

# A historical and contemporary study on annotations to derive key features for systems design

Maristella Agosti · Giorgetta Bonfiglio-Dosio ·  
Nicola Ferro

Published online: 13 February 2007  
© Springer-Verlag 2007

**Abstract** This paper provides a comprehensive study on annotations by defining their contours and complexity. This work adds a new complementary approach to the usual case and user studies, and also investigates history in order to benefit from previous knowledge and our cultural heritage. This study emphasizes an aspect which has never previously been taken into account: the temporal dimension involved in annotations. Moreover, it discusses both the notion of hypertext between documents and annotations and the idea of annotations as context for documents. The study gives the necessary historical and cultural background to derive a set of key features of annotations that must be taken into account when designing systems that have to support the management of digital annotations on digital contents.

**Keywords** Annotations · Contents · Metadata · Hypertext · Context · Temporal dimension

## 1 Introduction and motivations

The study and the research concerning the annotation of digital contents are an active field of investigation,

---

M. Agosti · N. Ferro (✉)  
Department of Information Engineering,  
University of Padua, Via Gradenigo, 6/a,  
35131 Padova, Italy  
e-mail: nicola.ferro@unipd.it

G. Bonfiglio-Dosio  
Department of History, University of Padua,  
Via del Vescovado, 30, 35141 Padova, Italy  
e-mail: giorgetta.bonfigliodosio@unipd.it

M. Agosti  
e-mail: maristella.agosti@unipd.it

which can be looked at from many different perspectives. The previous research work on annotations has left us with many open issues. These issues concern the lack of clarity about: what an annotation is, what the features and the way of using an annotation are, and what architecture and functionalities a system with annotation capabilities has to provide. These issues are mainly due to the fact that until now models and systems for annotations have been developed for specific purposes. As a result, there is a fragmentary picture on the annotation and its management, since the models and systems are all tied to specific usages and lack a general validity.

The aim of this paper is to discuss the different perspectives regarding annotations in order to gather some key features about them. These key features can help us to better distinguish among the different uses of annotations and to understand the case at hand when dealing with annotations. Furthermore, they can be used as a support if we need to make design choices for developing an annotation management system. In addition, they can also be seen as the groundwork needed to properly delineate the complexity of this problem, which is often underestimated and only partially addressed. Finally, these key features open the way for defining a comprehensive formal model of annotations on digital contents, which has not been dealt with in previous research on this topic and to date is still lacking. In conclusion, we aim at contributing to solving the issue pointed out by [29]:

strangely enough, there is not an agreement yet on the definition of digital annotation, or on how to distinguish it from other digital entities (e.g. hyperlinks, metadata, newsgroup messages).

Furthermore, an analysis of the basic operations, to be enabled by a digital annotation system, seems to be lacking.

This paper does not intend to be a fully exhaustive survey on annotations or a comparison of all the existing annotation systems along some predefined set of features. The reader interested in these issues can refer to [36, 51, 71, 74]

In order to achieve our aim, we will proceed in a two-fold way by analysing annotations both from a historical perspective and from a contemporary one. The rationale for studying both the historical and the present viewpoints about annotations is that our current notion of annotation, even in a digital world, is strongly influenced by the long history of the annotation. Thus, looking at both aspects can provide us with a better understanding of annotations.

When we talk about annotations, we deal with a concept that has been stratified over a long period of time in our culture, and literary research is the most effective way to benefit from the pre-existing knowledge of our cultural heritage [5, 42]. On the other hand, analysing the present use of annotations help us to understand both the current trends in developing annotation management systems and user requirements and expectations.

With respect to user requirements, if we were to conduct a user study aimed at gathering the user requirements for the design of an annotation management system, users may not be able to express all their implicit requirements, because they often overlook what they have naturally absorbed from their cultural heritage or studies. On the other hand, the historical perspective on annotations provides us with the additional information needed to complete the picture when designing an annotation system. In a certain sense, we are conducting a user study, where our users are the history of the annotation, and they tell us what features of the annotation are relevant.

In conclusion, the outcomes of this research are key points concerning the features of the annotation that we should take into account when developing a model for the annotation and when designing a system capable of providing annotation functionalities.

The paper is organized as follows: firstly, we present different viewpoints on annotations, where Sect. 2.1 presents a thorough study from a historical point of view, and Sect. 2.2 analyses present perspectives. Then, we study the different uses of annotations: Sect. 3.1 discusses some relevant historical uses, while Sect. 3.2 introduces the current uses of annotation in different information management systems, and highlights how annotations can be used to search for relevant docu-

ments in response to a user query. After that, Sect. 4 gathers up the observations made in the previous sections in order to wrap up and point out key features about annotations that need to be taken into account when designing information management systems.

## 2 Viewpoints

### 2.1 Historical viewpoints

A basic step in approaching the problem of annotations is to define the meaning of the different terms that come into play and to investigate their historical use over the course of time. All of this is necessary in order to gather information to effectively outline the problem.

Annotations are used not only in the literary field but also for administrative and legal practice. In this section, we have conducted research about annotations that is an exhaustive analysis of the terms in different frameworks and in different disciplines. The approach is literary, because we firmly believe that studying the terms, their meanings, their etymologies, and the way they have been used can provide us with a solid groundwork on which we can base our design choices. We have adopted the following methodology: first, we look up the given term in the dictionary; then, we investigate its etymology and its historical usage; now that we have all the needed information, we can highlight some key points about the annotation to take into consideration during the design phase.

#### 2.1.1 Annotation

Both [52, p. 57] and [55, p. 198] define the word *annotation*, firstly, as the *act of annotating* and, secondly, as a *note added as an explanation especially of some literary work*. IEI [55, p. 198] further observes that the word annotation can also be used in sentences with a passive sense, as something worthy of annotating, in other words *noteworthy* and *worth remembering*. The word annotation is closely related to the verb *annotate* which, in turn, means *to supply (a written work, such as an ancient text) with critical or explanatory notes* [52, p. 57] and *to note down, to write down, to record something* [55, p. 198]. But an important observation must be added to this definition: the annotation can be used in many situations tied to legal and administrative practices.

Both [37, p. 107] and [52, p. 57] trace the etymology of annotation back to the Latin word *annōtātio*, which simply means *annotation, note* [34, p. 189]. The Latin word *annōtātio*, in turn, derives from the Latin verb *annōtāre*, which means *to annotate* and *to observe in writ-*

ing [34, p. 190]. Finally, the Latin verb *annōtāre* comes from the Latin word *nōta*, which means *note, mark* [34, p. 1823], plus the intensifying prefix *ād*, which in compound words means to approach, to tend and thus to add [34, pp. 37–40]. Both [34, p. 1823] and [37, p. 1047] agree that the Latin word *nōta* has an uncertain etymology.

We can distinguish different types of annotations. Firstly, annotations written by the authors themselves: these can be corrections (both formal and the content: for example, the proofs), additions (when the authors themselves add something to their written texts, after they have been published in a final version), and sometimes commentaries. Secondly, the reader adds some commentaries to the text written by the author. Thirdly, an annotation is made with an administrative function, also on the back of a document, as explained later in Sect. 3.1.

From this point of view studying how annotations were created, developed, and used during different historical moments means to understand how a text derives from a progressive stratification. In fact, a text can go through subsequent phases of life (preparatory, final, re-elaborated, interpreted), and in these phases the annotations play a very important part, because they not only show the editorial steps, but also show how the author divides different information in three groups: private information, information shared with certain people, and public information. These kinds of information are various and they can range from simple notes for personal and non-transferable use, which at times can be only vague ideas, to more formalized comments on texts which have been written by others.

This brief discussion highlights some interesting points about the annotation. Firstly, the annotation is not only an object or something that is passive, but it also contains the notion of activity, as its first meaning “act of annotating”. In this sense, the annotation calls for an active involvement by the subject who is engaged in the act of supplying explanatory matter or keeping record of something. Furthermore, the annotation also covers, in its second meaning, the purpose of this active involvement, as for example to produce an intellectual work in order to add an explanation to some literary work. This idea is further supported by the meaning of the verb to annotate, which broadens the spectrum of the word annotation towards keeping record of something. Therefore, on the whole, the annotation requires an active involvement in order to produce an intellectual work that has to be recorded. These facets of the annotation are also present in the etymology of the word annotation: the Latin verb *annōtāre* means to make written observations, comments or remarks in the long-lasting and recordable written form; on the other hand, the

Latin word *nōta* recalls the note or the mark put to remember or highlight something.

In conclusion, we can point out how, in an annotation, both the content of the annotation and the annotated object benefit from each other and mutually reinforce themselves. Indeed, the annotated object becomes noteworthy thanks to the content of the annotation on it; this represents a passive aspect of the act of annotating, since the annotated object undergoes the annotation. On the other hand, the content of the annotation, which represents the active part of the act of annotating, also becomes worth recording thanks to its link with the annotated object. As a concrete example, consider an authoritative commentary on a valuable literary work: the literary work benefits from being commented by an authoritative person; on the other hand, the commentary is more relevant since it deals with a valuable literary work.

### 2.1.2 Terms related to annotation

The range of our investigations can be widened to also take into consideration synonyms and terms related to the word annotation.

Spencer [82, p. 14] provides the following synonyms of the term *annotation*: *comment, commentary, elucidation, explanation, footnote, gloss, interpretation, and note*. We can also add the word *jotting* to this list, which is a very brief annotation [52, p. 789], and the word *scholium*, which is a particular kind of annotation [58, p. 158].

Table 1 provides the definitions for the different terms listed above. Note that these definitions often refer to printed documents or texts, since they are taken from an English dictionary [52]. On the other hand, we should consider that their validity is not limited only to printed documents, but can also be applied and extended to information resources in a digital context.

As can be noted from Table 1, these words are often defined by terms used to describe other words in the list, or they refer to the same notion of explaining, expounding, interpreting, clarifying, or recording something. In this way, they reveal how closely related they are.

In conclusion, the terms listed in Table 1 support, refine and strengthen what has been observed above about the word annotation. Therefore, the table introduces further kinds of annotation which cover different needs and tasks, such as the gloss, the postil, the note, the jotting, and so on.

Now we can move a step further and investigate in more detail the terms gloss, scholium, and postil in order

**Table 1** Definition of the terms related to annotation from [52]

Word	Definition
Comment	<i>A note explaining or criticizing a passage in a text and explanatory or critical matter added to a text</i>
Commentary	<i>An explanatory series of notes or comments</i>
Elucidation	<i>Making clear (something obscure or difficult)</i>
Explanation	<i>The act or process of explaining and a statement or occurrence that explains and a clarification of disputed terms or points</i>
Footnote	<i>A note printed at the bottom of a page, to which attention is drawn by means of a reference mark in the body of the text</i>
Gloss	<i>A short or expanded explanation or interpretation of a word, expression, or foreign phrase in the margin or text of a manuscript</i>
To jot	<i>To write a brief note of</i>
Jotting	<i>Something jotted down</i>
Interpretation	<i>The act or process of interpreting or explaining; elucidation and the result of interpreting; an explanation</i>
Note	<i>A brief summary or record in writing especially a jotting for future reference and a short written statement giving any kind of information and a critical comment, explanatory statement, or reference in the text of a book, often preceded by a number</i>
Notes	<i>Short descriptive or summarized jottings taken down for future reference</i>
Observation	<i>The act of observing or the state of being observed and a comment or remark and the facts learned from observing</i>
Postil	<i>A commentary or marginal note, as in a Bible</i>
Record	<i>An account in permanent form, especially in writing, preserving knowledge or information about facts or events</i>
Scholium	<i>A commentary or annotation, especially on a classical text</i>

to understand the rich semantics of the annotation, how it has evolved over time and its current consequences.

### 2.1.3 Gloss

As reported in [52, p. 620] and [37, p. 673], the word gloss derives from the ancient Greek word  $\gamma\lambda\omega\sigma\sigma\alpha$  (glōssa), which means *tongue, language, idiom, spoken word, foreign or obsolete word* [77, p. 393].

As reported in [56, pp. 652–653], at the time of the ancient Greeks, the term gloss meant an obscure, archaic, dialect, or rare locution that required an additional explanation. These locutions were object of study by grammarians or object of research by scholarly poets, especially the Alexandrine poets, who enriched their compositions with these terms. Then, gloss meant the explanations themselves of such locutions, either collected in wide-ranging lexicons or as interlinear notes placed above the words to be explained. This methodology of study and a lexicographical practice dates back to ancient times (there were glosses to Homer as early as the fifth century BC) and was fully developed by the grammarians of the Alexandrine age. During the Bizantine age and the Middle Ages, the term gloss meant an interlinear or marginal note to a biblical or juridical codex. For the biblical codices, the gloss was a very short paraphrase to explain a passage of the Bible, sometimes together with a mention to its allegorical interpretation. On the other hand, for the juridical codices, the glosses were explanatory annotations which constituted a thorough commentary to the text.

This discussion about the term gloss points out some interesting facets of the annotation which have not fully emerged in the previous observations about the term annotation. The intellectual work entailed by the gloss is of very high quality, because it is a method of both study and research. This kind of intellectual work gives us an idea of how strong the active involvement required by the gloss is: it does not concern only the author themselves, but it is also capable of involving and stimulating a wide community of people who work, study and do research on a subject. Therefore, it turns out that an annotation may comprise a public dimension, because it becomes the vehicle for carrying and transmitting ideas and knowledge to other people, or it may comprise a shared dimension only, if the recipients of the annotation are less numerous.

The research or study aspects and the public or shared dimension entailed by the gloss help us to understand how long-lasting and recordable the annotations are. Indeed, they are not only comments and remarks to a text, but also an autonomous intellectual work which is worth recording. Finally, note how also in the case of the gloss there is an intrinsic mutual reinforcement between the annotations and the annotated objects which strengthen each other by deepening the knowledge on a subject.

### 2.1.4 Scholium

As reported in [52, p. 1305] and [37, p. 1479], the word scholium derives from the ancient Greek word  $\sigma\chi\acute{o}\lambda\iota\omicron\nu$

(schólion), which means *comment, explanation* [77, p. 1793]. The ancient Greek word *σχόλιον* (schólion), in turn, comes from the ancient Greek word *σχολή* (scholḗ), which means *scholarly activity and school* [77, p. 1793].

IEI [53, pp. 198–199] reports that the word scholium designates short annotations or explanations written by a reader in the margin of a manuscript. The distinguishing features of the scholium are the fact that they are anonymous and fragmentary. The scholium is anonymous because, initially, the reader writes their own observations in the margin of the manuscript or passages taken from a commentary for both personal use and scholastic needs. The next owner of the manuscript often extends the scholium or modifies it. Therefore, the lack of organic unity is explained in this way. Often scholia also contain citations by the authors from which the observations are taken; in this way, they are very useful for reconstructing the doctrines and the works of ancient grammarians that may no longer exist. Note that the term scholium is used only in the literary field and always means comment to the work of different authors. Today it is not used in current languages, excepted Rumanian and Russian, where it means commentary.

The term scholium suggests another facet of the annotation: it may be created for personal purposes, and therefore may entail a private dimension, since the main recipient of the annotation is the author themselves. However, the private dimension may represent only the initial intention of the annotation, because other people reading an annotated text can also benefit from existing annotations and can modify or extend them; thus, the annotation passes from a private dimension to a shared one. Taken to the extreme this process encompasses the possibility that an annotation becomes the means for studying the thought of authors that otherwise would be lost; thus, the annotation passes from a private dimension to a public one. In conclusion, private annotations are part of this spectrum of possibilities and this makes us aware of the necessity to carefully preserve private annotations, because they may also become worth recording for reasons different from the ones that motivated their creation.

### 2.1.5 Postil

As introduced in Table 1, a postil is a short annotation—often a marginal or interlinear note—to a text, handwritten by a scholar or by the authors themselves to express observations, explanations, or criticisms. During the Middle Ages, postils were a scholastic practice

and they sometimes represented comments that were more explanatory than simple notes [57, p. 1030].

Both [52, p. 1145] and [37, p. 1239] trace the etymology of postil back to the Latin terms *pōst illā* (*verbā textūs*) which mean *after those (words in the text)*, which often was the opening phrase of such annotation.

Therefore, in the etymology itself the word postil emphasizes one of the main aspects concerning the annotation: the annotation is the result of an intellectual work on an existing text and is a follow-up to an already existing text. Therefore, the annotation comprises a temporal dimension that is often not explicit but that limits the creation of the annotation to the existence of another text. This temporal relationship between the annotation and the annotated text does not mean that the annotation cannot be considered as a stand-alone intellectual work—and some glosses and scholia are by right considered autonomous pieces of knowledge—but it imposes a temporal ordering between the existence of an annotated text and the annotation on it. This temporal ordering cannot be overlooked since, for example, it allows us to reconstruct how a piece of knowledge has been created. Note that this temporal dimension turns out to be fundamental if we think about the mutual reinforcement between annotations and annotated objects, since this does not happen immediately but on the contrary needs time.

## 2.2 Present viewpoints

Many user studies are aimed at understanding annotation practices and discovering common annotation patterns. [64] studied personal annotative practices of American college students to point out the form the annotations have in the textbooks and the function of the annotations, which is derived from their form. [64, pp. 237–238] discovered that:

First, annotations are *procedural signals*, cluing in the student to where an assignment starts, what material is important (and as we will see, unimportant), and what material might require a second (or successive readings). Second, annotations are *placemarks*; they hold the quotes that are being reserved for the paper that the student will write at the end of the term, the chemical reactions and term definitions the student must memorize for the final, the theorem that is key to the proof in the homework assignment. Third, they are an *in situ way of working problems*. Fourth, annotations record *interpretive activity*, either from another reader (e.g. a professor's explanation), or as the result of careful reading (the student has inter-

preted it him or herself). Fifth, and most elusively, these markings act as a *visible trace of a reader's attention*, a focus on the passing words, and a marker of all that has already been read (as if these words are now possessed). Finally, the markings may just be incidental, *reflecting the material circumstance of reading*.

Marshall [65] carries on her research work and categorizes annotations along several dimensions which reflect the form annotations may take on: formal versus informal annotations, explicit versus tacit annotations, annotations as writing versus annotations as reading, hyperextensive versus extensive versus intensive annotations, permanent versus transient annotations, published versus private annotations. Finally, [66,67,81] investigate the relationship between private, shared and public annotations and how they can be exploited to find useful passages in the text.

It is worth noting how the findings of [64] and [65] agree with the outcomes of the study conducted in the previous section about the historical perspective on annotations. Indeed, both glosses and scholia are, to some extent, *placemarks*, an in situ way of working problems, they record an *interpretive activity*, and so on. The different dimensions of the annotation are also taken into account by the historical perspective: glosses are often more formal annotations than scholia and postils, which are usually informal; scholia can be tacit annotations due to their fragmentariness while glosses can be explicit annotations; all the kinds of annotations described in Sect. 2.1 act as a bridge between reading and writing; glosses may be considered intensive annotations, postils may be more extensive annotations and both glosses and scholia often contain references to other authors and quotations of other texts, which is a way of being hyperextensive annotations; the stratification of glosses and scholia in our cultural heritage is a clear sign of the passage from transient to permanent annotations; finally, the difference between postils, scholia, and glosses comprises the distinction between private and public annotations. On the other hand, neither [64] nor [65] explicitly points out the temporal dimension entailed by annotations and the temporal ordering between annotations and annotated objects, which has been discussed in Sect. 2.1.5 with regard to the term postil.

Phelps and Wilensky [74] suggest a list of desirable properties for annotations: annotations should appear in situ, that is on the documents themselves; they should be *highly expressive*; they should be *format and platform independent*; they should be *extensible, yet composable*, in other words they should allow different styles of anno-

tation; they should be *distributed, open, and robust*: they may reside in a place while referring to documents in another one.

Over the years, a lot of research work concerning annotations has been done where the main focus of this work has been on the employment of ad hoc devices or handheld devices which enable reading appliances with annotation capabilities, or on the design and development of document models and systems which support annotations in specific management systems, in particular:

- in digital libraries;
- in the Web;
- in collaborative systems and working groups;
- in databases.

All of this research work has led to different viewpoints about what an annotation is [7,10]; these different viewpoints are taken into consideration in the following sections.

### 2.2.1 Annotations are metadata

Annotations can be considered as additional data which concern an existing content, i.e. they are metadata, as they clarify in some way the properties and the semantics of the annotated content. For example, the Annotea<sup>1</sup> project developed by the *World Wide Web Consortium (W3C)* [61] sees annotations as metadata and interprets them as the first step in creating an infrastructure which will handle and associate metadata with content and will lead to the Semantic Web.<sup>2</sup>

Another example is MPEG-7 [60], formally named “Multimedia Content Description Interface”, which is a standard for annotating and describing multimedia content data. MPEG-7 to some degree supports the interpretation of the information meaning, which can be passed onto, or accessed by, a device or a computer code. MPEG-7 is not aimed at any one application in particular; rather, the elements that MPEG-7 standardizes support as many broad ranges of applications as possible. As a further example, in the context of *DataBase Management System (DBMS)* [24] sees annotations as “information about data such as provenance, comments, or other types of metadata.”

<sup>1</sup> <http://www.w3.org/2001/Annotea/>.

<sup>2</sup> <http://www.w3.org/2001/sw/>.

### 2.2.2 Annotations are contents

Differently from the previous case, annotations are additional contents which concern an existing content [71]; indeed, they increase existing content by providing an additional layer of content that elucidates and explains the existing one.

This viewpoint about annotations entails an intrinsic dualism between annotation as content enrichment and annotation as stand-alone document [5]:

- *annotation as content enrichment*: in this view annotations are considered as mere additional content regarding an existing document and as a result they are not autonomous entities but in fact they rely on previously existing information resources to justify their existence;
- *annotation as stand-alone document*: in this view annotations are considered as real documents and are autonomous entities that maintain some sort of connection with an existing document.

This twofold nature of the annotation is clear if we think about the process of studying a document: firstly, we can start annotating some interesting passages that require an in-depth investigation which can be considered content enrichment; then we can reconsider and collect our annotations and we can use them as a starting point for a new document which covers the points we would like to explain better, all of which is an annotation as a stand-alone document. In this case the annotation process can be seen as an informal, unstructured elaboration which could lead to a rethinking of the annotated document and to the creation of a new one. Bottoni et al. [29] also agree with this viewpoint about annotations and consider them to be reliant on the annotated objects; in this way, [29] consider annotations as content enrichment.

### 2.2.3 Annotations constitute a hypertext

Annotations allow the creation of new relationships between existing contents, by means of links that connect annotations together with existing content. In this sense we can consider that existing content and annotations constitute a hypertext [8,9], according to the definition of hypertext provided in [1]. This hypertext can be exploited not only for providing alternative navigation and browsing capabilities, but also for offering advanced search functionalities. Furthermore, [65] considers annotations as a natural way of creating and increasing hypertexts that connect information reso-

urces in a *Digital Library Management System (DLMS)* by actively engaging users. Finally, the hypertext which exists between information resources and annotations enables different annotation configurations: the first are *threads of annotations*, i.e., an annotation made in response to another annotation, and the second are *sets of annotations*, i.e. a bundle of annotations on the same information resource [5,10].

### 2.2.4 Annotations are a kind of context

Annotations introduce a new content layer aimed at elucidating the meaning of underlying documents, so that annotations can make hidden facets of the annotated documents more explicit. We can consider that annotations constitute a special kind of context which we call *annotative context*, for the documents of a *DLMS*, because they provide additional content which is related to the annotated documents. This viewpoint about annotations covers a wide range of annotations, ranging from personal jottings in the margin of a page to scholarly comments made by an expert to explain a passage of a text. Therefore, these different kinds of annotations involve different scopes for the annotation itself and, consequently, different kinds of annotative context. If we deal with a personal jotting, the recipient of the annotation is usually the author himself and so this kind of annotation involves a *private annotative context*; on the other hand, the recipients of a scholarly annotation are usually people who are not necessarily related to the author of the annotation, which thus involves a *public annotative context*; finally, a team of people can work together on a shared topic and can exchange annotations related to the topic in question: thus, in this case we have a *collaborative annotative context* [8].

Neuhold et al. [72] make use of annotations for exploiting and reconstructing different types of context. For example, annotations can be useful for reconstructing the *creation context* of historical documents, when this context is not available any more. For example, systems like *Collaboratory for Annotation Indexing and Retrieval of Digitized Historical Archive Material (COLLATE)* [45,84] or *Imaginum Patavinae Scientiae Archivum (IPSA)* [2,11,12] support researchers in performing this task. Furthermore, [72] exploit annotation for gathering information about the *interpretation context* of a document, so that it is possible to provide users with a framework where the interpretation of documents takes place. Finally, also the *collection context*, which provides information about documents in relation to other documents within a collection, can be enhanced by using the hypertext that exists between annotations and annotated documents [72].

### 2.2.5 Annotations are dialog acts

Annotations are part of a discourse with an existing content, where not only the single statement within the discourse is taken into consideration, but also its position and type within the discourse structure. This approach is taken by [45,46,84] who consider annotations as a means for placing a document in the collaborative discourse to which it is related.

Bottoni et al. [28,29] rely, in some way, on this notion of annotation, because they make use of the *Rhetorical Structure Theory (RST)* [63] to define a set of annotation types. Note that also [84] use the *RST* for the same purpose.

Finally, [43] agree, to some extent, with this viewpoint about annotations. Indeed, they interpret annotations as a means that allow a “two way exchange of ideas between the authors of the documents and the document users”.

## 3 Usages

### 3.1 Historical uses

#### 3.1.1 Historical annotations

Besides literary and philological annotations, there are also annotations concerning administrative documents and procedures. As a significative example of this kind, during the Middle Ages notaries and archivists usually adopted a three-step procedure for preparing legal papers.

The procedure was influenced by the characteristics of the parchment, which was used for writing and recording the information. Indeed, the parchment was made of sheep skin and was partially tanned. As a consequence, the front of the parchment, which was the internal part of the skin, was whiter and more polished than the back of the parchment, which was the part of the skin towards the coat.

Therefore, in the first step of the procedure, notaries recorded a draft of the names and other data, summarizing the legal action, on the back of the parchment. These annotations are called *back-annotations*, since they were written on the back of the parchment, and they were written beforehand [76, p. 51]. In the second step, these annotations were used for writing the notarial document in complete and legal form (*in publica forma*)<sup>3</sup> on

the front of the parchment [22]. In the third and final step, since the notarial documents were preserved by rolling them up, the archivists usually wrote a brief summary on the back of the parchment, in order to avoid opening it and to facilitate the consultation. In addition, on the back of the parchment, the archivists used to write a position-marker (*segnatura*) to show the logical and physical position of this document inside the archive. These annotations came after the document and now serve as a trace for archival management showing the ancient structure and vicissitudes of archives [73, pp. 243].

In other cases, the annotations highlight the stratifications of an administrative text due to subsequent administrative and legal actions. For example, both in Middle Ages and Modern Ages, when people enrolled in organizations, they were registered in special books, called rolls (*matriculae*), and these registrations were usually self-written, in other words standardized forms did not exist and thus each registration took on a life of its own. This situation and this special kind of annotation is well represented in Fig. 1, where a page from a roll of students is shown. The student wrote his personal data directly in the roll, as reported below from annotation  $a_1$  of Fig. 1 (the translation is in brackets):

*Godefridus Woyszel, filius Sigismundi Woysselii reipublicae Vratislaviensis archiatri, inclitae Germanorum nationi nomen suum dedit et solvit more solito sex libras. Patavii anno 1605 die 29 augusti* (Godfrey Woyszel, Sigmund's son, from Bratislav, enrolled and paid 6 liras).

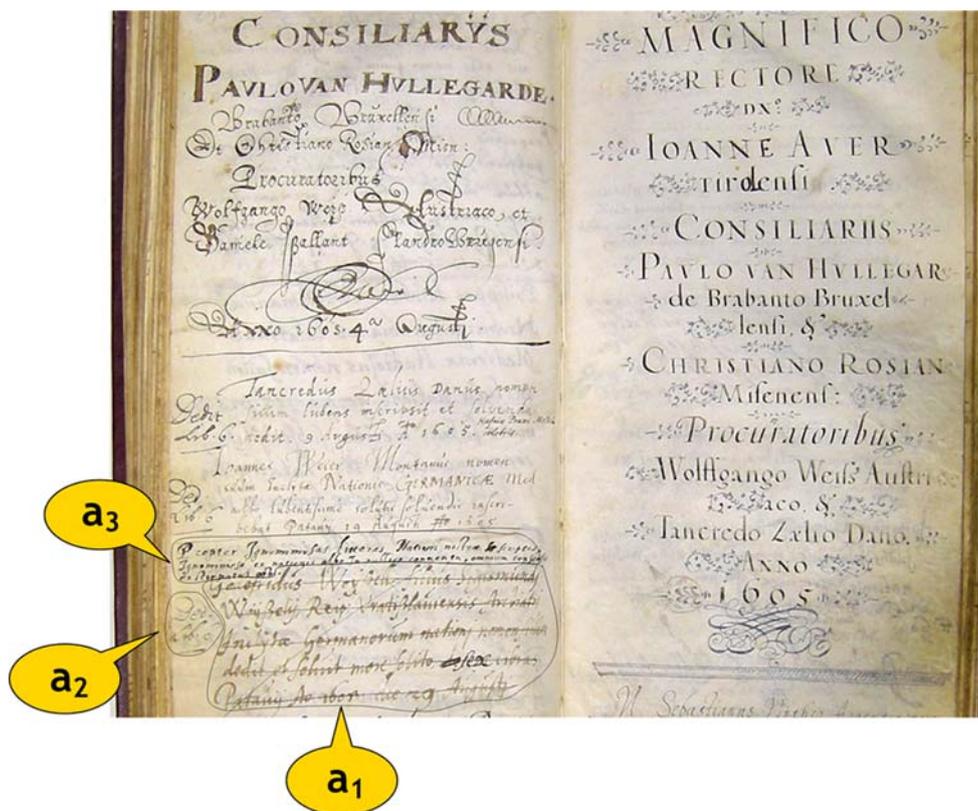
After this registration, the chancellor of the University recorded the tuition fee as paid; in annotation  $a_2$  of Fig. 1 there is written *Dedit libras 6* (he paid 6 liras).

With respect to the notion of time, it must be noted that the annotation of the chancellor occurred contemporary to or just after the student's enrolment. Afterwards, the chancellor could add other annotations about some episodes or incidents in the academic life of a student, as from annotation  $a_3$  of Fig. 1:

*Propter ignominiosas litteras Nationi nostrae scriptas ignominiose ex Nationis albo in publico conventu, omnium consensu, extirpatus est* (he was ignominiously expelled by the Council of Association, because he wrote an ignominious letter to the Association) [78, p. 134 n. 1125].

Furthermore, note how the physical placement of the annotations does not follow their temporal ordering but instead is constrained by the free space available on the

<sup>3</sup> The italicized words in brackets are the Latin technical terms used to indicate the specific technique. Note that they are the translation of the word which precedes them.



**Fig. 1** The photo shows subsequent annotations on a roll of students (Italia, Padova, Archivio dell'Università di Padova, Archivio antico, *Matricula Nationis Germanicae artistarum*, reg. 465, c. 69v—published with the permission of Archivio Generale di Ateneo di Padova)

paper; in the example of Fig. 1, annotation  $a_3$  is written above annotation  $a_1$ , even though  $a_3$  comes after  $a_1$ , because the space below  $a_1$  was already filled with the information about another student. Finally, note how annotation  $a_1$ , which reported the personal data of the student, had been crossed out as a consequence of the expulsion of the student from the University, as reported in annotation  $a_3$ . We can consider the cross-out as an additional annotation  $a_4$ , which modifies the status of an existing annotation after its creation.

It is worth pointing out that  $a_4$  neither delete nor remove  $a_1$  but, on the contrary, the personal data of the student are still readable under the cross-out. This is also particularly important in the case of digital annotations, because an annotation management system should imitate the approach described above and avoid completely deleting its annotations, since they allow us to trace how a piece of information has been created and to provide the lineage of each piece of information managed by the system.

As a final remark, in the rolls there were also annotations about the professional career of a student after his degree, such as *Medicinam exercet in patria* (he has a medical practice in his country of origin). This suggests

the use of annotations as a means of carrying out a data curation approach with respect to the information resources managed by our system [3,4], because annotations allow us to keep the information resources updated.

On the documents created by complex chanceries numerous annotations were used, each one with its specific purpose for denoting a precise step of a complex procedure; for an analytic list of these annotations, please see the one of Pope's chanceries [44, pp. 86–89]. For example, we can find some annotations which are a concrete trace of collation between the first draft and the final version of a document: in order to state that a document was in the final version and it conformed to the will of the author, the chancellor annotated some words, which authorized the document being sent. As a further example, the parchment of the Pope's letter was usually folded in order to affix the hanging seal. In this parchment fold, we can find the annotations about a chancery tax which the addressee must have paid before receiving the document. Also the different kinds of seal were a further formal signal and tool which the document author used to inform the addressees about significant contents.

As a further example, from Middle Ages until today public chancelleries usually annotate analogical documents in different forms. During the late Middle Ages the annotation *Registratum* (registered) written on the document meant that this document was transcribed in public form on a special book, together with other annotations written by the chancellor and chancery notaries [30, pp. 679–691] and [20,21]. Since the nineteenth century integral transcription has been replaced by the summary of the data of the document. In fact, at the end of eighteenth and nineteenth centuries, public administrations adopted a registration system for records management [41, pp. 165–169] and [25, pp. 355–357], based on the following tools: register of inward and outward letter-book; classification plan; and file list. Each of these tools provides us with some kind of annotation, as discussed in the following. The register contains the descriptive data of each document, such as number and date, which are also written on the document itself by using a stamp as a form to be filled in with the descriptive data. The classification plan allows for the logical organization of archival documents and determines the classification index which is written on each document. The aggregation of documents into files gives a physical order to the archives which reflects a logical order. Each file is enclosed in a paper container where identification data on the file are written [26, pp. 45–60].

Documents often contain a lot of annotations which show the subsequent steps of administrative procedures. For example, on petition papers the sovereign wrote a word of approval (usually *fiat*) [40, pp. 243–245] and [23, tav. 34]. The administrative annotation can also be a sign or a number: for example, in the books of deliberations of the Venetian Republic, a cross written in the left margin of the text meant that the proposal had been accepted. In accounting books using the double-entry system where debits and credits must be equal, this is carried out by making corresponding notes, such as ticks or marks, on both sides of the register. With regard to assets and inventories, the items are ticked off when the book keepers perform an audit [70, pp. 49–74]. All these administrative and managing annotations can be considered both descriptive metadata and context [75, pp. 169–195]. This use of annotation can also be found today in the public administration.

A further use of annotations was in administrative documents where head clerks wrote annotations to indicate who was responsible for a task. This kind of annotation demonstrates the concrete organization of administrative offices and work-flow among employees. Moreover, on analogical documents in the nineteenth and twentieth centuries there were annotations about archival management, indicating when a document had

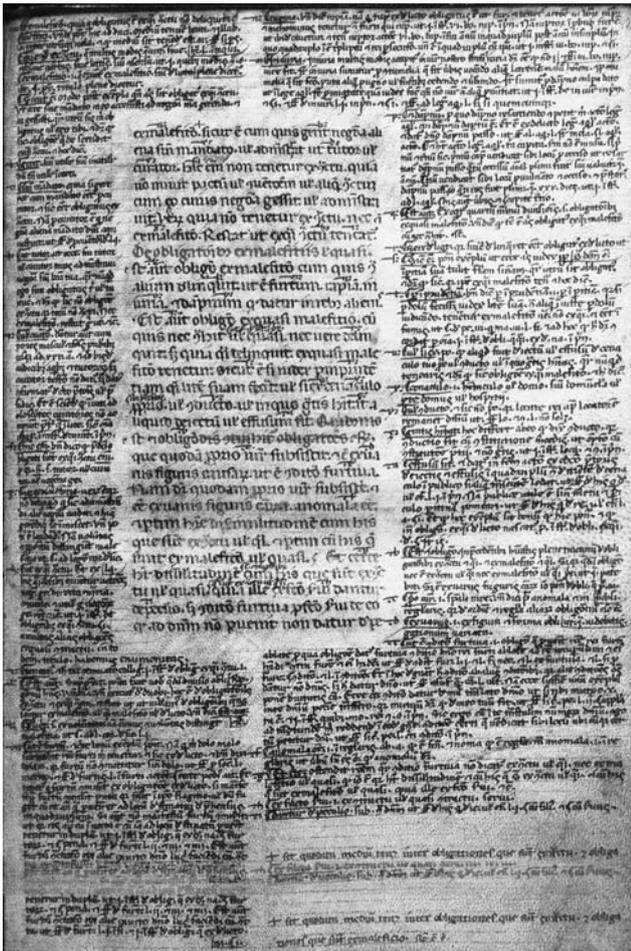
to be discarded from the archive. As a final example, some stamps can be considered a further kind of annotations, because they certify the subsequent steps of different administrative procedures of public or private administrations [35, pp. 111–113].

All these examples of administrative annotations show that on analogical documents the same paper has to be used in order to represent different subsequent actions. Administrative people, especially in the Roman law context, were used to producing documents by using different forms of annotation: numbers, stamps, seals, signatures, signs, and so on. Therefore, a challenge for contemporary annotation systems is to provide support for such advanced uses of annotations, which heavily rely on an implicit temporal dimension.

In the bibliographical field there are some specific annotations, called property notes, which can be a name, a motto, a coat of arms, a symbolic drawing, and which are used by owners of the book to make explicit their property [18,19]. In this context, annotations become an important tool for bibliographical researches, who aim at rebuilding, at least virtually, the ancient dispersed libraries and exploit such kinds of annotations as a means for achieving their objective.

### 3.1.2 Historical glosses

The gloss was a practice that flourished especially in the juridical context, as reported by [54, pp. 427–429]. During the Roman Empire, one of the usual literary forms of Roman jurisprudence was the comment to the works of former jurists, so that it is often possible to distinguish the annotated text from the annotation to the text; furthermore, the glosses were sometimes physically separated from the annotated text. However, the most famous use of this kind of method of study is credited to the Bolognese school: indeed, the word gloss denoted the way of studying the Justinian Code practised in Bologna, which began in the twelfth century AD. The Bolognese gloss passed from a simpler form to a more complex one by passing from simple interlinear notes to a real theoretical treatment of the subject. The glossarist reveals the contradictions (*contrarietates*) of the Justinian books, raises doubts (*dubiationes* or *dubietates*), which often give rise to controversies (*dissensiones*). The contradictions often find an explanation (*solutio*) and the doubts disappear by means of an appropriate distinction (*distinctio* or *differentia*). The glossarist teaches the Justinian books and creates cases in point and examples that originate glosses pointing out the different cases (*casus*); furthermore, the glossarist establishes and defines rules derived from the texts he studies, and, accordingly, creates glosses that report such



**Fig. 2** The photo shows the typical structure of the gloss: the author’s commentary, written in smaller characters, is placed around the text, written in greater characters (Italia, Bologna, Biblioteca Comunale dell’Archiginnasio, Salatiele, *Ars notariae*, ms B 1484, c. 12r—published with the permission of Biblioteca Comunale dell’Archiginnasio di Bologna)

rules (*regulae*) and definitions (*definitiones*). In conclusion, the Bolognese gloss was a way of doing research aimed at defining and elucidating the law.

The glosses were usually arranged on a page around the text, as shown in Fig. 2, which reproduces a page from the work *Ars notariae* (Handbook for notaries) written by Salatiele, one of the professors at the University of Bologna in the Middle Ages [48, 79]. The logical structure of glosses, seen as a professorial tool for high teaching and as a system adopted by jurisprudence to comment on codes, has become a solid method in the scientific society. The notes, both footnotes and final notes, are used by authors in order to justify assertions written in the text through quotations of sources and bibliography, to discuss assertions of other authors, and to give some practical examples.

In the philological and diplomatic fields the editors of literary texts and historical sources used a double array

of notes: on the one side, the notes on the text illustrate the different text-witness and justify the editor’s choices; on the other side, the historical notes precisely identify and correct names of people and places, which may be erroneously reported in the text or may not be explicitly mentioned [17, 39]. Therefore, the use of notes turns out to be a work method and a scientific mentality.

### 3.1.3 Historical notes

The scientific use of notes derived from the glosses is connected to a new kind of book created in medieval Universities, where Roman law was taught. Ancient books were compact and unitary objects, they were not organized into chapters and sections, and the whole page was used for writing the text. They were intended for reading and not for studying and, as a consequence, they had to be read from the beginning to the end without the possibility of skimming irrelevant content. In contrast, in the Universities books were used by professors and students to study and, especially, to comment and annotate the codes. To this end, the University book was written in two columns with large margins for commentaries. This book, particularly consecrated to handbooks and codes, develops into different chapters and paragraphs in order to offer readers a structured and easy to consult text which can be used without the need to read the entire book [48].

The logical structure of University books and, especially, the possibility of unambiguously identifying each piece of text by a hierarchical system of numbers and characters can be considered a marking system ahead of its time. Indeed, the medieval books were created by academic people of the Universities in order to exactly quote a specific passage of a code [41]. This logical system of quotation is a necessary pre-requisite for the following development of the scientific way of working.

In addition, the structure of the new University book can be considered as a starting-point towards the systematic and alphabetic index. Since the birth of University studies, the index system has become more complicated and refined. In short, the exact quotation cannot exist without the structure of text, based on logical subdivision of text and on ontological recognition of structural elements.

This important lesson learned from the past should also be taken into account when we design an annotation management system, because the possibility of digitally annotating and quoting can be a valid support to the research work in a networked and distributed environment.

### 3.2 Present uses

In the following sections, we will go into more detail about the current viewpoints concerning annotations and we will present some interesting cases of the use of annotations and *Information Management Systems (IMs)* with annotation capabilities in the context of digital libraries, the Web, and databases.

#### 3.2.1 Digital libraries

Digital libraries are not only the digital versions of traditional libraries and archives, but also they offer means which go beyond the mere presentation of the content stored in digital repositories. In the following we point out this fact by discussing two definitions of digital libraries which come from two different fields. The more computer science oriented view is expressed in the introduction in the first issue of the *International Journal on Digital Libraries*, cited by [47]:

Digital Libraries are concerned with the creation and management of information resources, the movement of information across global networks and the effective use of this information by a wide range of users.

Librarians have a different definition of *Digital Library (DL)*, as proposed by the Digital Library Federation, 1998, cited by [47]:

Digital Libraries are organisations that provide the resources, including the specialised stuff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.

Annotations can be exploited in order to provide users with the distinguishing features of a *DLMS*, as highlighted above. Note, however, that also archives have to be taken into account in this context, although they are not explicitly mentioned in the previous definitions. The *creation* of new information resources is supported by annotations in two ways. First, when users add annotations to existing information resources, they in turn become new information resources themselves. Second, annotations can also assist in the creation of new information resources. Through annotations, new ideas and concepts can be discussed and the results of such

a discussion can then be integrated into the newly created object. Annotations might increase and expand the information resources managed by the digital library. In this way, they may provide *interpretations* of information resources. User communities benefit from such interpretations in that they help the understanding of the annotated resource and contain additional information about it. As an example, in the Humanities, interpretation is one of the basic tasks scholars perform: systems like *COLLATE* [45,84] or *IPSA* [2,11,12] support this task through annotations. Annotations support user communities in *accessing* the information resources provided by the digital library in a personalized and customized way: indeed, users can create annotations that link different documents, enabling alternative paths for browsing digital contents and thus structuring them in alternative ways.

Different layers of annotations can coexist on the same document: a private layer of annotations accessible only by the annotations author themselves, a collective layer of annotations, shared by a team of people, and finally a public layer of annotations, accessible to all the users of the digital library. In this way, user communities can benefit from different views of the information resources managed by the digital library [64,66,67]. A *DL* can encourage cooperative work practices, enabling the sharing of documents and annotations, also with the aid of special devices, such as *XLibris* [80]. Finally, as suggested in [68,69], searching, reading and annotating a *DL* can be done together with other activities, for example working with colleagues. This may also occur in a mobile context, where merging content and wireless communication can foster ubiquitous access to *DLMS*, improving well established cooperative practices of work and exploiting physical and digital resources. The wireless context and the small form factor of handheld devices challenge our technical horizons for information management and access and require specialized solutions to overcome the constraints imposed by such kinds of devices, as reported in [6].

In the context of a *DLMS* it is also possible to create automatic annotations, which may facilitate the user's first approach with a document. Automatic annotations can be created by using topic detection techniques to associate each annotation with its related topic, which constitutes the context of the annotation. In this way, a document can be re-organized and segmented into topics, whose dimension can range in many different sizes, and annotations can present a brief description of those topics. Then, by applying automatic hypertext construction techniques, similar to those presented in [14], those pairs of topics and annotations can be linked together,

proposing an alternative way of navigating the content of a digital library.

Moreover, *DLMS* are currently in a state of evolution: today they are simply places where information resources can be stored and made available, whereas by tomorrow they will become an integrated part of the way the user works. For example, instead of simply downloading a paper and then working on a printed version, a user will be able to work directly with the paper by means of the tools provided by the *DLMS* and share their work with colleagues. In this way, the user's intellectual work and the information resources provided by the *DLMS* can be merged together to constitute a single working context. Thus, the *DLMS* is no longer perceived as something external to the intellectual production process and nor as a mere consulting tool, but instead as an intrinsic and active part of the intellectual production process [7].

Annotations are effective means to enable the paradigm of interaction between users and *DLMS*s envisioned above, since they are very well-established practices and widely used. Annotations are not only a way of explaining and enriching an information resource with personal observations, but also a means of transmitting and sharing ideas to improve collaborative work practices. Therefore, annotations can be geared not only to the way of working of the individual and to a method of study but also to a way of doing research, as happens in the Humanities. Finally, annotations allow users to naturally merge and link personal contents with the information resources provided by the *DLMS* so that a common context that unifies all of these contents can be created.

In this last respect, documents managed by the *DLMS* and annotations constitute a hypertext [8,10], since annotations allow the creation of new relationships among existing objects, by means of links that connect annotations together with existing objects. Halasz [50] points out that annotations are one of the activities that form the basis of any collaborative effort and for which hypermedia systems are ideally suited, while [65] considers annotations as a natural way of creating and growing hypertexts that connect information resources by actively engaging users. Moreover, *DLMS*s do not normally have a hypertext connecting information resources with each other; thus, annotations can turn out to be an effective way of associating a hypertext to a *DLMS* to enable an active and dynamic use of information resources. This hypertext can span and cross the boundaries of the single *DLMS*, if users need to interact with the information resources managed by diverse *DLMS*s [7]. This latter possibility is quite innovative, because it offers the means for interconnecting various

*DLMS*s in a personalized way meaningful for the end-user and, as recognized also by [59], is a big challenge for the next generation *DLMS*s.

### 3.2.2 The Web

As previously introduced, the Annotea project [61] considers annotations as metadata. Annotea defines annotations as comments, notes, explanations, or other types of external remarks that can be attached to any Web document or a selected part of the document without modifying the document. Annotea uses *Resource Description Framework (RDF)*<sup>4</sup> and *eXtensible Markup Language (XML)*<sup>5</sup> for describing annotations as metadata and *XPointer*<sup>6</sup> for locating the annotations in the annotated document. Annotea employs a client-server architecture, where annotations reside in dedicated servers and a specialized browser is capable of retrieving them upon request, when visiting a Web page. [62] move a step further and employ annotations as an extension of the bookmarks to improve collaboration among users: indeed, the additional data provided by annotations are exploited to describe, organize, categorize, share, and search for the bookmarks.

Moreover, the W3C Multimodal Interaction Working Group<sup>7</sup> is developing a markup language called *Extensible MultiModal Annotation (EMMA)* [85]. *EMMA* is a markup language intended for providing semantic interpretations for a variety of inputs, such as speech, natural language text, and *Graphical User Interface (GUI)* input. The language is focused on annotating the interpretation information of single and composed inputs, and it is expected that this markup will be used primarily as a standard data interchange format between the components of a multimodal system. The general purpose of *EMMA* is to represent information automatically extracted from a user's input by an interpretation component. *EMMA* provides a simple structural syntax for the organization of interpretations and instances, and an annotative syntax derived from *RDF* to apply the annotation to the input data at any level.

As a further example, *Multimedia Annotation of Digital Content Over the Web (MADCOW)* is based on a client-server architecture as Annotea is. Servers are repositories of annotations to which different clients can connect, while the client is a plug-in for a standard Web browser [27]. *MADCOW* employs *HyperText*

<sup>4</sup> <http://www.w3.org/RDF/>.

<sup>5</sup> <http://www.w3.org/XML/>.

<sup>6</sup> <http://www.w3.org/XML/Linking>.

<sup>7</sup> <http://www.w3.org/2002/mmi/Group/>.

*Transfer Protocol (HTTP)* to annotate Web resources and allows both private and public annotations. Moreover, it allows different pre-established types of annotations, such as explanation, comment, question, solution, summary, and so on; in this respect, *MADCOW* opts for a solution similar to the one of *COLLATE*, which is not Web-based but it models annotations as different types of dialog acts [45].

### 3.2.3 Databases

Annotations are also used in the context of the DBMSs and, in particular, in the case of *curated databases* and *scientific databases*. *SWISS-PROT*<sup>8</sup> is a curated protein sequence database, which strives to provide a high level of annotation, such as the description of the function of a protein, its domains structure, and so on. In this case, the annotations are embedded in the database and merged with the annotated content. *BIODAS*<sup>9</sup> provides a *Distributed Annotation System (DAS)*, which is a Web-based servers system for sharing lists of annotations across a certain segment of the genome. In this case, the annotation are not mixed together with the content they annotate, but instead are separated from it. Annotations have types, methods, and categories: the annotation type is selected from a list of types that have biological significance; the annotation method is intended to describe how the annotated feature was discovered and may include a reference to a software program; the annotation category is a broad functional category that can be used to filter, group, and sort annotations. Finally, annotations may also be associated with Web Uniform Resource Locators (URLs) that provide additional human readable information about the annotation itself [83].

In the context of scientific databases, [31] propose an archiving technique for managing and archiving different versions of such kinds of databases, as time moves on. Buneman et al. [31] exploit the hierarchical structure of scientific data to represent the content and the different versions of the database with a tree structure. They attach annotations to the nodes of the tree, annotations that contain time-stamp and key information about the underlying data structure. Therefore, these annotations are metadata about the database itself. These annotations differ from the annotations contained in the database, which are metadata about genome sequences. In conclusion, this annotated tree structure provides an additional data layer which allows the development of

efficient algorithms in order to archive and search for the different versions of the database.

Buneman et al. [32,33] investigate the use of annotations with respect to the *data provenance* problem, sometimes also referred to as *data lineage* or *data pedigree*, which is the description of the origins of a piece of data and the process by which it arrived in a database. Buneman et al. [32] distinguish between *why-provenance*, which explains why a given piece of data is in the database, and *where-provenance*, which explains where a given piece of data comes from. Data provenance is a relevant issue in the field of curated and scientific databases, such as genome databases, because in this field there are few databases that are sources of data; all the other databases are in a certain sense views of these source databases or of other views. The distinguishing feature of these databases is the fact that they have to be curated: in fact, they provide corrections and annotations to the original source data made by experts. It is now clear that data provenance is essential to any user interested in the accuracy and timeliness of the data. In particular, where-provenance is important for understanding the source of errors in data and for carrying annotations through database queries, problems addressed in [33]. Bhagwat et al. [24] carries on the research about where-provenance and propose and implement an extension to a relational *DBMS* and an extension to *Structured Query Language (SQL)*, called *propagate SQL (pSQL)*, which provides a clause for propagating annotations to tuples through queries. Bhagwat et al. [24] intend annotations to be an information about data such as provenance, comments, or other types of metadata; they envisage the following applications of annotations in *DBMS*: tracing the provenance and flow of data, reporting errors or remarks about a piece of data, and describing the quality or the security level of a piece of data.

### 3.2.4 Using annotation for search purposes

Despite all of the research in modelling annotations and providing annotation-enabled systems, there is much less study regarding the use of annotations for retrieving documents.

Golovchinsky et al. [49] compare queries based on annotations with relevance feedback, and consider annotation-based queries as an automatic technique for query construction, since queries are automatically generated from annotated text, e.g. from highlighted text.

Frommholz et al. [45] consider annotations—specifically annotations threads—as an extension of the document they belong to, creating a discourse context, in which not only the annotation itself but also its position in the discourse and its type, are exploited for searching

<sup>8</sup> <http://www.expasy.org/sprot/>.

<sup>9</sup> <http://biodas.org/>.

and retrieving documents; this approach is revised and extended upon in [46] to probabilistic datalog.

Agosti and Ferro [8,9] address this issue in the context of data fusion [38]. In fact, annotations provide us with an additional source of evidence, which is complementary to the one already contained in the set of documents. Therefore, we can exploit annotations with two final objectives of retrieving more relevant documents and of ranking them better. Furthermore, the paths that connect annotations to documents in the hypertext become the vehicle for moving this further source of evidence from annotation towards the documents. Thus, both *Hypertext Information Retrieval (HIR)* techniques [16] and *link fusion* techniques [86] can be exploited in order to provide an advanced search strategy that involves annotations. In addition, the mutual reinforcement between annotations and annotated objects should also be taken into account, for example, by employing link analysis techniques, such as those described in [15].

## 4 Key features for system design

On the whole, the line of reasoning conducted in the previous sections provides us with some key features of the annotation that we should take into account when designing systems that have to support the management of digital annotations on digital contents. Those key features are presented in the following subsections.

### 4.1 First-class intellectual work

Annotations are a valuable intellectual work, as it emerges from the discussion in Sect. 2.1 and from the user studies reported in Sect. 2.2. The spectrum of this intellectual work is very broad, because it ranges from explaining and enriching an information resource with personal observations to transmitting and sharing ideas and knowledge on a subject. In conclusion, annotations can be geared not only to the way of working of the individual and to a method of study, but also to a way of doing research.

As a consequence, during the design of a system with annotation functionalities, great attention should be paid to providing the tools to support and facilitate the creation of such valuable intellectual work, from its early stages, such as informal jottings, to its more mature and structured formulations. This requires not only designing a system with the needed functionalities but also taking care in designing the user interface of the system which should make such functionalities easily accessible [13].

### 4.2 Various facets

Annotations comprise different viewpoints, as discussed in Sect. 2.2: they may be considered as metadata, content, hypertext, context, or dialog acts. Moreover, the boundaries between these viewpoints are not sharp and they may coexist. All of these viewpoints have to be taken into account, especially because they are tightly coupled with and are the expression of the different kinds of intellectual work that an annotation may bear.

This wide range of uses requires an additional flexibility to an annotation management system, which should seamlessly support the different ways of using annotations. As a consequence, the annotation management system should avoid tying the annotation to only one of the described viewpoints or, worse, offering a fixed set of predefined annotation types.

### 4.3 Different scopes

Annotations involve different scopes and different kinds of annotative context: they can be private, shared or public, according to the type of intellectual work that is carried out. Moreover, the boundaries between these scopes are not fixed rather they may vary and evolve as the time passes.

This feature requires the careful design of the management of both users and group of users in order to support different and fine-grained access policies. For example, a shared annotation could be visible only with read permissions for a group, while another group could have read/write permissions on the same annotation.

Moreover, the system should avoid *scope conflicts*. For example, a private annotation cannot be annotated by a public annotation; in this case, the author of the private annotation could see both the public and the private annotation, but another user could see only the public annotation which would be annotating something hidden to this user.

### 4.4 Active involvement

Annotations call for an active involvement, the degree of which varies according to the aim of the annotation: private annotations requires the involvement of the authors, although shared or public annotations involve the participation of a whole community. Therefore, annotations are suitable for improving collaboration and co-operation among users.

Many different IMSs can benefit from annotations and from the improved collaboration among their users. For these reasons, annotations functionalities should not

be embedded in any given system, but rather in a stand-alone service capable of providing annotation functionalities to other systems.

In this way, the user would benefit from a uniform way of interaction with annotation functionalities, without the need of changing their annotative practices only because a user works with different IMSs. Finally, such kinds of annotation services would act as a bridge between different IMSs and allows the hypertext between documents and annotations to cross the boundaries of the single *IMS*, in order to further exploit annotations as an active and effective collaboration tool for users.

#### 4.5 Temporal dimension

Annotations implicitly entail a temporal dimension which regulates the temporal ordering among annotations and annotated information resources.

This temporal dimension needs to be taken into account and to be explicitly modelled, since it implicitly influences each feature of the annotations and, in addition, it allows us to trace the provenance of each piece of knowledge managed by the system.

#### 4.6 Search by using annotations

Annotations provide an additional source of evidence that can be effectively exploited to search for documents. Furthermore, this additional source of evidence holds complex relationships with the annotated documents. These complex relationships can be made explicit by using the hypertext that connects documents and annotation. In this way, the hypertext becomes the vehicle for moving this source of evidence from annotations towards documents, so that it is possible to develop advanced search strategies. Moreover, the mutual reinforcement existing between annotations and annotated objects should also be taken into account when defining a search strategy that involves annotations.

As a consequence of making the annotation service a stand-alone service which can be used by different IMSs, we deal with a distributed search problem. Indeed, we aim at combining multiple sources of evidence which come from both documents and annotations. Since the source of evidence concerning the documents is completely managed by the IMSs, the annotations service has to query each *IMS* in order to obtain it. Only once the annotation service has acquired this information from the *IMS*, can it be combined with the source of evidence which comes from annotations to create a list of result documents that better satisfies the user's information needs.

**Acknowledgments** The work reported in this paper was partially funded by the Italian Ministry of Education (MIUR) under the Project of Relevant National Interest (PRIN) called "Methods for a digital corpus of illuminated manuscripts" (2005–2006). The work was also partially supported by the DELOS Network of Excellence on Digital Libraries, as part of the Information Society Technologies (IST) Program of the European Commission (Contract G038-507618).

#### References

1. Agosti, M.: An overview of hypertext. In: Agosti, M., Smeaton, A. (eds.) *Information Retrieval and Hypertext*. pp. 27–47. Kluwer Academic, Norwell (1996)
2. Agosti, M., Benfante, L., Orio, N.: IPISA: a digital archive of herbals to support scientific research. In: Sembok, T.M.T., Zaman, H.B., Chen, H., Urs, S.R., Myaeng, S.H. (eds.) *Proceedings of the 6th International Conference on Asian Digital Libraries—Digital Libraries: Technology and Management of Indigenous Knowledge (ICADL 2003)*, pp. 253–264. *Lecture Notes in Computer Science (LNCS)*, vol. 2911, Springer, Heidelberg (2003)
3. Agosti, M., Di Nunzio, G.M., Ferro, N.: A data curation approach to support in-depth evaluation studies. In: Gey, F.C., Kando, N., Peters, C., Lin, C.-Y. (eds.) *Proceedings of the International Workshop on New Directions in Multilingual Information Access (MLIA 2006)*, pp. 65–68. <http://ucdata.berkeley.edu/sigir2006-mlia.htm> [last visited 2006 October 2], 2006
4. Agosti, M., Di Nunzio, G.M., Ferro, N.: Scientific data of an evaluation campaign: do we properly deal with them? In: Nardi, A., Peters, C., Vicedo, J.L. (eds.) *Working Notes for the CLEF 2006 Workshop*. [http://www.clef-campaign.org/2006/working\\_notes/workingnotes2006/agostiCLEF2006.pdf](http://www.clef-campaign.org/2006/working_notes/workingnotes2006/agostiCLEF2006.pdf) [last visited 2006, October 2], 2006
5. Agosti, M., Ferro N.: Annotations: enriching a digital library. In: Koch, T., Sølvberg, I.T. (eds.) *Proceedings of the 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003)*, pp. 88–100. *Lecture Notes in Computer Science (LNCS)*, vol. 2769. Springer, Heidelberg (2003)
6. Agosti, M., Ferro, N.: Managing the interactions between handheld devices, mobile applications, and users. In: Lim, E.P., Siau, K. (eds.) *Advances in Mobile Commerce Technologies*, chap. 10, pp. 204–233. Idea Group, Hershey (2003)
7. Agosti, M., Ferro, N.: A system architecture as a support to a flexible annotation service. In: Türker, C., Agosti, M., Schek, H.-J. (eds.) *Peer-to-Peer, Grid, and Service-Oriented in Digital Library Architectures: 6th Thematic Workshop of the EU Network of Excellence DELOS. Revised Selected Papers*, pp. 147–166. *Lecture Notes in Computer Science (LNCS)*, vol. 3664. Springer, Heidelberg (2005)
8. Agosti, M., Ferro, N.: Annotations as context for searching documents. In: Crestani, F., Ruthven, I. (eds.) *Proceedings of the 5th International Conference on Conceptions of Library and Information Science—Context: Nature, Impact and Role*, pp. 155–170. *Lecture Notes in Computer Science (LNCS)*, vol. 3507. Springer, Heidelberg (2005)
9. Agosti, M., Ferro, N.: Search Strategies for finding annotations and annotated documents: the FAST Service. In: Legind Larsen, H., Pasi, G., Ortiz-Arroyo, D., Andreasen, T., Christiansen, H. (eds.) *Proceedings of the 7th International Conference on Flexible Query Answering Systems (FQAS 2006)*, pp. 270–281. *Lecture Notes in Artificial Intelligence (LNAI)*, vol. 4027. Springer, Heidelberg (2006)

10. Agosti, M., Ferro, N., Frommholz, I., Thiel, U.: Annotations in digital libraries and laboratories—facets, models and usage. In: Heery, R., Lyon, L. (eds.) Proceedings of the 8th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2004), pp. 244–255. Lecture Notes in Computer Science (LNCS), vol. 3232. Springer, Heidelberg (2004)
11. Agosti, M., Ferro, N., Orio, N.: Annotating illuminated manuscripts: an effective tool for research and education. In: Marlino, M., Sumner, T., Shipman, F. (eds.) Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2005), pp. 121–130. ACM Press, New York (2005)
12. Agosti, M., Ferro, N., Orio, N.: Graph-based automatic suggestion of relationships among images of illuminated manuscripts. In: Haddad, H.M., Liebrock, K.M., Chbeir, R., Palakal, M.J., Ossowski, S., Yetongnoon, K., Wainwright, R.L., Nicolle, C. (eds.) Proceedings of the 21st ACM Symposium on Applied Computing (SAC 2006), pp. 1063–1067. ACM Press, New York (2006)
13. Agosti, M., Ferro, N., Panizzi, E., Trinchese, R.: Annotation as a support to user interaction for content enhancement in digital libraries. In: Celentano, A., Mussio, P. (eds.) Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2006), pp. 151–154. ACM Press, New York (2006)
14. Agosti, M., Melucci, M.: Information retrieval techniques for the automatic construction of hypertext. In: Kent, A., Hall, C.M. (eds.) Encyclopedia of Library and Information Science, vol. 66, pp. 139–172. Marcel Dekker, New York, (2000)
15. Agosti, M., Pretto, L.: A theoretical study of a generalized version of Kleinberg's HITS algorithm. *Inf. Retr.* **8**(2), 219–243 (2005)
16. Agosti, M., Smeaton, A. (eds.): Information Retrieval and Hypertext. Kluwer Academic, Norwell (1996)
17. Avalle, d'A.S.: Principi di critica testuale. *Vulgares Eloquentes*, vol. 7. Antenore, Padova (1972)
18. Barbieri, E. (ed.): Nel mondo delle postille: i libri a stampa con note manoscritte. Una raccolta di studi. *Humanae litterae*, 6, CUSL, Milano, Italia, (2002)
19. Barbieri, E., Frasso, G. (eds.): Libri a stampa postillati. *Atti del Colloquio internazionale (Milano, 3–5 lug. 2001)*. *Humanae litterae*, 8, CUSL, Milano, Italia (2003)
20. Bartoli Langelì, A.: La documentazione degli stati italiani nei secoli XIII–XV: forme, organizzazione, personale. In: *Culture et idéologie dans la genèse de l'État moderne. Actes de la table ronde organisée par le Centre National de la recherche scientifique et l'École Française de Rome (Roma, 15–17 ottobre 1984)*, pp. 35–55. Collection de l'École Française de Rome, 82, École Française, Roma, Italia (1985)
21. Bartoli Langelì, A.: La documentazione degli stati italiani nei secoli XIII–XV: forme, organizzazione, personale. In: Albinì, G. (ed.) *Le scritture del Comune. Amministrazione e memoria nelle città dei secoli XII e XIII*, pp. 155–171. I Florigeri, XII, Paravia, Torino, Italia, (1998)
22. Bartoli Langelì, A.: Notai. Scrivere documenti nell'Italia medievale. *I libri di Viella*, vol. 56. Viella, Roma (2006)
23. Battelli, I.: *Exempla scripturarum edita consilio et opera procuratorum Bibliothecae et Tabularii Vaticani. Fasciculus III. Acta pontificum (editio secunda aucta)*. apud Bibliothecam Vaticanam, Vaticani (1965)
24. Bhagwat, D., Chiticariu, L., Tan, W.-C., Vijayvargiya, G.: An annotation management system for relational databases. In: Nascimento, M.A., Özsu, M.T., Kossmann, D., Miller, R.J., Blakeley, J.A., Schiefer, K.B. (eds.) Proceedings of the 30th International Conference on Very Large Data Bases (VLDB 2004), pp. 900–911. Morgan Kaufmann San Francisco (2004)
25. Bonfiglio-Dosio, G.: Un nuovo titolare per gli archivi dei comuni italiani? Un contributo per la soluzione di una questione archivistica nazionale. *Bollettino della Deputazione di storia patria per l'Umbria*, vol. XCIX, pp. 355–380, 2002
26. Bonfiglio-Dosio, G.: *Primi passi nel mondo degli archivi. Temi e testi per la formazione archivistica di primo livello (seconda edizione riveduta e ampliata)*. CLEUP, Padova (2005)
27. Bottoni, P., Civica, R., Levialdi, S., Orso, L., Panizzi, E., Trinchese, R.: MADCOW: a multimedia digital annotation system. In: Costabile, M.F. (ed.) Proceedings of Working Conference on Advanced Visual Interfaces (AVI 2004), pp. 55–62. ACM Press, New York (2004)
28. Bottoni, P., Civica, R., Levialdi, S., Orso, L., Panizzi, E., Trinchese, R.: Storing and retrieving multimedia web notes. In: Bhalla, S. (ed.) Proceedings of the 4th International Workshop on Databases in Networked Information Systems (DNIS 2005), pp. 119–137. Lecture Notes in Computer Science (LNCS), vol. 3433. Springer, Heidelberg (2005)
29. Bottoni, P., Levialdi, S., Rizzo, P.: An analysis and case study of digital annotation. In: Bianchi-Berthouze, N. (ed.) Proceedings of the 3rd International Workshop on Databases in Networked Information Systems (DNIS 2003), pp. 216–230. Lecture Notes in Computer Science (LNCS), vol. 2822. Springer, Heidelberg (2003)
30. Bresslau, H.: *Manuale di diplomazia per la Germania e l'Italia*. Traduzione di Anna Maria Voci-Roth. Pubblicazioni degli Archivi di Stato, Sussidi, vol. 10. Ministero per i Beni Culturali e Ambientali—Ufficio Centrale Beni Archivistici, Roma (1998)
31. Buneman, P., Khanna, S., Tajima, K., Tan, W.-C.: Archiving scientific data. *ACM Trans. Database Syst. (TODS)* **29**(1), 2–42 (2004)
32. Buneman, P., Khanna, S., Tan, W.-C.: Why and where: a characterization of data provenance. In: Van den Bussche, J., Vianu, V. (eds.) Proceedings of the 8th International Conference on Database Theory (ICDT 2001), pp. 316–330. Lecture Notes in Computer Science (LNCS), vol. 1973. Springer, Heidelberg (2001)
33. Buneman, P., Khanna, S., Tan, W.-C.: On propagation of deletions and annotations through views. In: Abiteboul, S., Kolaitis, P.G., Popa, L. (eds.) Proceedings of the 21st ACM SIGMOD–SIGACT–SIGART Symposium on Principles of Database Systems (PODS 2002), pp. 150–158. ACM Press, New York (2002)
34. Calonghi, F.: *Dizionario latino-italiano*, 3 edizione (interamente rifusa ed aggiornata del dizionario Georges — Calonghi) — 15 tiratura. Rosenberg and Sellier, Nuova Offlito, Mappano, Torino (1986)
35. Carucci, P.: *Il documento contemporaneo. Diplomatica e criteri di edizione*, Beni culturali, vol. 1. NIS, Roma (1987)
36. Constantopoulos, P., Doerr, M., Theodoridou, M., Tzobanakis, M.: On information organization in annotation systems. In: Griesser, G., Tanaka, Y. (eds.) Proceedings of the International Workshop on Intuitive Human Interfaces for Organizing and Accessing Intellectual Assets, pp. 189–200. Lecture Notes in Computer Science (LNCS), vol. 3359. Springer, Heidelberg (2004)
37. Cortelazzo, M., Zolli, P.: *DELI: Dizionario Etimologico della Lingua Italiana*, 2 edizione a cura di M. Cortelazzo e M. A. Cortelazzo. Zanichelli, Bologna (1999)

38. Croft, W.B.: Combining approaches to information retrieval. In: Croft, W.B. (ed.) *Advances in Information Retrieval: Recent Research from the Center for Intelligent Information Retrieval*, pp. 1–36. Kluwer Academic, Norwell (2000)
39. Commission Internationale de Diplomatique Commission Internationale de Sigillographie: *Diplomatica et sigillographica. Folia Caesaraugustana*, vol. 1. Institucion Fernando el Catolico, Zaragoza (1984)
40. de la Sala, F., Rabikauskas, P.: *Il documento medievale e moderno. Panorama storico della diplomazia generale pontificia*. Pontificia Università Gregoriana — Istituto portoghese di Sant’Antonio, Roma (2003)
41. Delsalle, P.: *Une histoire de l’archivistique*. Presses de l’Université du Québec, Québec (1998)
42. Ferro, N.: *Formal model and conceptual architecture of the annotation service for dynamic ubiquitous knowledge environments*. PhD thesis, Department of Information Engineering, University of Padua, Italy, December 2004
43. Fogli, D., Fresta, G., Mussio, P.: On electronic annotation and its implementation. In: Costabile, M.F. (ed.) *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2004)*, pp. 98–102. ACM Press, New York (2004)
44. Frenz, T.: *I documenti pontifici nel Medioevo e nell’età moderna*. Edizione italiana a cura di Sergio Pagano. *Littera antiqua*, vol. 6. Scuola Vaticana di paleografia, diplomazia e archivistica, Città del Vaticano (1989)
45. Frommholz, I., Brocks, H., Thiel, U., Neuhold, E., Iannone, L., Semeraro, G., Berardi, M., Ceci, M.: Document-centered collaboration for scholars in the humanities — The COLLATE system. In: Koch, T., Sølvberg, I.T. (eds.) *Proceedings of the 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003)*, pp. 434–445. *Lecture Notes in Computer Science (LNCS)*, vol. 2769. Springer, Heidelberg (2003)
46. Frommholz, I., Thiel, U., Kamps, T.: Annotation-based document retrieval with four-valued probabilistic datalog. In: Baeza-Yates, R., Maarek, Y., Roelleke, T., de Vries, A.P. (eds.) *Proceedings of the 3rd XML and Information Retrieval Workshop and the 1st Workshop on the Integration of Information Retrieval and Databases (WIRD2004)*, pp. 31–38. <http://homepages.cwi.nl/~arjen/wird04/wird04-proceedings.pdf> [last visited 2006, October 2], 2004
47. Fuhr, N., Hansen, P., Micsik, A., Sølvberg, I.: Digital libraries: a generic classification scheme. In: Constantopoulos, P., Sølvberg, I.T. (eds.) *Proceedings of the 5th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2001)*, pp. 187–199. *Lecture Notes in Computer Science (LNCS)*, vol. 2163. Springer, Heidelberg (2001)
48. Giovè Marchioli, N.: *Gli strumenti del sapere. I manoscritti padovani tra tipizzazioni generali e peculiarità locali*. In: Piovani, F., Sitran Rea, L. (eds.) *Studenti, Università, città nella storia padovana. Atti del convegno (Padova, 6–8 febbraio 1998)*, pp. 47–71. Centro per la storia dell’Università di Padova. *Contributi*, 35, Lint, Trieste, Italia, 2001
49. Golovchinsky, G., Price, M.N., Schilit, B.N.: From reading to retrieval: freeform ink annotations as queries. In: Gey, F., Hearst, M., Tong, R., (eds.) *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 1999)*, pp. 19–25. ACM Press, New York (1999)
50. Halasz, F.G.: Reflections on note cards: seven issues for the next generation of hypermedia systems. *Commun. ACM (CACM)* **31**(7), 836–852 (1988)
51. Handschuh, S., Staab, S. (eds.): *Annotation for the Semantic Web*. IOS Press, Amsterdam (2003)
52. Hanks, P. (ed.): *Collins Dictionary of the English Language*. William Collins, Glasgow (1979)
53. IEI.: *Enciclopedia Italiana di scienze, lettere ed arti*, vol. XXIV (SCAR–SOC). Istituto della Enciclopedia Italiana, Istituto Poligrafico dello Stato, Roma (1950)
54. IEI.: *Enciclopedia Italiana di scienze, lettere ed arti*, vol. XVII (GIAP–GS). Istituto della Enciclopedia Italiana, Istituto Poligrafico dello Stato, Roma (1951)
55. IEI.: *Vocabolario della lingua italiana*, vol. I (A–C). Istituto della Enciclopedia Italiana, Arti Grafiche Ricordi, Monotopia Olivieri, Milano (1986)
56. IEI.: *Vocabolario della lingua italiana*, vol. II (D–L). Istituto della Enciclopedia Italiana, Arti Grafiche Ricordi, Monotopia Olivieri, Milano (1987)
57. IEI.: *Vocabolario della lingua italiana*, vol. III\*\* (Pe–R). Istituto della Enciclopedia Italiana, Arti Grafiche Ricordi, Monotopia Olivieri, Milano (1991)
58. IEI.: *Vocabolario della lingua italiana*, vol. IV (S–Z). Istituto della Enciclopedia Italiana, Arti Grafiche Ricordi, Monotopia Olivieri, Milano (1994)
59. Ioannidis, Y., Maier, D., Abiteboul, S., Buneman, P., Davidson, S., Fox, E.A., Halevy, A., Knoblock, C., Rabitti, F., Schek, H.-J., Weikum, G.: Digital library information-technology infrastructures. *Int. J. Digit. Libr.* **5**(4), 266–274 (2005)
60. ISO.: *MPEG Standards—Coded representation of video and audio*. <http://www.iso.ch/iso/en/prodservices/popstds/mpeg.html> [last visited 2006, October 2], July 2004
61. Kahan, J., Koivunen, M.-R.: Annotea: an open RDF infrastructure for shared Web annotations. In: Shen, V.Y., Saito, N., Lyu, M.R., Zurko, M.E. (eds.) *Proceedings of the 10th International Conference on World Wide Web (WWW 2001)*, pp. 623–632. ACM Press, New York (2001)
62. Koivunen, M.-R., Swick, R., Prud’hommeaux, E.: Annotea Shared Bookmarks. <http://www.w3.org/2001/Annotea/Papers/KCAP03/annotaebm.html> [last visited 2006, October 2], October 2003
63. Mann, W.C., Thompson, S.A.: *Rhetorical structure theory: a theory of text organization*. Technical Report ISI/RS-87-190, Information Sciences Institute, University of Southern California, Los Angeles (CA), USA, June 1987
64. Marshall, C.C.: Annotation: from Paper Books to the Digital Library. In: Allen, R.B., Rasmussen, E. (eds.) *Proceedings of the 2nd ACM International Conference on Digital Libraries (DL 1997)*, pp. 233–240. ACM Press, New York (1997)
65. Marshall, C.C.: Toward an ecology of hypertext annotation. In: Akscyn, R. (ed.) *Proceedings of the 9th ACM Conference on Hypertext and Hypermedia (HT 1998): links, objects, time and space-structure in hypermedia systems*, pp. 40–49. ACM Press, New York (1998)
66. Marshall, C.C., Brush, A.J.B.: From personal to shared annotations. In: Terveen, L., Wixon, D. (eds.) *Proceedings of the Conference on Human Factors and Computing Systems (CHI 2002)—Extended Abstracts on Human Factors in Computer Systems*, pp. 812–813. ACM Press, New York (2002)
67. Marshall, C.C., Brush, A.J.B.: Exploring the relationship between personal and public annotations. In: Chen, H., Wactlar, H., Chen, C.-C., Lim, E.-P., Christel, M. (eds.) *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2004)*, pp. 349–357. ACM Press, New York (2004)
68. Marshall, C.C., Golovchinsky, G., Price, M.N.: Digital Libraries and Mobility. *Commun. ACM* **44**, 55–56 (2001)

69. Marshall, C.C., Ruotolo, C.: Reading-in-the-small: a study of reading on small form factor devices. In: Hersh, W., Marchionini, G. (eds.) Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2002), pp. 56–64. ACM Press, New York (2002)
70. Melis, F.: Documenti per la storia economica dei secoli XIII–CVI, con una nota di paleografia commerciale a cura di Elena Cecchi, volume 1 of Istituto internazionale di storia economica “F”. Datini Prato. Pubblicazioni—Serie I. Documenti. Olschki, Firenze (1972)
71. Nagao, K.: Digital Content Annotation and Transcoding. Artech House, Norwood (2003)
72. Neuhold, E., Niederée, C., Stewart, A., Frommholz, I., Mehta, B.: The role of context for information mediation in digital libraries. In: Chen, Z., Chen, H., Miao, Q., Fu, Y., Fox, E.A., Lim, E.-P. (eds.) Proceedings of the 7th International Conference on Asian Digital Libraries. Digital Libraries: International Collaboration and Cross-Fertilization (ICADL 2004), pp. 133–143. Lecture Notes in Computer Science (LNCS), vol. 3334. Springer, Heidelberg (2004)
73. Paoli, C.: Diplomatica. Nuova edizione aggiornata da G. C. Bascapè, con 220 disegni e fac-simili. Le lettere, Firenze (1987)
74. Phelps, T.A., Wilensky, R.: Multivalent annotations. In: Peters, C., Thanos, C. (eds.) Proceedings of the 1st European Conference on Research and Advanced Technology for Digital Libraries (ECDL 1997), pp. 287–303. Lecture Notes in Computer Science (LNCS), vol. 1324. Springer, Heidelberg (1997)
75. Pigiapoco, S.: La memoria digitale delle amministrazioni pubbliche. Requisiti metodi e sistemi per la produzione, archiviazione e conservazione dei documenti informatici. Progetto ente locale, vol. 148. Maggioli editore, Santarcangelo di Romagna (2005)
76. Pratesi, A.: Genesi e forme del documento medievale, Guide, vol. 3. Jouvance, Roma (1979)
77. Rocci, L.: Vocabolario greco italiano, 34 edizione. Società Editrice Dante Alighieri, Italia, 1989
78. Rossetti, L. (ed.): Matricula Nationis Germanicae artistarum in Gymnasio Patavino (1553–1721). Con la collaborazione di Giorgetta Bonfiglio-Dosio. Fonti per la storia dell’Università di Padova, vol. 10. Antenore, Padova (1986)
79. Salatiere.: Ars notariae. In: Orlandelli, G. (ed.) Opere dei maestri, vol. 2. Giuffrè, Milano (1961)
80. Schilit, B.N., Price, M.N., Golovchinsky, G.: Digital library information appliances. In: Witten, I., Akscyn, R., Shipman, F.M. (eds.) Proceedings of the 3rd ACM International Conference on Digital Libraries (DL 1998), pp. 217–226. ACM Press, New York (1998)
81. Shipman, F., Price, M.N., Marshall, C.C., Golovchinsky, G.: Identifying useful passages in documents based on annotation patterns. In: Koch, T., Sølvberg, I.T. (eds.) Proceedings of the 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003), pp. 101–112. Lecture Notes in Computer Science (LNCS), vol. 2769. Springer, Heidelberg (2003)
82. Spooner, A.: A Dictionary of Synonyms and Antonyms. Oxford University Press, New York (1999)
83. Stein, L.D., Eddy, S., Dowell, R.: Distributed sequence annotation system (DAS), version 1.53. <http://www.biodas.org/documents/spec.html> [last visited 2006, October 2], March 2002
84. Thiel, U., Brocks, H., Dirsch-Weigand, A., Everts, A., Frommholz, I., Stein, A.: Queries in context: access to digitized historic documents in a collaboratory for the humanities. In: Hemmje, M., Niederée, C., Risse, T. (eds.) From Integrated Publication and Information Systems to Virtual Information and Knowledge Environments. Essays Dedicated to Erich J. Neuhold on the Occasion of His 65th Birthday, pp. 117–127. Lecture Notes in Computer Science (LNCS), vol. 3379. Springer, Heidelberg (2005)
85. W3C.: EMMA: Extensible MultiModal Annotation markup language—W3C Working Draft 16 September 2005. <http://www.w3.org/TR/emma/> [last visited 2006, October 2], September 2005
86. Xi, W., Zhang, B., Chen, Z., Lu, Y., Yan, S., Ma, W.-Y., Fox, E.A.: Link fusion: a unified link analysis framework for multi-type interrelated data objects. In: Feldman, S., Uretsky, M., Najork, M., Wills, C. (eds.) Proceedings of the 13th International Conference on World Wide Web (WWW 2004), pp. 319–327. ACM Press, New York (2004)