXP#14_141 (Miracoli)	4	20	5							6	30	6			
EXP#14_170 (Montepiani)													5	20	6
EXP#14_171 (Morandi)	6	50	6							6	50	6			
EXP#14_176 (Pagano)	4	25	6							6	20	6	6	25	6
EXP#14_178 (Pagliarulo)	5	25	8										3	25	6
EXP#14_110 (Palmieri)	4	30	6							6	30	6	7	30	8
EXP#14_127 (Passoni)	4	25	6							4	25	6			
EXP#14_123 (Pavan)	4	100	6												
EXP#14_162 (Pegorer)	5	25	7							6	25	7			
EXP#14_139 (Perrone)	7	25	8							7	25	7			
EXP#14_159 (Piras)	7	25	8										3	25	6
EXP#14_152 (Pistolesi)	7	25	7										5	25	6
EXP#14_177 (Polizzotto)	4	25	6										4	25	6
EXP#14_147 (Porrello)	6	30	7							4	40	6			
EXP#14_130 (Puccio)	4	20	5							4	20	6	4	20	6
EXP#14_169 (Pulcini)	6	25	6							5	25	6			
EXP#14_138 (Romanelli)	6	20	6							6	20	6	6	20	6
EXP#14_161 (Rosso)	6	20	6							7	20	6	5	10	6
EXP#14_128 (Rudello)	6	45	6							6	45	6			

Benefits	Media literacy			ICT literacy			Initiative			Design capability			Leadership			Motivation		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_107 (Ghezzi)				4	15	6				4	15	6	4	15	6			
EXP#14_114 (Grimaldi)				5	25	8										8	25	6
EXP#14_143 (Gualdandri)				4	25	6										4	25	6
EXP#14_150 (Lambertini)				4	20	6										5	20	6
EXP#14_165 (Lattanzio)				3	30	6										4	40	6
EXP#14_124 (Laudano)																3	100	5
EXP#14_102 (Lo Monaco)				5	25	6										6	20	7
EXP#14_148 (Lorini)				3	20	5										4	20	6
EXP#14_164 (Maione)																3	30	6
EXP#14_167 (Mancin)	5	20	6	5	10	6										4	10	7
EXP#14_146 (Mancuso Malaspina)				5	10	6										5	30	8
EXP#14_117 (Mangano)				6	20	6										6	20	6
EXP#14_179 (Mariani)	5	40	5													3	30	6
EXP#14_141 (Miracoli)				6	30	6										5	20	6
EXP#14_170 (Montepiani)				5	20	6				5	20	6	5	20	6	5	20	6
EXP#14_171 (Morandi)																		
EXP#14_176 (Pagano)										5	10	6	5	10	6	4	10	6
EXP#14_178 (Pagliarulo)				6	25	6										4	25	6
EXP#14_110 (Palmieri)				2	10	3												
EXP#14_127 (Passoni)				5	25	6										3	25	5
EXP#14_123 (Pavan)																		
EXP#14_162 (Pegorer)				5	25	7										7	25	7
EXP#14_139 (Perrone)				6	25	8										4	25	6
EXP#14_159 (Piras)				3	25	2										4	25	6
EXP#14_152 (Pistolesi)	3	25	5													5	25	6
EXP#14_177 (Polizzotto)				5	25	6										5	25	6
EXP#14_147 (Porrello)																6	30	6
EXP#14_130 (Puccio)				4	20	5										3	20	6
EXP#14_169 (Pulcini)				6	25	6										5	25	6
EXP#14_138 (Romanelli)				4	20	6										3	20	3
EXP#14_161 (Rosso)										5	20	6	5	20	6	2	10	2
EXP#14_128 (Rudello)																1	10	1

Figure C.3: EDOC repository data (3/8)

	TPACK														
	Delta Teacher Tk			Delt Teacher Pk			Delta Teacher Ck			Delta Student Tk			Delta Student Ck		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_155 (Sacchi)	5	100	6							5	30	6	5	70	6
EXP#14_134 (Sai)				1	100	2				3	100	3	4		5
EXP#14_121 (Sampaolesi)				8	100	8				1	20	5	4	80	5
EXP#14_168 (Sanseviero)	4	50	6	5	50	6				4	50	6	5	50	6
EXP#14_132 (Santanicola)	1	100	2							1	50	2	2	50	3
EXP#14_103 (Scuppa)	6	100	6							6	50	6	6	50	6
EXP#14_113 (Severini)	3	100	4							3	50	3	4	50	5
EXP#14_140 (Sorge)	4	50	6	3	50	3				5	50	6	6	50	7
EXP#14_125 (Tarquini)	1	30	2	4	35	2	6	35	6	6	50	6	6	50	6
EXP#14_133 (Toscano)	6	60	6	5	40	5				6	50	6	6	50	6
EXP#14_151 (Tranchina)	7	50	6	7	50	7				6	50	6	6	50	6
EXP#14_180 (Tribuno)				1	100	2				7	50	7	8	50	9
EXP#14_145 (Troia)	6	50	6	4	50	6				6	50	6	6	50	6
EXP#14_104 (Trombetti)	1	50	1	7	50	7				6	50	6	2	50	2
EXP#14_144 (Trotta)	1	50	1	1	50	2							6	100	6
EXP#14_149 (Zamagni)	1	10	1	7	90	7				6	50	6	6	50	6
EXP#14_111 (Zardini Lacadelli)	6	60	8	5	40	5				1	10	2	6	90	6

	Organization						Site								
	Team work			Individual work			At home			At school			External		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_155 (Sacchi)	8	50	8	5	50	6	4	50	7	4	50	6			
EXP#14_134 (Sai)	5	100	5				5	80	6	7	20	6			
EXP#14_121 (Sampaolesi)	7	100	6							6	100	6			
EXP#14_168 (Sanseviero)	7	100	6				6	20	6	6	80	6			
EXP#14_132 (Santanicola)	7	100	7							7	40	7	8	60	8
EXP#14_103 (Scuppa)	6	100	6							6	100	6			
EXP#14_113 (Severini)	4	100	7				2	100	5						
EXP#14_140 (Sorge)	5	30	6	5	70	6	5	50	6	5	50	6			
EXP#14_125 (Tarquini)				5	100	5				5	100	5			
EXP#14_133 (Toscano)	5	100	5							5	100	5			
EXP#14_151 (Tranchina)	5	50	5	5	50	5				5	50	5	8	50	8
EXP#14_180 (Tribuno)	6	100	6							6	100	6			
EXP#14_145 (Troia)	6	50	6	6	50	6				5	100	5			
EXP#14_104 (Trombetti)	5	100	5							5	100	5			
EXP#14_144 (Trotta)	6	50	6	6	50	6	6	50	7	6	50	7			
EXP#14_149 (Zamagni)	5	100	5							5	100	5			
EXP#14_111 (Zardini Lacadelli)				5	100	5	5	50	5	5	50	5			

	Benefits																	
	Content comprehension			Curiosity			Critical thinking			Communication skills			Collaboration skills			Media literacy		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_155 (Sacchi)	7	30	7							7	20	6						
EXP#14_134 (Sai)										3	10	2	8	50	8			
EXP#14_121 (Sampaolesi)	5	20	6							5	20	6	5	20	6			
EXP#14_168 (Sanseviero)	6	25	6							5	25	6				6	25	6
EXP#14_132 (Santanicola)													4	50	6			
EXP#14_103 (Scuppa)	6	10	6							6	10	7	8	30	8			
EXP#14_113 (Severini)										5	35	6	4	35	6			
EXP#14_140 (Sorge)	6	10	6							7	30	7	7	30	7			
EXP#14_125 (Tarquini)	6	25	6							6	25	6						
EXP#14_133 (Toscano)	6	30	6							7	35	6						
EXP#14_151 (Tranchina)	6	5	6							7	30	7	9	30	9			
EXP#14_180 (Tribuno)	6	7	6							8	27	8	8	28	8			
EXP#14_145 (Troia)	5	25	6							5	25	6						
EXP#14_104 (Trombetti)	6	20	6							6	20	6						
EXP#14_144 (Trotta)																		
EXP#14_149 (Zamagni)	6	25	6													6	25	6
EXP#14_111 (Zardini Lacadelli)	6	40	6							6	40	6						

	Benefits														
	ICT literacy			Initiative			Design capability			Leadership			Motivation		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
EXP#14_155 (Sacchi)	5	30	6										6	20	7
EXP#14_134 (Sai)	3	20	3										3	20	3
EXP#14_121 (Sampaolesi)	2	20	2										5	20	6
EXP#14_168 (Sanseviero)													8	25	8
EXP#14_132 (Santanicola)													6	50	6
EXP#14_103 (Scuppa)							5	10	6	5	10	6	8	30	9
EXP#14_113 (Severini)	1	10	1										4	20	6
EXP#14_140 (Sorge)	4	20	6										7	10	2
EXP#14_125 (Tarquini)	6	25	6										5	25	5
EXP#14_133 (Toscano)	6	35	6												
EXP#14_151 (Tranchina)	6	5	6										8	30	8
EXP#14_180 (Tribuno)	8	28	8										6	10	6
EXP#14_145 (Troia)	5	25	6										5	25	6
EXP#14_104 (Trombetti)	4	20	6				9	20	9				6	20	6
EXP#14_144 (Trotta)													4	100	4
EXP#14_149 (Zamagni)	5	25	5										5	25	7
EXP#14_111 (Zardini Lacadelli)	2	20	3												

Figure C.4: EDOC repository data (4/8)

	School level
EXP#14_302 (Sozzi)	Primaria
EXP#14_306 (Artale)	Primaria
EXP#14_314 (Bossi)	Primaria
EXP#14_315 (Cappelli)	Liceo Artistico
EXP#14_318 (Castellani)	Primaria
EXP#14_311 (Di Camillo)	Secondaria 2° grado
EXP#14_317 (Di Ronza)	Primaria
EXP#14_309 (Gavazzi)	Secondaria 2° grado
EXP#14_308 (Mascape)	Infanzia
EXP#14_310 (Massa)	Secondaria 2° grado
EXP#14_320 (Nuovo)	Primaria
EXP#14_313 (Pandiani)	Primaria
EXP#14_319 (Tria)	Secondaria 1° grado
EXP#14_207 (Andriani)	Secondaria 2° grado
EXP#14_211 (Baccaro)	Secondaria 1° grado
EXP#14_217 (Bertilaccio)	Secondaria 1° grado
EXP#14_216 (Bleve)	Secondaria 2° grado
EXP#14_201 (Calisi)	Primaria
EXP#14_215 (Casavecchia)	Secondaria 2° grado
EXP#14_205 (Danisi)	Primaria
EXP#14_213 (De Giorgi)	Primaria
EXP#14_218 (De Iaco)	Secondaria 1° grado
EXP#14_202 (De Nicolo)	Primaria
EXP#14_210 (Del Prete)	Secondaria 2° grado
EXP#14_204 (Di Nunzio)	Secondaria 1° grado
EXP#14_203 (Gatti)	Primaria
EXP#14_208 (Landi)	Primaria
EXP#14_219 (Licchelli)	Secondaria 1° grado
EXP#14_214 (Martina)	Primaria
EXP#14_220 (Palmieri)	Primaria
EXP#14_206 (Parisi)	Secondaria 1° grado
EXP#14_209 (Pezzuto)	Primaria

(a) 5/8

	School level
EXP#14_212 (Pico)	Infanzia
EXP#14_175 (Andreoni)	Primaria
EXP#14_172 (Arosio)	Infanzia
EXP#14_116 (Baima)	Infanzia
EXP#14_173 (Bernardi)	Secondaria 2° grado
EXP#14_101 (Bidoia)	Secondaria 1° grado
EXP#14_118 (Bosoni)	Secondaria 1° grado
EXP#14_156 (Buzzi)	Infanzia
EXP#14_106 (Campanale)	Primaria
EXP#14_157 (Capalbo)	Primaria
EXP#14_115 (Carnevale)	Secondaria 1° grado
EXP#14_153 (Carta)	Infanzia
EXP#14_122 (Casero)	Secondaria 1° grado
EXP#14_160 (Casettari)	Secondaria 1° grado
EXP#14_135 (Celotto)	Secondaria 2° grado
EXP#14_129 (Cerchi)	Infanzia
EXP#14_131 (Chirico)	Secondaria 1° grado
EXP#14_137 (Cimadomo)	Primaria
EXP#14_158 (Concina)	Secondaria 2° grado
EXP#14_166 (D'aria)	Secondaria 2° grado
EXP#14_174 (D'Armi)	Secondaria 1° grado
EXP#14_109 (D'Oria)	Secondaria 2° grado
EXP#14_120 (De Bellis)	Secondaria 2° grado
EXP#14_108 (De Luca)	Primaria
EXP#14_136 (Del Vecchio)	Secondaria 2° grado
EXP#14_142 (Delalle)	Secondaria 2° grado
EXP#14_112 (D'Eletto)	Primaria
EXP#14_119 (Faroldi)	Secondaria 1° grado
EXP#14_105 (Fedi)	Primaria
EXP#14_154 (Fois)	Primaria
EXP#14_163 (Garofano)	Secondaria 2° grado
EXP#14_126 (Gatto)	Secondaria 1° grado

(b) 6/8

Figure C.5: EDOC repository data

	School level
EXP#14_107 (Ghezzi)	Secondaria 1° grado
EXP#14_114 (Grimaldi)	Secondaria 1° grado
EXP#14_143 (Gualdandri)	Secondaria 2° grado
EXP#14_150 (Lambertini)	Primaria
EXP#14_165 (Lattanzio)	Primaria
EXP#14_124 (Laudano)	Secondaria 1° grado
EXP#14_102 (Lo Monaco)	Primaria
EXP#14_148 (Lorini)	Infanzia
EXP#14_164 (Maione)	Infanzia
EXP#14_167 (Mancin)	Secondaria 2° grado
EXP#14_146 (Mancuso Malaspina)	Infanzia
EXP#14_117 (Mangano)	Secondaria 2° grado
EXP#14_179 (Mariani)	Secondaria 1° grado
EXP#14_141 (Miracoli)	Secondaria 1° grado
EXP#14_170 (Montepiani)	Secondaria 2° grado
EXP#14_171 (Morandi)	Infanzia
EXP#14_176 (Pagano)	Secondaria 1° grado
EXP#14_178 (Pagliarulo)	Primaria
EXP#14_110 (Palmieri)	Infanzia
EXP#14_127 (Passoni)	Infanzia
EXP#14_123 (Pavan)	Infanzia
EXP#14_162 (Pegorer)	Secondaria 1° grado
EXP#14_139 (Perrone)	Secondaria 2° grado
EXP#14_159 (Piras)	Primaria
EXP#14_152 (Pistolesi)	Secondaria 2° grado
EXP#14_177 (Polizzotto)	Primaria
EXP#14_147 (Porrello)	Secondaria 2° grado
EXP#14_130 (Puccio)	Primaria
EXP#14_169 (Pulcini)	Secondaria 1° grado
EXP#14_138 (Romanelli)	Primaria
EXP#14_161 (Rosso)	Secondaria 1° grado
EXP#14_128 (Rudello)	Primaria

(a) 7/8

	School level
EXP#14_155 (Sacchi)	Primaria
EXP#14_134 (Sai)	Secondaria 2° grado
EXP#14_121 (Sampaolesi)	Secondaria 1° grado
EXP#14_168 (Sanseviero)	Secondaria 1° grado
EXP#14_132 (Santanicola)	Secondaria 2° grado
EXP#14_103 (Scuppa)	Secondaria 1° grado
EXP#14_113 (Severini)	Secondaria 1° grado
EXP#14_140 (Sorge)	Secondaria 2° grado
EXP#14_125 (Tarquini)	Secondaria 1° grado
EXP#14_133 (Toscano)	Secondaria 1° grado
EXP#14_151 (Tranchina)	Secondaria 2° grado
EXP#14_180 (Tribuno)	Infanzia
EXP#14_145 (Troia)	Primaria
EXP#14_104 (Trombetti)	Secondaria 1° grado
EXP#14_144 (Trotta)	Secondaria 2° grado
EXP#14_149 (Zamagni)	Secondaria 1° grado
EXP#14_111 (Zardini Lacadelli)	Secondaria 2° grado

(b) 8/8

Figure C.6: EDOC repository data

D

MUSEUM DATA

This chapter reports data collected in the Museum of Musical Instrument case study.

	KEYBOARD MATERIALS									KEYBOARD MATERIALS									KEYBOARD MATERIALS									KEYBOARD MATERIALS					
	ebony			fir			maple			mahogany			cypres			walnut			beech			rosewood			poplar			linden			boxwood		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
447. Clavicembalo				3	35	4	2	5	2										4	25	1										2	30	5
448. Clavicembalo	5	10	5				1	10	6							10	25	8	6	25	5												
449. Clavicembalo	8	5	5	1	35	6							1	4	4	2	35	5	3	5	7										3	5	6
450. Clavicembalo	9	10	7	6	5	5							3	10	5	7	10	6	6	40	4							4	10	5	7	10	7
451. Clavicembalo	2	10	7	8	10	8										6	15	6	4	40	2												
452. Clavicembalo in miniatura				7	90	3																									5	10	6
453. Clavicembalo	7	5	7	3	35	6							8	10	6				1	10	6				5	30	8						
454. Clavicembalo convertito in pianoforte	10	5	7	7	15	8										6	45	5							2	25	3						
455. Clavicembalo	7	10	6	1	90	4																											
456. Clavicembalo																												4	30	5			
457. Clavicembalo	7	20	4	5	50	8																									1	30	2
458. Virginale poligonale				1	10	7							6	30	5	5	10	5													10	3	3
459. Virginale poligonale	6	10	6	8	25	3	6	5	4										8	8	5							6	25	7	2	5	7
460. Virginale doppio				7	3	1													2	10	9				3	85	4						
461. Virginale rettangolare				10	10	2							9	40	8	6	20	5	4	10	5										3	15	3
462. Virginale ottavino	7	5	4	1	15	4	2	40	5				6	10	7	5	10	4													3	5	3
463. Virginale ottavino				6	45	5										9	5	2	8	30	5										5	10	4
464. Virginale pentagonale	1	5	7	7	20	4										1	20	8	8	40	7										10	5	4
465. Virginale rettangolare				4	40	5										9	40	9													6	10	6
466. Virginale in miniatura													9	35	7	10	40	6													1	10	3
467. Virginale pentagonale				3	25	7	7	7	7							9	13	9	4	25	1				2	10	9				1	5	4
468. Virginale rettangolare	1	10	6	9	50	6				8	10	8																					
469. Spinetta traversa				8	8	7				9	10	7				5	30	6	7	30	9	3	3	2	4	5	4						
470. Spinetta ottavina				8	15	7	5	5	6				8	15	6	9	10	6	7	5	1				4	5	6	7	40	7	7	5	6
471. Spinetta traversa				9	30	7										2	8	8	6	40	2										4	5	6
472. Spinetta. traversa																1	34	8													1	33	5
473. Spinetta traversa				7	30	6							7	5	6	3	30	7				7	7	3									
474. Spinetta traversa							1	100	1																								
475. Spinetta traversa	8	5	7	8	35	3										7	35	6															
476. Clavicordo libero	6	10	6	3	60	5																											
477. Clavicordo libero	4	15	7																									10	60	6	6	20	5
478. Pianoforte a tavolo	10	15	7																									4	50	6			
479. Pianoforte a tavolo	2	7	6	4	40	7										6	33	6															
480. Pianoforte a tavolo	8	10	5							9	35	8																6	40	6			
481. Pianoforte o tavolo ottavi no										2	70	1																					
482. Pianoforte a tavolo	9	10	6							4	35	6																5	35	5			
483. Pianoforte a tavolo	1	10	5				10	10	6	6	30	6																6	30	6			
484. Pianoforte a tavolo	9	5	6							5	35	5																9	50	4			
485. Pianoforte a tavolo																									7	70	2						
486. Pianoforte (l tavolo	8	5	6	3	40	7																											

				KEYBOARD MATERIALS									KEYBOARD MATERIALS									KEYBOARD MATERIALS									KEYBOARD MATERIALS											
	chestnut			durmast			oak			fruit			tin			copper			steel			lead			brass			iron			bone			felt			leather			ivory		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
447. Clavicembalo																																					2	5	6			
448. Clavicembalo										9	10	9																4	10	6										2	10	6
449. Clavicembalo																																		3	3	5	7	3	6	7	5	6
450. Clavicembalo																															3	5	6									
451. Clavicembalo																									4	15	6				4	10	5									
452. Clavicembalo in miniatura																																										
453. Clavicembalo																									5	5	7				6	5	6									
454. Clavicembalo convertito in pianoforte																																					4	5	6	1	5	6
455. Clavicembalo																																										
456. Clavicembalo							9	50	6																6	5	5										2	15	6			
457. Clavicembalo																																										
458. Virginale poligonale							2	3	4	8	30	8													6	10	5	3	4	2												
459. Virginale poligonale										7	10	6																6	7	6				7	5	6						
460. Virginale doppio																																					9	2	6			
461. Virginale rettangolare																															9	5	5									
462. Virginale ottavino										8	15	6																														
463. Virginale ottavino										1	10	6																														
464. Virginale pentagonale																																		10	10	6						
465. Virginale rettangolare																																		4	10	5						
466. Virginale in miniatura										1	10	6													2	5	6															
467. Virginale pentagonale										9	7	7													7	6	5										3	2	6			
468. Virginale rettangolare																									10	20	6				8	10	6									
469. Spinetta traversa							9	7	7																			5	5	6	7	2	5									
470. Spinetta ottavina																																										
471. Spinetta traversa							5	5	7																4	8	6										4	4	5			
472. Spinetta. traversa																															8	33	6									
473. Spinetta traversa										1	7	7																5	10	6	6	6	4							10	5	7
474. Spinetta traversa																																										
475. Spinetta traversa										8	5	8													7	5	5	8	5	6	2	5	6				7	5	5			
476. Clavicordo libero																									4	15	7				10	10	6				3	5	4			
477. Clavicordo libero																									5	5	8															
478. Pianoforte a tavolo																																		9	10	5	7	10	5	1	15	7
479. Pianoforte a tavolo										5	7	6																			9	5	6				5	8	6			
480. Pianoforte a tavolo																																					2	5	5	6	10	5
481. Pianoforte o tavolo ottavi no																															5	15	5				8	15	6			
482. Pianoforte a tavolo																																					1	10	5	9	10	7
483. Pianoforte a tavolo																																					9	10	5	7	10	6
484. Pianoforte a tavolo																															4	5	6				1	5	6			
485. Pianoforte a tavolo																															1	15	6				8	15	5			
486. Pianoforte (l tavolo										8	40	8																						2	5	5	6	5	5	2	5	6

Figure D.1: Museum repository data (1/6)

	KEYBOARD MATERIALS									KEYBOARD MATERIALS								
	ebony			fir			maple			mahogany			cypres			walnut		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
486. Pianoforte (l tavolo	8	5	6	3	40	7												
487. Pianoforte a tavolo	1	10	4	2	30	6												
488. Pianino in forma di mobile per toeletta				8	80	4												
489. Pianoforte o tavolo	3	5	7	6	35	3	10	35	1									
490. Pianoforte a tavolo				10	30	5	6	25	5							10	5	7
491. Pianoforte a tavolo	2	5	3				5	10	8	10	25	9						
492. Pianoforte a coda	9	10	9	8	30	7												
493. Pianoforte a coda				3	30	2												
494. Pianoforte verticale a forma di giraffa	1	5	5	6	35	7												
495. Pianoforte a tavolo informa d'ala	5	5	5	2	70	7												
496. Pianoforte u tavolo	2	5	3							8	35	7						
497. Pianoforte a coda	6	25	6															
498. Pianoforte a tavolo	8	20	7				4	10	7	9	30	6						
499. Pianoforte verticale	5	5	5	6	40	4												
500. Pianoforte a tavolo				1	35	2	3	15	7									
501. Pianoforte a tavolo	5	5	6															
502. Pianoforte a tavolo esagonale	5	5	7															
503. Pianoforte a tavolo				7	25	8												
504. Pianoforte a coda	1	12	8															
505. Pianoforte verticale	5	15	5															
506. Pianoforte a coda	5	15	6															
507. Pianoforte verticale	1	20	8															
508. Pianoforte a coda	2	15	3															
509. Pianoforte a coda	2	40	8															
510. Pianoforte verticale	4	40	6															
511. Pianoforte verticale	9	20	7															
512. Pianoforte verticale informa di armadio																		
513. Pianoforte verticale	7	40	7															
514. Pianoforte verticale	2	40	7															
515. Pianoforte a coda	10	20	6															
516. Pianoforte a coda																		
518. Tastiera da digitazione	7	50	5															
519. Organo	10	25	8															
520. Organo processionale																		
521. Organo positivo	8	25	9	5	25	6												
522. Organo positivo				8	20	6										4	20	2

	KEYBOARD MATERIALS									KEYBOARD MATERIALS								
	beech			rosewood			poplar			linden			boxwood			chestnut		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
486. Pianoforte (l tavolo																		
487. Pianoforte a tavolo																		
488. Pianino in forma di mobile per toeletta																		
489. Pianoforte o tavolo																		
490. Pianoforte a tavolo																		
491. Pianoforte a tavolo										7	30	7						
492. Pianoforte a coda																		
493. Pianoforte a coda																		
494. Pianoforte verticale a forma di giraffa																		
495. Pianoforte a tavolo informa d'ala																		
496. Pianoforte u tavolo										5	40	6						
497. Pianoforte a coda																		
498. Pianoforte a tavolo										5	30	4						
499. Pianoforte verticale																		
500. Pianoforte a tavolo																		
501. Pianoforte a tavolo										10	35	6						
502. Pianoforte a tavolo esagonale										8	30	6						
503. Pianoforte a tavolo																		
504. Pianoforte a coda																		
505. Pianoforte verticale																		
506. Pianoforte a coda										8	60	5						
507. Pianoforte verticale																		
508. Pianoforte a coda																		
509. Pianoforte a coda																		
510. Pianoforte verticale																		
511. Pianoforte verticale																		
512. Pianoforte verticale informa di armadio													5	30	6	7	40	6
513. Pianoforte verticale																		
514. Pianoforte verticale																		
515. Pianoforte a coda																		
516. Pianoforte a coda																		
518. Tastiera da digitazione																		
519. Organo																		
520. Organo processionale							4	20	5				4	20	6			
521. Organo positivo																		
522. Organo positivo																		

	KEYBOARD MATERIALS									KEYBOARD MATERIALS								
	durmast			oak			fruit			tin			copper			steel		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
486. Pianoforte (l tavolo							8	40	8									
487. Pianoforte a tavolo							2	30	7									
488. Pianino in forma di mobile per toeletta																		
489. Pianoforte o tavolo							3	5	4									
490. Pianoforte a tavolo																		
491. Pianoforte a tavolo																		
492. Pianoforte a coda							9	30	6									
493. Pianoforte a coda							1	30	6									
494. Pianoforte verticale a forma di giraffa							4	35	5									
495. Pianoforte a tavolo informa d'ala																		
496. Pianoforte u tavolo																		
497. Pianoforte a coda																		
498. Pianoforte a tavolo																		
499. Pianoforte verticale							3	10	6									
500. Pianoforte a tavolo																		
501. Pianoforte a tavolo																		
502. Pianoforte a tavolo esagonale							9	25	6				9	15	7			
503. Pianoforte a tavolo							3	40	5				5	8	6	6	13	5
504. Pianoforte a coda																		
505. Pianoforte verticale													9	15	6	1	25	2
506. Pianoforte a coda																		
507. Pianoforte verticale													4	20	7	10	40	7
508. Pianoforte a coda													2	20	6	7	40	6
509. Pianoforte a coda																		
510. Pianoforte verticale																		
511. Pianoforte verticale													6	20	7	5	30	5
512. Pianoforte verticale informa di armadio																		
513. Pianoforte verticale																		
514. Pianoforte verticale																		
515. Pianoforte a coda													4	20	6	10	30	6
516. Pianoforte a coda										8	20	6	5	20	7	2	40	4
518. Tastiera da digitazione	9	50	9															
519. Organo																		
520. Organo processionale							4	20	5									
521. Organo positivo																		
522. Organo positivo							3	20	7									

	KEYBOARD MATERIALS									KEYBOARD MATERIALS											
	lead			brass			iron			bone			felt			leather			ivory		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
486. Pianoforte (l tavolo													2	5	5	6	5	5	2	5	6
487. Pianoforte a tavolo										7	10	7	10	10	6	5	10	5			
488. Pianino in forma di mobile per toeletta										3	10	6				10	10	5			
489. Pianoforte o tavolo				3	5	6				9	5	5	10	5	6	3	5	6			
490. Pianoforte a tavolo				5	30	4	3	5	7							5	5	5			
491. Pianoforte a tavolo				10	15	6							6	5	4	4	5	5	4	5	5
492. Pianoforte a coda				9	15	5	10	10	6							7	5	6			
493. Pianoforte a coda	5	10	6	6	10	4				9	5	6	2	5	6	7	5	5	4	5	6
494. Pianoforte verticale a forma di giraffa				7	10	6				3	5	5	1	5	5	5	5	5			
495. Pianoforte a tavolo informa d'ala				1	10	6				10	5	4				4	10	4			
496. Pianoforte u tavolo	2	5	5										9	5	5	8	5	6	3	5	4
497. Pianoforte a coda				1	25	7										9	25	5	6	25	6
498. Pianoforte a tavolo																			1	10	4
499. Pianoforte verticale				2	5	2	6	25	5	9	5	7				8	5	5	4	5	6
500. Pianoforte a tavolo				4	20	6	9	20	7	7	5	7				10	5	5			
501. Pianoforte a tavolo				4	15	7	2	15	6				2	10	5	2	10	6	1	10	7
502. Pianoforte a tavolo esagonale							2	15	6				4	5	6				7	5	6
503. Pianoforte a tavolo				5	9	7				1	5	6									
504. Pianoforte a coda				9	22	6	5	30	5	2	12	6				4	12	5	5	12	5
505. Pianoforte verticale				2	15	7							6	15	5				8	15	7
506. Pianoforte a coda													4	10	6	2	5	4	2	10	9
507. Pianoforte verticale													4	10	6				10	10	6
508. Pianoforte a coda				1	8	5				6	10	6				4	7	6			
509. Pianoforte a coda													9	30	5				8	30	5
510. Pianoforte verticale													1	30	6				1	30	7
511. Pianoforte verticale										7	10	7	9	20	5						
512. Pianoforte verticale informa di armadio																3	30	6			
513. Pianoforte verticale										2	30	6	8	30	6						
514. Pianoforte verticale										2	30	6	7	30	5						
515. Pianoforte a coda													4	20	6				9	10	5
516. Pianoforte a coda													8	20	5						
518. Tastiera da digitazione																					
519. Organo				1	25	6	10	25	8	8	25	6									
520. Organo processionale				1	20	6	1	20	6												
521. Organo positivo				2	25	8	5	25	5												
522. Organo positivo							1	20	8	5	20	7									

Figure D.2: Museum repository data (2/6)

	SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS								
	ebony			fir			maple			mahogany			cypress			walnut			beech			cherry			rosewood		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
447. Clavicembalo				10	30	6	4	70	2																		
448. Clavicembalo	3	5	2	2	30	8							3	30	6	9	5	5				3	20	8			
449. Clavicembalo	4	10	7	2	10	7							2	80	4												
450. Clavicembalo	7	5	6	6	20	6							1	50	5	8	10	6	5	10	4						
451. Clavicembalo	4	14	4	6	30	6							2	55	5												
452. Clavicembalo in miniatura													10	100	7												
453. Clavicembalo				10	5	8				7	5	5	4	50	5	2	5	6									
454. Clavicembalo convertito in pianoforte				4	50	8										6	20	5									
455. Clavicembalo	9	20	5	1	50	6	4	10	5							8	10	5									
456. Clavicembalo				7	40	5													4	10	7						
457. Clavicembalo				6	30	5							8	30	6	5	30	6									
458. Virginale poligonale				7	45	6							9	20	5	7	10	7									
459. Virginale poligonale	1	10	6	1	20	6							4	35	5												
460. Virginale doppio				1	10	6													8	3	5						
461. Virginale rettangolare				5	45	6							3	20	4	5	25	8									
462. Virginale ottavino				10	25	8	1	15	8	4	15	5				9	25	6									
463. Virginale ottavino				10	35	7				8	10	7	8	35	8	3	10	4									
464. Virginale pentagonale																											
465. Virginale rettangolare	10	5	7	10	40	6							1	40	7												
466. Virginale in miniatura				8	30	6							1	30	6	9	30	7									
467. Virginale pentagonale	4	10	7	5	20	7							8	30	5	3	20	8									
468. Virginale rettangolare				5	10	7										8	40	7				6	10	8	4	8	6
469. Spinetta traversa				4	40	7										5	10	5									
470. Spinetta ottavina	9	15	6	9	30	6							5	25	8	1	20	6									
471. Spinetta traversa				10	50	8							4	10	8	2	15	6									
472. Spinetta. traversa				4	90	6																					
473. Spinetta traversa				3	70	6							8	10	4	6	10	3									
474. Spinetta traversa				10	35	5							7	20	5	2	20	3	5	8	7						
475. Spinetta traversa				1	45	4										4	45	5									
476. Clavicordo libero				6	40	8				4	40	6							4	10	5						
477. Clavicordo libero				9	60	6																1	40	5			
478. Pianoforte a tavolo				8	30	6	10	30	1							6	35	7									
479. Pianoforte a tavolo				5	30	8										5	40	5									
480. Pianoforte a tavolo	1	5	5	5	30	9	2	8	1	5	30	6							3	8	6						
481. Pianoforte o tavolo ottavi no				7	15	6										3	30	7				1	15	5			
482. Pianoforte a tavolo				8	25	6	2	10	6	1	40	6													2	10	7
483. Pianoforte a tavolo	9	5	4	10	20	7	7	5	7	2	40	6				9	15	7	7	5	5						
484. Pianoforte a tavolo				10	40	7	1	5	1							10	5	5				3	40	7	9	5	6
485. Pianoforte a tavolo				6	35	5	2	10	5							7	35	4							9	10	7
486. Pianoforte (l tavolo	6	10	7	7	30	7																1	50	3			

	SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS								
	poplar			linden			boxwood			chestnut			oak			fruit			elm			birch			tin			brass			iron			bone			leather			ivory			n.a.		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
447. Clavicembalo																																													
448. Clavicembalo																												6	5	6										5	5	7			
449. Clavicembalo																																													
450. Clavicembalo	5	5	3																																										
451. Clavicembalo																																		8	1	4									
452. Clavicembalo in miniatura																																													
453. Clavicembalo	4	35	5																																										
454. Clavicembalo convertito in pianoforte	8	20	6	3	10	8																																							
455. Clavicembalo																												3	10	5															
456. Clavicembalo	4	30	8	9	10	7							4	10	1																														
457. Clavicembalo																9	10	9																											
458. Virginale poligonale																6	10	5										6	10	6	3	5	6												
459. Virginale poligonale													1	7	7	1	8	9										3	5	7				10	15	3									
460. Virginale doppio	5	83	5																												2	2	3				9	2	6						
461. Virginale rettangolare																															2	5	7							1	5	4			
462. Virginale ottavino																															1	10	4	9	10	4									
463. Virginale ottavino																												7	10	7															
464. Virginale pentagonale																																											0	100	1
465. Virginale rettangolare																												5	5	6	7	5	7	9	5	1									
466. Virginale in miniatura																6	10	7																											
467. Virginale pentagonale							5	10	3																			10	5	6	8	5	7												
468. Virginale rettangolare																5	20	8										7	5	6				3	7	4									
469. Spinetta traversa				10	40	5																									7	10	6												
470. Spinetta ottavina																												3	10	6															
471. Spinetta traversa																10	10	6													3	10	4				4	5	5						
472. Spinetta. traversa																															8	10	5												
473. Spinetta traversa																															1	10	3												
474. Spinetta traversa	5	8	8																												7	9	5												
475. Spinetta traversa																												4	5	7	8	5	4												
476. Clavicordo libero																															9	10	5												
477. Clavicordo libero																																													
478. Pianoforte a tavolo																															6	5	6												
479. Pianoforte a tavolo																												8	15	6	3	15	4												
480. Pianoforte a tavolo																												9	7	7	5	7	5	10	5	1									
481. Pianoforte o tavolo ottavi no																3	10	5	4	20	2										5	10	2												
482. Pianoforte a tavolo							3	5	4																						5	10	1												
483. Pianoforte a tavolo							3	5	9																						9	5	6												
484. Pianoforte a tavolo																															3	5	3												
485. Pianoforte a tavolo																															1	10	6												
486. Pianoforte (l tavolo																															10	10	7												

Figure D.3: Museum repository data (3/6)

	SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS								
	ebony			fir			maple			mahogany			cypress			walnut			beech			cherry			rosewood		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
487. Pianoforte a tavolo				2	30	4	5	10	9													5	50	4			
488. Pianino in forma di mobile per toeletta				5	30	2	2	30	4							4	30	6									
489. Pianoforte o tavolo				10	35	6	10	5	5							5	40	8									
490. Pianoforte a tavolo				2	30	8	7	20	6							2	40	5									
491. Pianoforte a tavolo				7	25	6	2	10	6	2	40	4							6	10	7				5	5	8
492. Pianoforte a coda				5	30	4	8	30	6													10	30	6			
493. Pianoforte a coda				8	30	8				6	40	5										6	20	6			
494. Pianoforte verticale a forma di giraffa				6	40	7	6	50	2																		
495. Pianoforte a tavolo informa d'ala				5																							
496. Pianoforte u tavolo				9	30	4	8	5	5	3	35	6													4	20	3
497. Pianoforte a coda				4	30	6	7	30	7	10	30	6															
498. Pianoforte a tavolo				8	25	6	6	10	6	6	25	7													10	20	6
499. Pianoforte verticale				9	40	6	9	20	6																6	40	7
500. Pianoforte a tavolo				5	25	1	7	30	6							3	30	7	4	7	5						
501. Pianoforte a tavolo				8	30	8										8	30	6									
502. Pianoforte a tavolo esagonale				8	25	4	5	7	5	1	20	6													3	40	4
503. Pianoforte a tavolo				9	30	2										8	60	4									
504. Pianoforte a coda				5	35	9				3	65	4															
505. Pianoforte verticale				9	40	5																			8	50	4
506. Pianoforte a coda				1	40	2				6	50	6															
507. Pianoforte verticale				5	40	5																			4	50	6
508. Pianoforte a coda				2	45	8										6	45	6									
509. Pianoforte a coda				9	30	3	8	30	5																5	30	7
510. Pianoforte verticale				1	45	8																			2	45	4
511. Pianoforte verticale				7	30	6				1	30	5													1	30	3
512. Pianoforte verticale informa di armadio				10	40	6							7	30	6	5	30	6									
513. Pianoforte verticale				10	45	3																			1	45	5
514. Pianoforte verticale				5	40	5													2	10	6				7	40	4
515. Pianoforte a coda				6	90	9																					
516. Pianoforte a coda				10	90	7																					
518. Tastiera da digitazione																											
519. Organo													9	15	5												
520. Organo processionale				8	50	2																					
521. Organo positivo																											
522. Organo positivo																											

	SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS									SOUNDBOARD MATERIALS								
	poplar			linden			boxwood			chestnut			oak			fruit			elm			birch			tin			brass			iron			bone			leather			ivory			n.a.		
	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N	C	P	N
487. Pianoforte a tavolo																															6	10	8												
488. Pianino in forma di mobile per toeletta																															8	10	6												
489. Pianoforte o tavolo																												9	10	5	7	10	5												
490. Pianoforte a tavolo																															2	10	5												
491. Pianoforte a tavolo																															8	10	4												
492. Pianoforte a coda																															1	10	7												
493. Pianoforte a coda																															7	10	5												
494. Pianoforte verticale a forma di giraffa																															10	10	5												
495. Pianoforte a tavolo informa d'ala																																											0	100	1
496. Pianoforte u tavolo																															6	10	6												
497. Pianoforte a coda																															1	10	8												
498. Pianoforte a tavolo																												6	12	7	2	8	3												
499. Pianoforte verticale																																													
500. Pianoforte a tavolo																															7	8	4												
501. Pianoforte a tavolo																						1	30	8							2	10	7												
502. Pianoforte a tavolo esagonale																															10	8	5												
503. Pianoforte a tavolo																															4	10	7												
504. Pianoforte a coda																																													
505. Pianoforte verticale																															1	10	5												
506. Pianoforte a coda																															10	10	8												
507. Pianoforte verticale																															5	10	6												
508. Pianoforte a coda																															5	10	7												
509. Pianoforte a coda																												3	10	4															
510. Pianoforte verticale																															7	10	5												
511. Pianoforte verticale																															1	10	6												
512. Pianoforte verticale informa di armadio																																													
513. Pianoforte verticale																															5	10	4												
514. Pianoforte verticale																															5	10	6												
515. Pianoforte a coda																															2	10	8												
516. Pianoforte a coda																															9	10	6												
518. Tastiera da digitazione																																											0	100	1
519. Organo	6	40	7																						10	30	6	2	15	5															
520. Organo processionale	1	50	7																																										
521. Organo positivo	8	30	7							4	30	7													6	40	6																		
522. Organo positivo	8	50	9																						1	50	8																		

Figure D.4: Museum repository data (4/6)

	Conservation state	Artistic merits	Place	Time
447. Clavicembalo	9	9	Napoli	prima metà del XVI secolo
448. Clavicembalo	6	2	Venezia	fine del XVI secolo
449. Clavicembalo	6	10	Scuola veneta	metà XVII secolo
450. Clavicembalo	2	10	Italia centro-meridionale	1690
451. Clavicembalo	1	3	Bologna	1700
452. Clavicembalo in miniatura	5	3	Venezia	prima metà del XVIII secolo
453. Clavicembalo	4	1	Italia	fine XVII inizio XVIII secolo
454. Clavicembalo convertito in pianoforte	4	5	Piemonte	1750 ca.
455. Clavicembalo	3	8	Milano	1758
456. Clavicembalo	4	3	Parigi	metà del XVIII secolo
457. Clavicembalo	8	7	Londra	1895
458. Virginale poligonale	1	8	Lombardia	prima metà del XVI secolo
459. Virginale poligonale	4	4	Milano	metà del XVI secolo
460. Virginale doppio	2	9	Anversa	1600 ca.
461. Virginale rettangolare	8	1	Scuola veneta	prima metà del XVII secolo
462. Virginale ottavino	1	8	Scuola veneta	metà
463. Virginale ottavino	1	6	Scuola veneta	1649 (Î')
464. Virginale pentagonale	3	8	Italia centrale (Toscana?)	seconda metà del XVII secolo
465. Virginale rettangolare	1	4	Aquileia	1659
466. Virginale in miniatura	9	8	Bologna ('?)	seconda metà del XVII secolo
467. Virginale pentagonale	2	3	Italia	inizio del XVIII secolo
468. Virginale rettangolare	4	7	Bergamo	1836
469. Spinetta traversa	1	7	Italia	seconda metà del XVII secolo
470. Spinetta ottavina	4	5	Venezia	1692
471. Spinetta traversa	1	9	Italia	inizio del XVIII secolo
472. Spinetta. traversa	10	1	Italia	XVIII (7) secolo
473. Spinetta traversa	1	1	Milano	1735
474. Spinetta traversa	9	4	Italia	XVIII (7) secolo
475. Spinetta traversa	8	10	Scuola tedesca	ultimo decennio del XVIII secolo
476. Clavicordo libero	3	3	Stoccolma	1792
477. Clavicordo libero	1	2	Londra	1894
478. Pianoforte a tavolo	4	2	Scuola tedesca	seconda metà del XVIII secolo
479. Pianoforte a tavolo	9	4	Scuola tedesca	seconda metà del XVIII secolo
480. Pianoforte a tavolo	5	10	Londra	1778
481. Pianoforte o tavolo ottavi no	1	9	Milano	1778
482. Pianoforte a tavolo	1	6	Londra	1780 ca.
483. Pianoforte a tavolo	2	5	Londra	1782
484. Pianoforte a tavolo	1	9	Ginevra	1788 (1789?)
485. Pianoforte a tavolo	9	9	Torino	1797
486. Pianoforte (l tavolo	4	1	Varese	ca. 1800

Figure D.5: Museum repository data (5/6)

	Conservation state	Artistic merits	Place	Time
487. Pianoforte a tavolo	1	6	Milano	1804
488. Pianino in forma di mobile per toeletta	6	5	Vienna	primo quarto del XIX secolo
489. Pianoforte o tavolo	7	6	Milano (?)	inizio del XIX secolo
490. Pianoforte a tavolo	9	8	Milano	inizio del XIX secolo
491. Pianoforte a tavolo	5	1	Londra	ca. 1805
492. Pianoforte a coda	6	3	Vienna	1810 ca.
493. Pianoforte a coda	4	3	Vienna	ca. 1815
494. Pianoforte verticale a forma di giraffa	9	4	Scuola viennese	1815-1820
495. Pianoforte a tavolo informa d'ala	3	5	Vienna	ca. 1818
496. Pianoforte u tavolo	3	5	Londra	secondo decennio del XIX secolo
497. Pianoforte a coda	7	8	Vienna	1820
498. Pianoforte a tavolo	1	9	Londra	1820-1821
499. Pianoforte verticale	9	8	Vienna	1829
500. Pianoforte a tavolo	3	3	Milano	ca. 1830
501. Pianoforte a tavolo	1	5	Parigi	1835-1840
502. Pianoforte a tavolo esagonale	3	4	Parigi — Londra	1840 ca.
503. Pianoforte a tavolo	3	3	Bologna	1850 ca.
504. Pianoforte a coda	5	8	Vienna	1840 ca.
505. Pianoforte verticale	7	7	Parigi	1843
506. Pianoforte a coda	1	3	Parigi	ca. 1845
507. Pianoforte verticale	8	1	Parigi	ca. 1850
508. Pianoforte a coda	1	10	Milano	metà del XIX secolo
509. Pianoforte a coda	3	7	Parigi	1858
510. Pianoforte verticale	4	5	Londra	1860 ca.
511. Pianoforte verticale	5	4	Francia	1860 ca.
512. Pianoforte verticale informa di armadio	10	10	Italia	XIX (?) secolo
513. Pianoforte verticale	1	5	Milano	seconda metà del XIX secolo
514. Pianoforte verticale	5	4	Milano	ultimo quarto del XIX secolo
515. Pianoforte a coda	7	9	Berlino	1913 ca.
516. Pianoforte a coda	3	7	Vienna	inizio del XX secolo
518. Tastiera da digitazione	2	8		XX secolo
519. Organo	7	3	Emilia	prima metà del XVIII secolo
520. Organo processionale	7	9		
521. Organo positivo	6	2	Napoli	1839
522. Organo positivo	10	2	Toscana	XVIII secolo

Figure D.6: Museum repository data (6/6)