Passively acquiring information must end

Hermann Maurer, hmaurer@iicm.edu
Bilal Zaka, bilal.zak@gmail.com
Sonja Eisenberger, sonja.eisenberger@tugraz.at

Abstract:
When we read a printed paper, document or book we sometimes come to points that we do not understand or where we feel important information is missing, or plain wrong, as we can show by examining a particularly reliable collection of data. Yet we can just shrug our shoulders, since contacting the source of the information is at least very cumbersome.

This has changed surprisingly little even if the information does not come in a printed version, but as a digitized book, a Web-page or a PDF File. Usually, there is no easy way to point out mistakes or missing information or asking for more extensive explanations. It seems that the possibilities provided by digitized information, particularly if it is on the Internet, are not used in most environments.

In this paper we argue that it is strange that the power of networks is not used to allow all kinds of interaction, involving authors, readers and the information at issue, in collections of digital documents, while communication is often considered to be over boarding in some social networks. We discuss one concrete solution for alleviating this problem in collections of digital documents.

Note: This paper is based on [10] and particularly the paper [18], but puts more emphasis on the real reason for NID (as already hinted at in the abstract), and on usability and application aspects. Note that basic ideas of NID were already addressed in [19].

1.Introduction

When we read some information that we find on the Internet or that was sent to us electronically, it is often very annoying that we cannot suggest additions, corrections of the material, or ask questions concerning some specific issue.

Notice that this lack of easy communication is very strange, considering social networks and discussion forums, where many even far out contributions are found.

In contrast, consider some of the following examples.

You read in a substantial news-paper like The Guardian [1]. Maybe you want to ask a question or contribute some further information, but you will have a hard time finding someone to contact. Even if you are successful, after a longish search for an e-Mail, chances that your message will achieve anything are slim. But even if you get an answer or explanation no other readers who might also profit from your remarks and the answers will see it.

When you read a digitized book available form e.g. one of the many sites offering free books to read, a few suggested in "12 sites for free books" [3] or in the very prestigious "Project

Gutenberg" [4] you may wonder why it even would make sense to add something in a classic book. But this just shows lack of imagination, as we will try to show in our exposition in Section 2.

If you follow Steingarts excellent daily newsletter [5] you do have the possibility to send a comment on his long reports, yet you have to identify yourself, you have to specify the section you want to address, and since unfortunately the contributions are not archived, any additional material you think should be added will never be seen.

The situation is a bit better if you read a Wikipedia [2] article: At least there is a contact you send a message to, and it may even result in a change of the Wiki-entry. Still, we believe this is neither enough, reactions if any are slow, nor is it always desirable to identify yourself.

In Austria-Forum [6] this is not necessary. If you look at one of the ten-thousands of pictures in the geography section, like the one shown in Fig. 1 below, you always get a green button labelled "Feedback" on the right side.

Pula - Church of St. Francis; Cloister

Außen romanisch, innen gotisch und mit beeindruckendem Kreuzgang ausgestattet, bildet die Franziskanerkirche einen Ort der Ruhe inmitten des hektischen Touristengewimmels in Pulas Altstadt.

Outside romanesque, inside gothic and endowed with an impressive cloister, the Franciscan church is a place of quiet inmid the hectic tourist bustle in the old town of Pula.

Pula - Church of St. Francis; Cloister, May 2016, © Gerhard Huber, under (cc) BY-NC +Edu
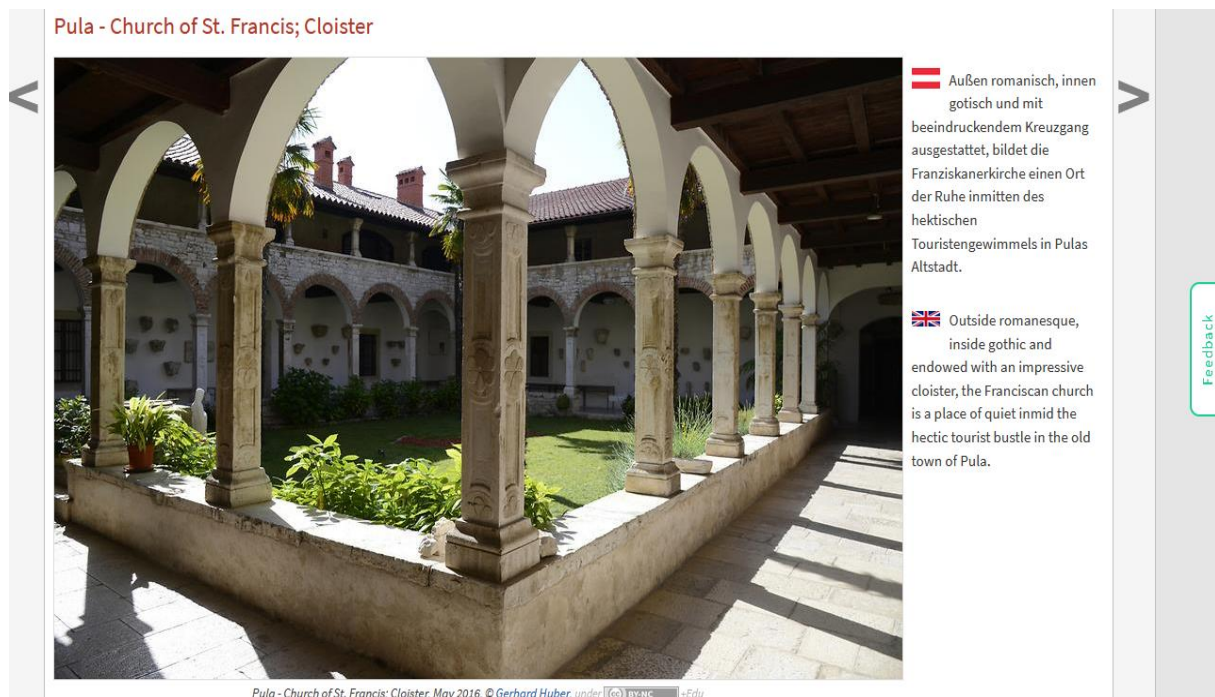
Fig.1. Cloister in St. Francis church in Pula, Croatia. Notice the green Feedback button on the right.

A click at it lets a window pop up where anyone without a need for identification can leave a message, as you see in Fig.2. Of course, if the sender wants an answer, an e-Mail address has to be specified. It is important to realize that such a Feedback does not have to go to the administrator of the site at issue (who is responsible for some 1,4 million media objects), but will go to the person responsible for the page the Feedback is made on, hence preventing any person or group to be overwhelmed with too many feedback messages.
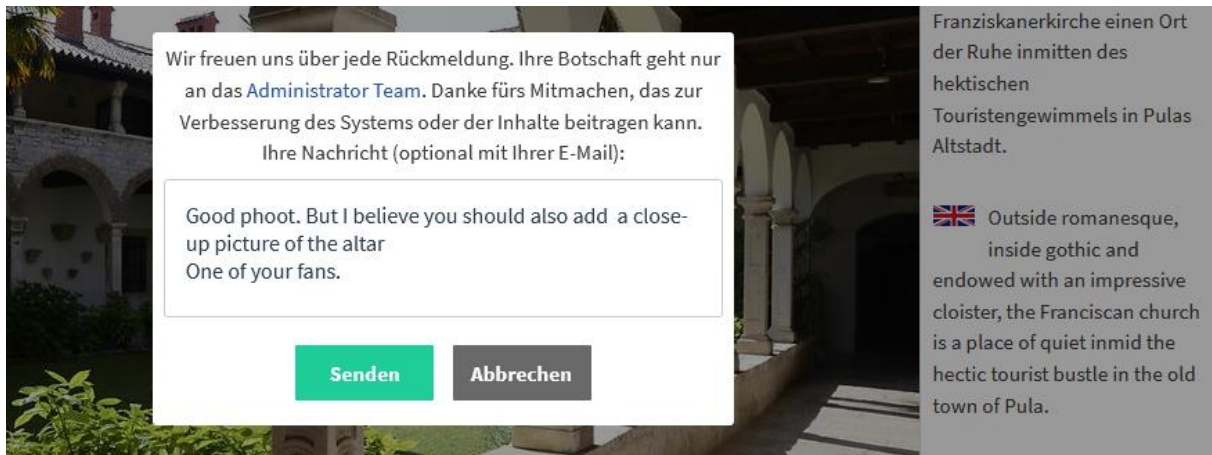
Fig.2. Anonymous Feedback-form

Above examples can be arbitrarily extended. If you go to the Website of any major company and you have a specific question you will be hard pressed to get an answer. The FAQ list, more often than not, does not contain your specific question. A contact e-Mail address is usually missing. You may be able to find a phone number: calling it, you often will end up in long waiting loop, and then the person answering is probably not at specialist in the area you are interested in.

How to use and avoid misuse of anonymous feedback has been discussed in the paper [7].

Thus, it should be clear that a certain amount of communication between users and providers of information makes sense. We will presenting a concrete implementation in Section 2, and will provide what we hope is a stunning array of applications in Section 3, applications that will be helpful in many situations, yet have been overlooked in the past.

2. Networked Interactive Digital documents - NID

Documents of any size (from a few pages to large books) can be easily converted into NID documents if they exist as Word- or PDF-Files or PPTs.

(Before we go deeper: If you want to try out examples, look at e.g. one of the books on [9], maybe the short 14-page booklet [8] is a good starting point. For still more information, look, as pointed out at [18].)

A collection of NID documents allows to find a particular document or documents by specifying author, title, keywords ("metadata") or a category or subcategory as defined by the administrator of the system. This is shown in Fig. 3 using as example part of a screenshot of the NID collection of Graz University of Technology mentioned in [10]. Before working with NID documents one can choose the interface language: Here, English was chosen. One can switch easily to other languages. If not yet available it can be added easily.

Within a document, relevant parts can again be found using keywords, a table of contents (that opens at a click on any page of the document) or full-text search. Particularly convenient is the fact that full-text search can be applied not just to one document but to a

number of them, avoiding the necessity to e.g. repeatedly search in more than one volume of a series of books or more than one issue of a journal, or more than one brochure.
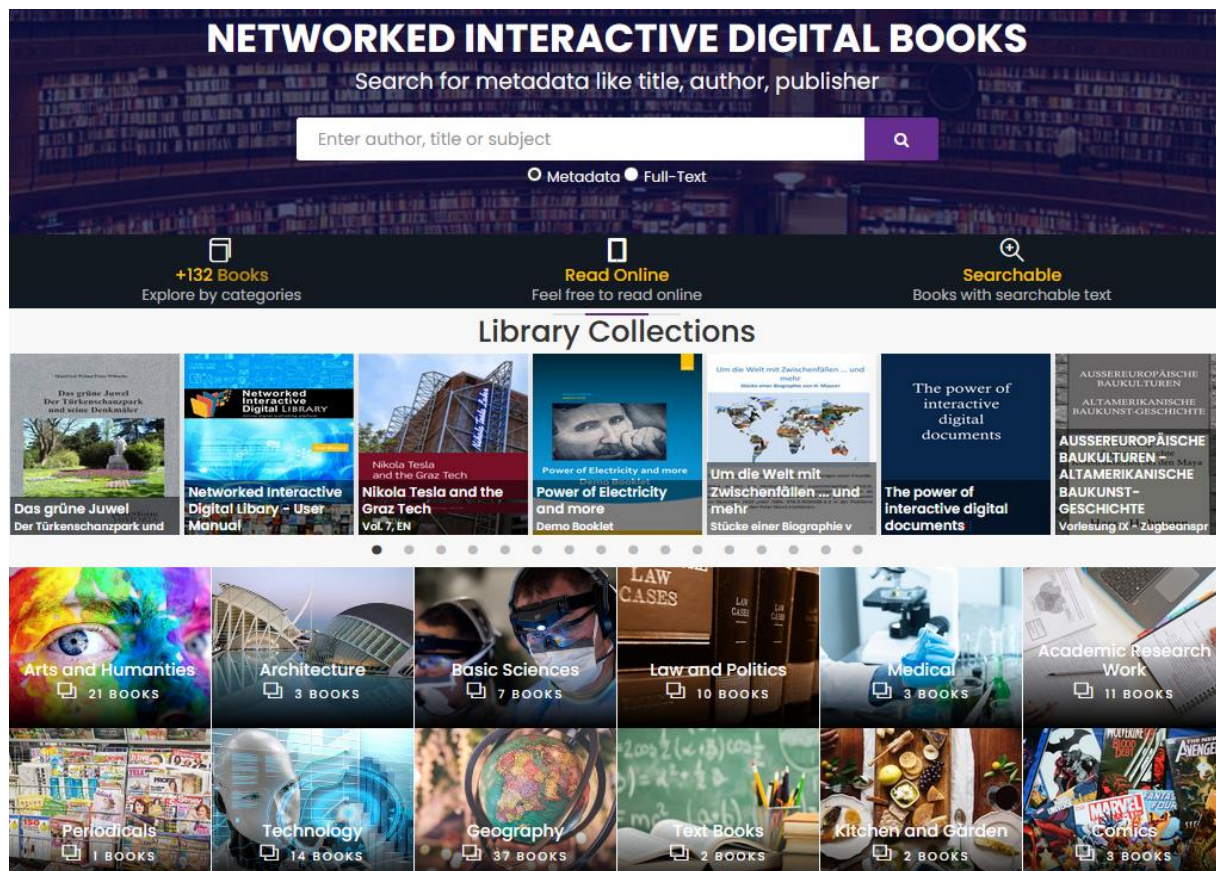

Fig.3.: Entry page of a typical NID collection

On can use NID document in "Simple Mode" or "Full Mode". In Simple Mode the menu options are **Feedback, Search** (full text), **Annotation, Discussions, Settings, Login/Logout** and **Full Menu.** We will now discuss those options and then discuss the Full Mode, obtained by clicking at Full Menu. Note: When a book is uploaded it can be decide that the book starts in Simple Mode or in Full Mode.

Feedback needs no further explanation: See the discussion for Fig.2. Search has been mentioned before. A search for "power" in the short booklet [8] on Tesla as genius of early electricity finds 3 relevant pages, see Fig.4 a, and clicking at the first option, yields Fig. 4b.
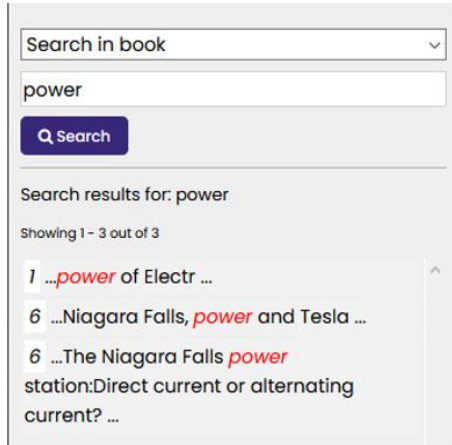
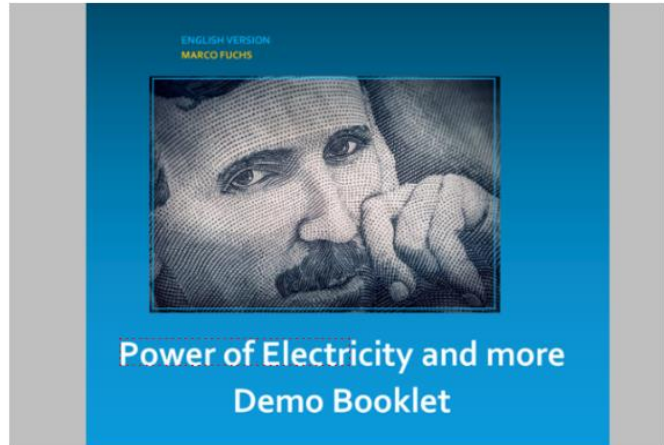Fig. 4a   Fig.4 b

The page whose part is shown in Fig. 4b is interesting for another reason. The face of the person (Tesla at the age he worked in the United States) has a light-blue rectangle around it, indicating that here is an "annotation". Moving the mouse into that rectangle we obtain a page part of which is shown in Fig. 5.
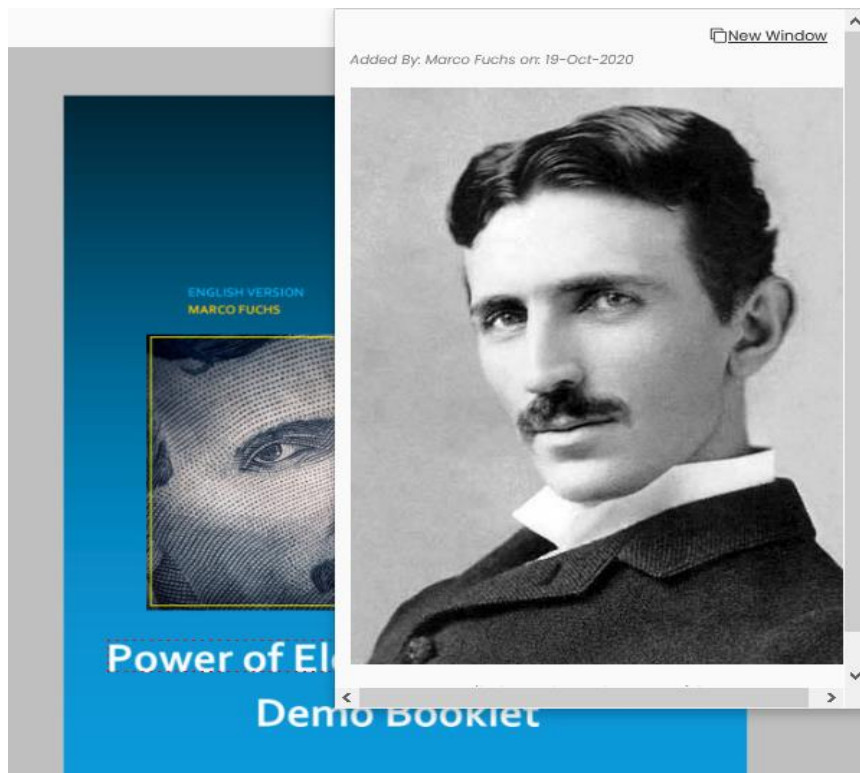

Fig.5

Fig. 5 shows Tesla at age 34, as can be seen using the scroll bar or opening "New Window". This Picture was, as itr says at the top, added by Marco Fuchs on 18-Oct-2020. This means that Marco Fuchs was logged-in, otherwise it would say "Added By: Anonymous on 19-Oct-2020".

When choosing the menu-item "annotation" one can draw (after clicking at the pencil icon) the outline of a light blue rectangle wherever one wants on the screen and then use the form shown in Fig. 6 to add an arbitrary piece of text that pops up in an additional window,

or a picture (as in Fig.5) with or without additional text, the first frame of a video, or some audio-clip, or a link to an arbitrary URL, or a combinations thereof. Unless specified differently, when uploading a document, no registration is needed for annotation or contributions in discussions (see below), i.e. user can stay anonymous. However, whoever is responsible for a document may have specified that registration is necessary (for annotations and dontributing to discussions), so whoever made an addition or change can be contacted if necessary: The persons E-Mail is only "know to the system", but this can be used to contact the person.

In this way a document can be enriched by many further additions, as was indeed done in [8]. (Give it a try!). Of course, if someone makes an annotation the administrator and the person responsible for the document receive a note and can delete the entry; more politely, they can turn it into "Private" so only the person who made it can still see it, or leave it as public, or make it visible to just specific groups (as the Annotator may already have done. A convenient group-hierarchy management system makes this possible). By the way, this is one good reason why one should log-in before making an annotation, since if it is considered unsuitable, it will have to be removed, as it cannot be turned into "Private".)
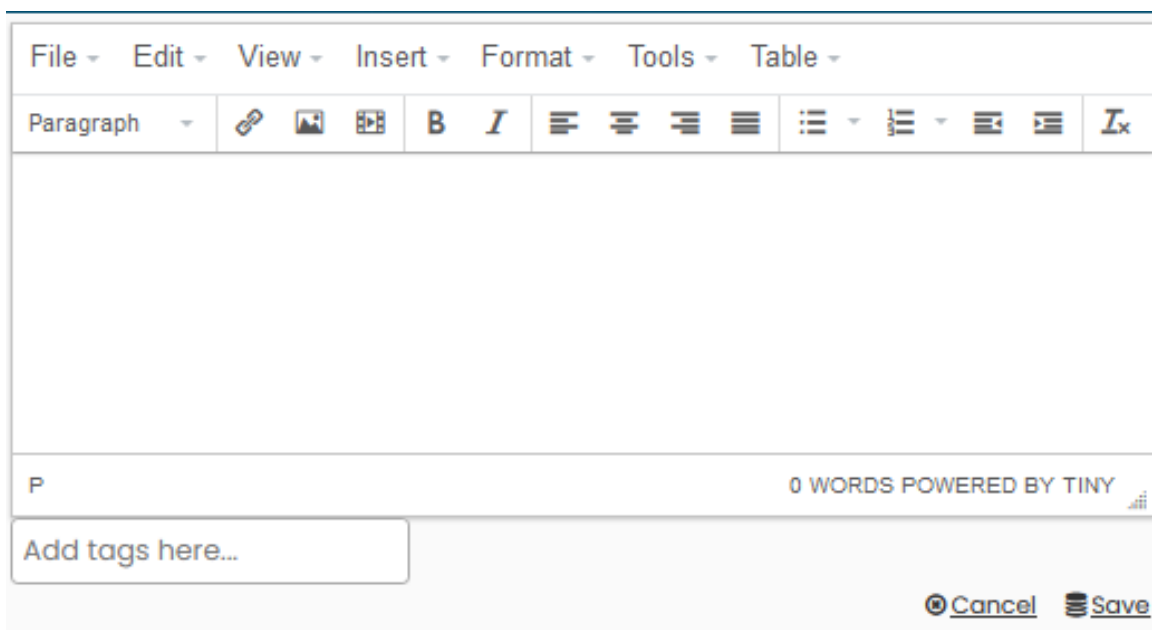

Fig.6: Annotation Form

Annotations are very powerful tools. They allow adding material to a document for all or just for some group(s), linking to related material, alerting others of similar or controversial information and can also work as extensions of discussions.

In passing, please note that also "Fancy annotation" can be activated. Since this is a standard feature of of the Full Mode, we will discuss it when we come to Full Mode.

On the top of every page of a NID document one can see "Discussions". If this word is lit up in blue, then a discussion exists on that page and one can look at it and add a comment by just scrolling to the end, and posting some material using the empty window. Look at the discussion on page 6 of [8]. Independent of whether the word "Discussions" is lit or not, you

find a list of all discussion in this document, and you can start a new one: If so, the first few words should be descriptive of the discussion to ease finding it. As mentioned for anntoations it is possible to demand registration for adding to discussions when uploading a book if one wants to avoid controversial statements and is not willing to checht all changes reported around the clock.

The function of the Login menu item is clear (the first time you sign in you have to register as usual in most systems). Signing up is e.g. necessary if you want to add a quiz on some page, or if you want to add annotations only for a certain group or just for yourself, or if the person responsible for a document has made a log-in necessary for making annotations or adding discussions. Also, contributions showing who made them may increase their credibiltity, and makes it possible to send information about them to a certain group.

The menu-item "settings" allows you to switch to full-screen, or choose other viewing modes, but the standard mode is the "image view", so usually you should rerun to it, sicne only in iamge view all menu options are avialble, anntoations and discussions are visible, etc.

Very important is the choice "Full Menu", that gets you into Full Mode (You can switch back any time to Simple Mode by clicking at the now appearing item "Short Menu"). As you choose Full Menu, you suddenly get further menu options:

Table of contents, Console, Fancy annotation, **Change layout, Transclusion, Image Manipulation, Book Details, Metadata, Citation, A.I.** Note also that the list of options unter Settings is a bit longer and the choice Full Menu has been replaced by Short Menu (to switch back if desired), much as Login is replaced by Logout aftera log-in, and Logout is replaced by Login if the Logout option is chosen.

The Table of Content (ToC) option shows the table of contents on any page by clicking, if such a ToC was created any time after uploading the document. If none was created, ToC just shows a pointer to the first and last page.

Console requires to be logged-in. It provides a number of administrative tools, like defining user groups (important if e.g. some annotations should only be available for certain users) or to define quizzes, and more. NID offers two fundamentally different types of quizzes: test quizzes and information quizzes.

A test quiz consists of one or more questions, usually with one correct answer for each question. The (in this case) registered user chooses the answer that he thinks is correct. At a later point in time, NID enables persons with special rights to determine for each person that has taken part in the quiz how many questions were answered correctly and how many incorrectly. Thus, this is then primarily used to evaluate the performance of the user (e.g. as part of an examination or instead of an examination), but can also lead to adjustments of the quiz in some cases.

An information quiz also consists of one or more questions, each with one or more correct answers. The aim is not to test users, but to inform them: If they give a wrong answer, it is explained why it is wrong, what the correct answers are, possibly with a link to information that explains the situation in more detail. In this way, users can e.g. realize at the end of a

section, whether they should continue, or not: if a lot was misunderstood it makes sense to repeat certain passages or to receive additional information before continuing in the document.

Fancy annotations allow to point to annotations not just by the thin outline of a light blue rectangle, but by using other colors, shapes, different types of lines or special icons indicating that users will be presented with e.g. a video, an audio clip, or just textual information, etc.

Change layout allows to place more than one page on the screen at the same time, and work with them independently. This allows to e.g. compare pictures, or look at a page in English next to the page from a German version, etc. Pages can come from any server providing information following the International Image Interoperability Framework (IIIF) standard [11]. Many large libraries use the standard, and examples of what can be done are e.g. found in [12].
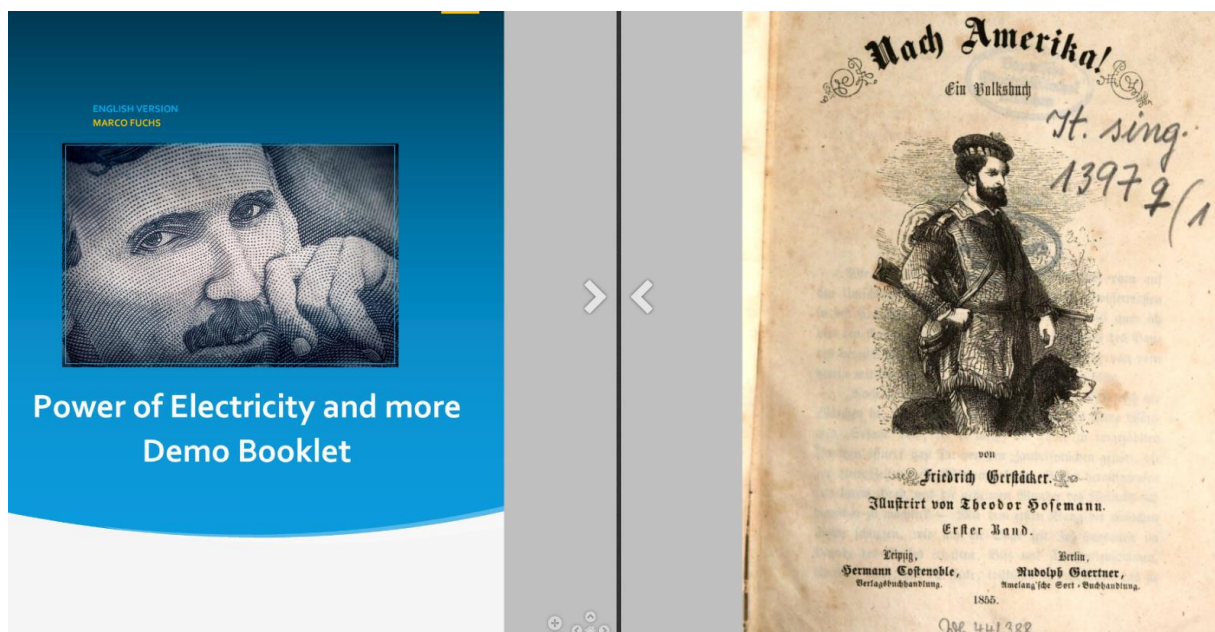

Fig.7: Two pages of two different NID libraries (Graz and Munich)

Fig.7 shows Nikola Tesla on the first page of book [8] who went to America to continue to work there, and next to it a page of the book "Nach Amerika!" (i.e. "Let's go to America!) from the State-Library Bavaria [20]: It is now possible to turn pages in both books, or change them in any other way NID allows, the two sides completely independent of each other.

Transclusion allows to define a rectangular area of any screen and creates a link to just that area of the screen, e.g. to incorporate it into a document or an E-Mail. Thus, NID allows the creation of links to part of a page (!).

Image Manipulation allows to change the appearance of a screen, by e.g. dimming, switching to black-white, and much more. Try it out!

Book Details provides what it says, general info on the book (document) at hand.

Metadata shows terms associated with the current document.

Citation is nice if one wants to cite a page of a document according to one of the many versions of citations in the language of choice.

A.I. is a demo of what is in the works: Under **"Topics"** you can find all terms considered of interest extracted from the current page of the document, if the document went through appropriate pre-processing when uploading it. This will be used in the future to intelligently link segments of one book to segments treating similar topics in the same or different books. The main challenge in implementing semi-automatic linking of NID books with other information (potentially also outside the NID books) and vice versa is that this requires an expansion of the full-text search (from word-oriented to phrase-oriented), including taking care of synonyms, homonyms, stemming etc. To be successful, methods of AI and NPL [16] must be used to recognize "important" words or phrases. Even aspects of sentiment analysis could also be considered. I.e., for example, you might not be looking for "electric cars" but "arguments against electric cars" or "arguments for electric cars". Under **"Object detection"**
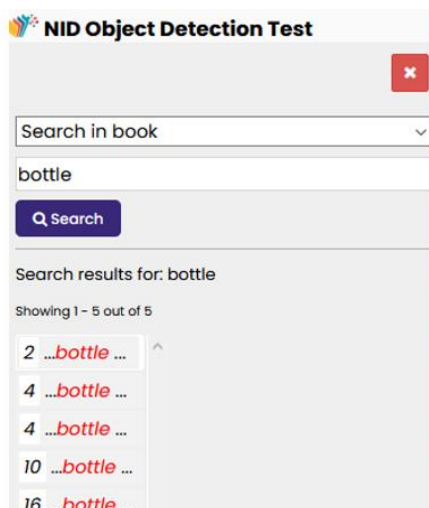

Fig. 8a   Fig. 8b

you can locate all pages containing a picture in one of now 80 categories, again if the book has been preprocessed appropriately. NID uses state-of-the-art computer vision methods to do so. This, being in an experimental stage was only done for very few publicly accessible books, the best example is [13]. When entering the book, you find in a green frame all object categories you can search for. Searching for e.g. "bottle" gives five results as shown in Fig. 8a and choosing 2 shows part of   page 2 as seen in Fig. 8b, in which the bottle is highlighted with a dotted rectangular outline. Everyone is familiar with the search for similar images, for example using Google Images [17]. However, object detection here is based not on similarity of pictures, but on a desired "category".

3. Applications of collections of NID documents

In this section we describe some of the many applications of collections of NID collections that we have come across. We are sure the list is not exhaustive, yet shows that with a NID-like approach a new way of using digital documents is emerging.

3.1 NID for digital libraries

To offer a collection of documents like books using NID has the advantage that the library can expect some feedback from users, some users may have questions or remarks or will start discussions, or enrich the material by adding textual or multimedia information or linking an entry in a book to another relevant entry in the same or another book, or to some information on the Internet. Note in passing that research is continuing how to create such links with little effort using advanced A.I. techniques.

Libraries who are already offering digitized versions of some of what they hold will clearly be reluctant to go to the trouble of converting all material to NID format. Also, they may be offering e.g. E-Books under some special license from publishers, or do not have the material as such on their server, but users access - via the library - a database run by a publisher with which the library has some agreement.

Hence such libraries are likely to make only a limited number of documents available via NID, often because of licensing or copyright issues or because the expected benefit will not be sufficient to justify the effort necessary.

The situation will be different depending on whether one considers a general public library, a library of a university or research institute, or a library just useful for a certain organization, or some other library (= collection of digitized material).

One main reason why NID is sometimes justified is the fact that groups of users can be defined for whom sharing collective knowledge by feedback, annotations and discussions is very much desirable.

In case of a public library, there might be a reading club that at any time is just interested in examining a few books that all members of the club should read and discuss with other members. In such a case the library may make those books available on NID only for the group involved: Since NID books can be stored so that they can be accessed only by a certain number of users the library will be able to negotiate reasonable conditions with publishers, if necessary.

If we are talking about a university library, a typical group might be a class where members need access to some material just for one term. In such institutions some of the material might also originate as lecture notes of professors of the university, of books authored by staff members, or of books available under some free (typically Creative Common) license. In this case licensing issues are usually easily resolved. We are aware of instances where a professor has made a set of notes and a few supporting books available for a graduate course: students of the course are not just supposed to extend the material by stuff found or worked out by them, but the grade obtained for participating in the course depends (!) on how much and what each student has contributed. NID offers the possibility to find for a student and a set of documents all material that has been added by that student to this set, thus making it easy for the lecturer to judge the value of what the student has done.

Similar ideas apply to organizations of all kind that want to encourage a group to familiarize themselves with certain material by interacting with it and members of the group.

In other cases, the reason to use NID may be topic oriented. One could e.g. imagine a collection of documents and books on the topic of electric cars. Persons with very diverse background are allowed or asked to participate in studying the material, adding their own ideas and viewpoints as annotations or in discussions.

Another important application for some organization or companies of a limited set o NID documents is the dissemination of PR or even sales information. Consider two typical examples:

(a) Often research organizations or companies provide a monthly, quarterly or yearly report of particularly important developments. Such reports are typically sent to some users in "glossy" printed form, and to a larger group as e.g. PDF File. Unfortunately, feedback is very rare: To write a mail, one has to refer to the part one wants to criticize, ask a question about, add further information, often quite laborious. Also, and in some cases, anonymity is preferred. Well, NID allows anonymity, and when an addition is made, the person responsible for the document is informed automatically of the fact and the spot of change and change it, if deemed necessary or advantageous.

(b) This also applies if the report is really a sales instrument, showing some new product(s). Again, feedback of any kind is very much desired, but rarely obtained.

There is one general observation that must also be considered. The number of people willing to read detailed textual material is decreasing all the time. NID, by allowing to make material more attractive by adding to the text multimedia material to it at any later time, or even links to some interactive environment, makes reading potentially more interesting, hence NID documents enriched in this fashion will be studied more fully by more users.

3.2 NID for organizational knowledge management, knowledge sharing and knowledge growth

One of the possibilities that makes NID particularly interesting for such applications is the fact that questions or discussions can be asked in a document, such as in manuals, that only a certain group (e.g. all members of an organizational unit) can see and use and work with. Within such a group, one can also set up sub-groups with special rights.

It is particularly interesting that an NID document can complement communication media such as WebEx [14] or Zoom [15] very much: These communication platforms--- very popular in times of COVID--- have little temporal stability or usability. One can save a long Zoom Meeting as such, but it is hardly possible to find something there later, or ask a question, etc.

After such a meeting, it makes sense to put short minutes of the meeting available as a document in the NID system and thus offer the opportunity for further discussions, to ask questions, to provide additional sources, and more.

An organization-internal collection of important documents in NID with corresponding comments on the one hand preserves the knowledge beyond the time of a person's presence in the organization, and on the other hand is particularly valuable for newcomers,

because those who do not understand information can get help from others using the question and discussion facility.

3.3 E-learning features

Face-to-face teaching or teaching using communication tools such as Zoom etc. is meaningfully supported by NID documents that allow questions and discussions.

NID is also well suited for graduate courses (seminar-like work) as mentioned in 3.1, where the active participation of the user is important. A lecturer offers students material in the form of NID books, which the users are supposed to supplement through own work and research. This process can take several months. At the end, the lecturer can determine for each participant what new contributions this user has made: It is precisely the value of those contributions that is used (or is also used) for the assessment of the student.

4. Summary

NID is a system that can be used for many tasks and has many possible applications by now. Nevertheless, it is constantly being further developed with innovative features and for new applications, some of which are available for testing purposes.

The important point to understand is that NID allows the interaction of users with each other, the person(s) who have made the digital information available ("editors"), and that the digital information can be extended over time by annotations and discussion, all in a very controlled way. Just to indicate how far this can go, a teacher might ask pupils/students to add information to some documents and provide some bonus for good additions. At the end of the year the teacher might erase all annotations and discussions, so the next class starts again with a "clean slate". The opposite may also make sense: To have a document/ book that keeps growing every year!

This paper has not discussed all features available for communication purposes. Editors can e.g. send messages to a certain group of users (or to all registered users); users can define themselves arbitrary group to communicate with easily, etc.

Although NID can deployed and is deployed now with success in many situations, there are many innovations that are currently on the internal "to do"- list, from trivial extensions to a fuller set of quizzes, more extensive administrative possibilities, providing better integration of full-screen discussions and videos, special handling of certain annotations (like annotations that pop up automatically when turning a page), and much more. We hope that this exposition helps us to increase the number of NID users so that we can welcome them in the planned NID user-group!

5. References

[1] The Guardian (https://www.theguardian.com)
[2] Wikipedia on COVID (https://en.wikipedia.org/wiki/COVID-19)

[3] 12 sites with books to read free (https://ebookfriendly.com/sites-where-you-can-read-books-online/)

[4] Project Gutenberg (https://www.gutenberg.org/)

[5] Steingart (https://www.gaborsteingart.com/steingarts-morning-briefing/)

[6] Austria-Forum (https://austria-forum.org)

[7] A Note concerning Feedback and Queries for Web pages; N. Delilovic, H. Maurer, JUCS vol. 25, no. 7, 733-739 (2020) (https://austria-forum.org/attach/User/Maurer%20Hermann/Publikationen_2019_und_2020/19-07-30-Feedback-JUCS.pdf)

[8] Demo Book (https://nid.iicm.tugraz.at/Home/BookDetail/199)

[9] A collection of some NID books (https://nid.iicm.tugraz.at/)

[10] Information on NID on (https://austria-forum.org/af/AEIOU/NID-Books)

[11] IIIF-Consortium Homepage (https://iiif.io/)

[12] IIF Example (https://iiif.io/apps-demos/#awesome-iiif)

[13] NID Object Detection (https://nid.iicm.tugraz.at/Home/BookDetail/305)

[14] WebEx (https://www.webex.com/)

[15] Zoom (https://zoom.us/)

[16] NPL (https://en.wikipedia.org/wiki/Natural_language_processing).

[17] Google Images (https://images.google.com/)

[18] The power of interactive digital documents (https://nid.iicm.tugraz.at/Home/BookDetail/344)

[19] Investigating the interaction activities in digital libraries: The Networked Interactive Digital Books Project. (H. Maurer, B. Zaka, N. Delilovic), IPSI Transactions on Internet Research, vol. 16, no. 1 (Jan. 2020), 75- 82 (https://austria-forum.org/web-books/tir1601en2020isds/000070) , (https://nid.iicm.tugraz.at/Home/BookDetail/204)

[20] IIIF Bibliothek in München (https://app.digitale-sammlungen.de/bookshelf)