Int J Digit Libr (2012) 12:159–178
DOI 10.1007/s00799-012-0078-z

Video digital libraries: contributive and decentralised

James Lanagan · Alan F. Smeaton

Published online: 24 January 2012

Abstract Technology usage is changing rapidly and is becoming a more mobile, more social and more multimedia-based experience. This is especially true in the area of content creation where mobile social applications used by crowds of people are challenging traditional ways of creating and distributing content, especially for applications like news dissemination. Libraries have traditionally functioned as repositories where the information content of a society is analysed, curated, organised and stored, acting as a permanent record of what is to be remembered from a society. How can this function be achieved by present-day libraries attempting to cope with mobile, social, multimedia content who's nature and utility of which change the type of information we wish to curate and store? This information is both dynamic and organic, posing challenges to the more fixed models of information in digital libraries. In this article we describe two digital library systems that archive video content from the sports domain, and which support user annotations and merging of diverse information sources in an integrated way. We report on analysis of the deployment of these two systems and highlight how they extend the traditional role of a (digital) library.

Keywords Digital multimedia libraries · Annotation · User-generated content · Trust · Data quality

1 Introduction

Libraries have existed for millennia, providing civilisation with storehouses of information and reference. The information contained within each library may be sifted through, catalogued, and used by any person present. Libraries curate the most important content from within a society and provide support for cataloguing so that its storehouse may be used by any person. Their curated nature means that the information contained within them has been selected and approved, and characteristics such as its provenance and accuracy are guaranteed.

A Digital Library (DL) can been defined in many ways, but the over-arching theme appears to be that of 'a collection of electronic materials . . . created and managed for (and sometimes by) one or more user communities, and technical and user services are provided that add value to the materials' [65]. Two parts of this definition strike us as fundamentally and profoundly different from the basic understanding of what a physical or 'traditional' library is.

The creation and management of information by a community is at odds with the idea of a traditional library. Information may be borrowed, researched, copied, or cited from within the media of a library, but this information is permanent and unchanging. Levy [45] talks of the long-lived nature of information within the physical library, however, this is not necessarily the case for information in a DL. Information of an ephemeral nature may be created in this new context—created to service a current purpose but of no long-standing use. Whilst physical books are limited in their ability to change and adapt to the current usage context, digital sources can evolve and change at will. Indeed this is one of the key architectural requirements of all DLS.

The role of a librarian in the modern-day library is still that of custodian and archivist and is heavily reliant on technologies that have already been incorporated into the library space: database search, catalogue creation, and indexing are all supporting activities. All these place the librarian at the focal-point of content management and creation, but

J. Lanagan (✉) · A. F. Smeaton
CLARITY: Centre for Sensor Web Technologies,
School of Computing, Dublin City University, Dublin 9,
Ireland
e-mail: jlanagan@computing.dcu.ie

this should not necessarily be the case. With the move to digitalise resources and access there is a chance to further democratise and share the burden of content management and organisation with the community of users as a whole.

This content is in many different media and is created for a variety of reasons, many of them to do with social interaction. Assembling, curating, describing and validating the authenticity of content is now taken out of the hands of the professional, and the role is given to everybody. This has obvious advantages in that the volume and scale can grow hugely, but also means that many of the characteristics of traditional libraries do not exist within the digital realm.

Curation no longer stands as a filter or hurdle for inclusion within the library. The democratisation of creating, publishing, and sharing of information content through resources such as YouTube, Flickr, Facebook and Twitter now stretches the traditional role that DLs play in our world [56]. This is evidenced by recent information that shows an incredible 35 h of new video content uploaded to YouTube every second.[1] This is just one of a slew of repositories available on the web today through which users may catalogue, create and share information but whether these can be called DLs or not is questionable.

When Library 2.0 was first introduced by Casey [19], concerns were raised as to the actual meaning of the phrase [57]. Since then, however, and thanks in part to the widespread adoption of social media and computing, a clearer understanding has emerged. Library 2.0 does not refer to a completely new way of using and creating libraries, but instead is more about the augmentation and improvement of library usage through the adoption of digital technologies. Habib later provided a more concrete definition of the ideas of Library 2.0 by calling them a 'subset of library services designed to meet user needs through adoption and in acknowledgement of the basic tenants of the evolving Web 2.0 paradigm' [32].

In defining modern DLs, an important early step was the questioning of assumptions made about their nature. Early models of multimedia DLs took on the characteristics and feel of the physical libraries that came before them [21,75] replicating the role of archiving information that would remain constant and immune to change. While the digitisation and archiving of perishable commodities such as paper books and plastic video and CDs, are obviously important, it is noted in [45] that, in simply moving from digital to physical domains, we miss an opportunity to create something far better. Although as we have said DLs build in many ways on the physical equivalents, there are a number of ways in which they can exceed their physical counterparts.

[1] http://youtube-global.blogspot.com/2010/11/great-scott-over-35-hours-of-video.html.

We look at these ways now. They includes the following specific points:

Versioning
The transient and temporal nature of digital artefacts means that, for thorough archiving, there is the possibility for the storage of different article life-stages. It may not be enough to store just the original article, but there should also be provision for the storage of edits to it so as to preserve its life-cycle and allow its development to be re-played.

Annotation
The creation of metadata and annotations around the content in a DL can provide highly preferable new interaction paradigms for library content. Annotation, as discussed in Sect. 3, allows for the constant growth of information within the library, even when no artefacts are added. Through the use of annotation, users continue to create new information based on (or providing further understanding of) content already within the DL.

Collaboration
While the traditional use of libraries has been based on solitary discovery and use of information, the move to digital media creates an opportunity to aggregate and use every person's activities to the advantage of the library community as a whole. Marshall notes [52] that public annotation is created with an audience in mind, meaning that the creators themselves feel that the annotation is of use to their peers. Couple this with some of the more explicit forms of interaction, such as shared search, and we see the enormous opportunities available for collaboration.

Media Integration
Within the digital realm, it is possible to combine audio, visual, and written instances of an artefact into a single representation. An example of this might be a digital artefact that describes in text the life of a great composer. Within this written document, it is possible to embed audio, a visual tour of the country where s/he was brought up, etc. This type of integration is possible since the single representation of the instance may be re-arranged and collected within the architecture of the DL without any physical change.

Leveraging the power of user feedback can be useful for libraries in future when attempting to decide what books, videos or media they should continue to store. Concepts such as The Long Tail [7] require this feedback, but are also grounded in the concepts of infinite change and space. These are obviously not available within the physical world where space is at a premium within library premises. Digital Libraries need not suffer from this limitation, however, and are therefore able to offer more diverse catalogues to their communities. The philosophy of 'Library as a service' works around the idea of removing the innumerable opaque information silos

and creating a network. In doing so, the small groups of specialist users within each library community can find other groups of like-minded individuals.

Miller [55] describes several services that harness the 'library as platform', building on the services provided by the online catalog and other pre-existing library services [77]. By opening up the collections within the libraries themselves a more fluid exchange of information can take place between geographically and physically disparate locations, allowing the digital user access to far more information. Even this is a step forward allowing for additional statistics to be gathered without necessarily allowing for the whole sail personalised creation and re-organisation information that is available on the web as a whole.

It is these principles of shared content creation and access, coupled with the aforementioned characteristics of versioning, annotation, collaboration and media integration, which underpin our study. In this article, we shall introduce working exemplars of the share-everything principle and discuss how the ability to share content can lead to the creation of content which may not have been possible otherwise.

In the next section, we discuss how content creation has become a decentralised activity. We then extend this to discuss annotation as a form of decentralised content, and the various ways in which it can be created. We then summarise the most significant past study in creating digital video libraries, with a focus on the sports domain. Following this, we outline the two systems which are the focus of this article, (SPORTSANNO) and (ANNOBY) and outline some of the ways in which these systems differ from current web repositories or 'libraries' such as YouTube. The issues that are raised by such contributive and decentralised DLs including data quality, trust, and the value of authorship are then discussed.

2 Decentralising content creation

The driving force behind Library 2.0 and Social Computing is the consolidation and utilisation of a user community's interactions with media. In the DL context, this can mean utilising the access patterns of users to better organise and prioritise the filtering of information. Again the benefits of versioning and architecture isolation as seen in the ability to personalise the library experience to any user. The provision for services that allow users to communicate and gather information without the need for direct input by the librarian again go some way towards democratising DL artefact creation. These services can also be of benefit to the library itself in providing qualitative feedback into what users find attractive or intrusive within their current implementation.

These ideas are echoed by Marchionini in what he refers to as a 'sharium', a concept that makes important distinctions between itself and that of the traditional library model. 'A sharium goes beyond providing information in a curated collection, to inviting active participation in the form of collaboration and contributions from all users and to providing flexible means for reusing information resources' [46]. In extending the traditional model, a sharium looks for many of the same qualities that are provided by the Library 2.0 paradigm to facilitate the personal and collective augmentation of intelligence and content. The requirement for tools that may be used to both retrieve and use the multimedia content within a DL in the context of the current users' needs is a versatility that is possible through the lightweight nature of Library 2.0 applications. Shariums also recognise the clear value of allowing the users of a DL to contribute their own ideas, time, and expertise to the general information framework. They do not, however, make any specific mention of explicit content creation by users.

The requirements for a sharium to exist have not yet been fully met, but [46] gives many examples of platforms that are progressing in this direction. Platforms such as SourceForge (http://www.sourceforge.org) that allow users to upload code projects that may then be downloaded and used by any other web users. The system is not restricted to use by those registered, but is open to all. Another example that is closer to the concepts of both Library 2.0 and the systems we shall outline in Sect. 4 is Slashdot (http://www.slashdot.com) where an initial user uploads a link and abstract for an article posted elsewhere on the web. Others may then comment on the story, creating a thread of new information and discussion centred on the article in question. Many more examples exist, including the Internet Archive[2] that is a DL of Internet sites and other cultural artefacts in digital form. Although this is a DL project, the opportunities to create new content are there in the form of annotations and discussion forums.

2.1 Crowd-based metrics

Techniques originally used for distributed information gathering across small-scale intranets may now be applied to the web as a whole thanks to increased computing and networking power. TAPESTRY was the first system to employ the idea of collaborative filtering, helping users to filter a growing number of e-mails for the most interesting and useful [28]. The system relied on two types of users; eager annotators who would read the majority of messages, providing annotations to each, and users who would wait for these annotators to provide a guide as to what was useful. The GROUPLENS project [39] uses the aggregated recommendation ratings of users to provide a score for Newsnet postings. This extended

[2] http://www.archive.org/.

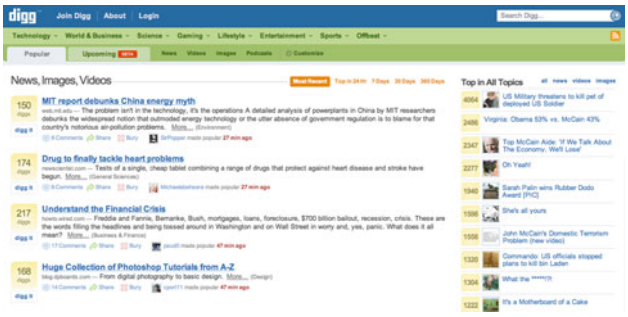


Fig. 1 The front page of Digg showing those stories that have the most votes

the work of TAPESTRY by removing the requirement of the user to choose which filters he/she wished to use, at the cost of any form of personalisation. This method when applied within the DL setting would draw parallels to overall ranking and popularity of artefacts [77] within the DL based on access, generated activity etc.

One of the most popular Web 2.0 social rating sites is Digg[3] which site allows users to post a link to web-pages, podcasts and digital content which they find interesting for users to vote on (i.e. the post is 'dugg' by other users). The more the people who vote for a post, the higher it is placed on the site's ranked list of posted 'diggs' with the goal being to have the post appear on the front/first page of the website as shown in Fig. 1. Within Digg, users are able to create lists of other users whom they wish to follow, being notified any time a post is made by those users whom they are following.

As with all open systems, the lack of centralised control leads to possibilities for abuse. For example, acquiring a lot of followers can lead to increased influence within the site and [44] has shown that friends, or followers of other users, prefer to digg the posts of each other which can lead to 'tyranny of the minority'. Unlike GroupLens, users' identities are explicit and so while ratings are effectively anonymous, the benefit of posts being 'dugg' is passed on to the user who first posted the item: the more often a post appears on the front page, the greater the likelihood of increasing the number of followers and prestige.

Within the context of the two DL systems that we present in Sect. 6 and the impact on DLs in general, the prestige or value of users to the community as a whole is measured not by voting but in terms of interaction. This idea has been used in the past to provide users with a trust or reputation score [83,87], enabling prioritisation of users and the limitation of privileges when using websites.

The idea of trust and reputation is widely used in sites such as eBay[4] to provide knowledge about other users based on past interactions with the community. While these interactions may be presented in aggregate form (number of positive recommendations, etc.) the information is present for users who wish to find it, and provides an implicit level of confidence that many others before us have had successful interactions with the party of interest. Complementary to this, Arms [9] states that libraries must be able to keep the information within them secure and safe. Levels of access to the artefacts within the library should be implemented so as to allow for the addition of completely new artefacts to the library, whilst not limiting the opportunities for additional metadata creation.

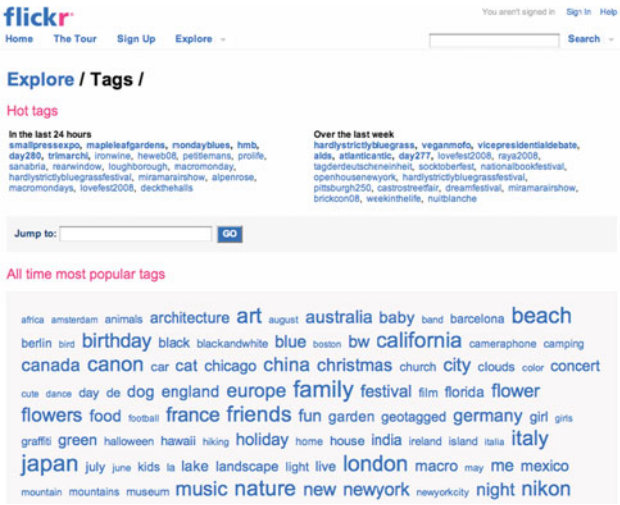


Fig. 2 Lists of tags which have been applied to photos in Flickr to aid in browsing and retrieval

Terveen et al [79] note that 'the distinct number of recommenders of a source is a plausible measure of resource quality'. This principle has been carried over to both Twitter—where the task is to find other users to 'follow' [33]—and in TripAdvisor—where the users' hotel ratings are modulated by the number of contributions they have made in the past [61]. It has also been used extensively within the bibliometric field where citations among scholarly articles create an implicit rating of a article, author, or a journal impact factor is an aggregate of the cites it receives from others. Again, this rating in the form of a bibliometric measure such as *h*-index [35] is an aggregate, but at the same time provides some explicit measure of value to a publication: citing a publication is a public statement of opinion much as an annotation is.

2.2 Tagging

Tagging of content refers to the association of descriptive words (tags) with objects to make them easier to find. Originally used by the photo-sharing website Flickr[5] to address

[3] http://www.digg.com/tour/.

[4] http://www.ebay.com.

[5] http://flickr.com.

the difficulties of searching for photos (Fig. 2), tagging has become a highly popular way of organising information. Social bookmarking site del.icio.us[6] has seen rapid growth in users since its introduction in 2003 and now has over 5 million users and 150 million tags (as of 2008). These datasets have been the focus of research into 'folksonomies' or social tagging [58,64].

A major drawback of tagging is the fact that tags are single words or concatenations of words to form single tags, e.g. 'highschool' or 'creditcrunch' and they lack depth of expression. Unless additional indicators such as the co-occurrence of tags are taken into account, there is no real way to differentiate two piece of media tagged with the same or similar tags. The reason for tagging is thus often lost, although it may in some cases be obvious (as is the case with noun tags) [12,29].

Organised ontologies may arise from the seemingly chaotic assignment of tags to resources by an uncontrolled and unrestricted user community [54]. Folksonomies—a portmanteau of the words folk and taxonomy—have been shown to aid in the retrieval process, both by providing keywords with which to search, and by using the tags given to a resource to provide context for a query [36]. Folksonomies are also becoming one of the focus-points of the Semantic Web [13] because tags created by users provide a better description of the resources they describe than anything created by a machine [54,85].

The value of tags is even more evident when browsing visual media. In conjunction with visual features, such as texture and colour, tags can provide the additional information required to provide meaningful results to image queries [10]. Within the DL domain, this additional organisation comes directly from the contributions of users and allows for both an easing of the curation burden for librarians, and the creation of a entirely new information architecture dependent on creators' needs. Casey makes the point that the addition of something as simple as tagging could greatly improve the utility of, say, an academic library by allowing users to add tags based on course numbers, assignment titles, or professors' names [19].

2.3 Social commentary

Much research has been done on using the past search/browsing behaviour of users to group them into communities who benefit from analysis of the group as a whole in recommendation of browsing or searching paths. I-Spy and HeyStaks use the idea of community to provide improved search results to a user [53,76]. This form of collaborative filtering relies on implicit links between users to group them. The very nature of annotation and the threading of those annotations however, provides an explicit link between disparate users that can be leveraged for the same purpose. This explicit communication between users shows a grouping that has been decided by the users themselves and does not rely on any computational approximation of group.

The best contemporary example of social commentary is microblogging through sites like Twitter where the creation of information is on the one-to-many model rather than the inverse [73]. This is a form of annotation, or simple commentary, on anything as diverse as reality TV shows or the availability of free WiFi in airports. A great deal of Twitter content is actually an annotation of real-world events, as they happen, and these are becoming metadata to the events that happen in our lives rather than metadata on a particular artefact in a DL [42]. Sakaki *et al.* [68] observed similar seismic activity within the Twittersphere as was happening in the real world as an earthquake and aftershocks hit the Japanese islands. It is clear that future DLs should accommodate an element of user annotation for reasons we address in the next section.

3 Annotation of content

While reading is an inherently passive activity, writing requires far more effort on the part of a writer. It is perhaps not surprising then that 'the most pervasive activity around documents is reading' [17], but the act of reading is often closely followed by annotating.

Annotation forms a bridge between the separate activities of reading and writing, allowing the reader to take an active role in the creation and dissemination of information. This *active reading* role [3] is something that has become more prominent with the advent of e-books, social computing, and specific digital annotation software. The smooth integration of annotation and reading is an essential and vital quality of any annotation system [60], something that in-context annotation provides. This is the approach we take in our own systems as detailed in Sect. 6. 'In-context' refers to the anchoring of comments in place and neighbouring the phrases/quotation to which they refer within the original document instead of being placed in a separate area. The additional information present within annotations helps users to focus users on the information that is of most interest to themselves and the user community.

In the same way as web users prefer to 'digg' their friends, they will visit websites that others have visited in the past. ASSIST [26] built on these assumptions, enabling users to see where others had browsed before them, while OATS [11] was designed to allow students to create and share annotations on course-work, augmenting the idea of tagging with free-text annotations. Free-text annotations or comments can be used to give a descriptive and semantically accurate impression of

[6] http://del.icio.us.

information that is being annotated. While the tags provide a means of categorising and clustering annotations, the annotations themselves provide information.

Annotating manifests itself in many different manners and for many different reasons. The annotation taxonomy presented by Ovsiannikov et al. [63] states that annotations may be created 'to remember, to think, to clarify, to share'. An annotation may be of any modality, audio (as in a sound-byte or song), visual (photographs or video) or most commonly written as text. It is often the case that the annotation itself is in the same modality as the document or source being annotated [4]. As is the definition, the purpose of an annotation is to provide additional explanation or clarification to the source. In doing so, a symbiotic relationship is created between annotation and source with the information in each re-enforcing the other.

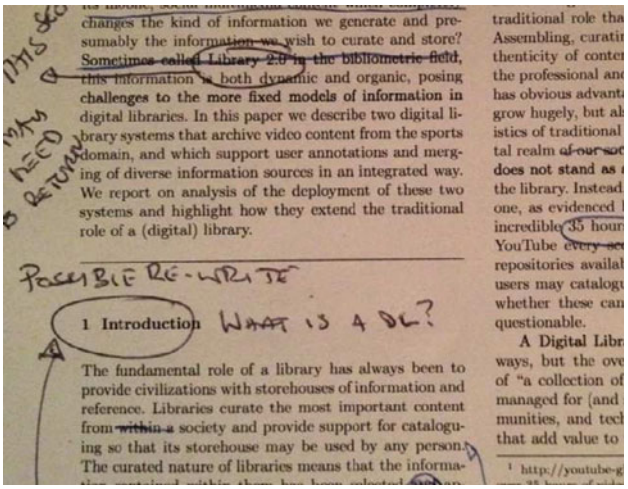
An annotation may aid in data-provenance helping to preserve information on the origins of a document, as well as interpreting it, and adding contextual information. Marshall [51] provides a thorough overview of many of the different ways in which annotations may be used.

3.1 The scope of annotations

While the purpose of an annotation may be to clarify and provide context information for an annotated source, the method of annotating can vary greatly. Annotations can be highly transient in nature, marking out a reader's current state-of-mind when reading a document. On the other hand, the persistent and permanent nature of a physical annotation can lead to a growing usefulness: notes added in a book to aid their creator may become of unexpected benefit to future readers. Duration also introduces issues with the life-span of annotations. Are annotations new, or are they just part of the metadata—as has been questioned by Agosti [6].

Most research has focussed on written documents annotated by written annotations. In the physical world, these annotations take the form of underlining, margilinia, highlighting, etc. with the exact method specific to each annotator. The vast majority of these annotations are anchored to specific points within the source such as phrases, words and paragraphs (see Fig. 3). This is mainly due to the increased effort required on the part of the annotator to recreate the context of an annotation that is recorded separately to its corresponding source [16,50]. This increased effort leads to a different style of annotation in which the information contained within the annotation is of a more general nature, recapping or summarising the source.

The use of anchored annotations gives new information about the source, as opposed to summaries of pre-existing information [84]. This distinction between in-context and separate annotations is important from the DL perspective since the choice of architecture within the DL can dictate


changes the kind of information we generate and presumably the information we wish to curate and store? Sometimes called Library 2.0 in the bibliometric field, this information is both dynamic and organic, posing challenges to the more fixed models of information in digital libraries. In this paper we describe two digital library systems that archive video content from the sports domain, and which support user annotations and merging of diverse information sources in an integrated way. We report on analysis of the deployment of these two systems and highlight how they extend the traditional role of a (digital) library.

1 Introduction

The fundamental role of a library has always been to provide civilizations with storehouses of information and reference. Libraries curate the most important content from within a society and provide support for cataloguing so that its storehouse may be used by any person. The curated nature of libraries means that the informa-

Fig. 3 Annotations made for private purpose can aid understanding when brought into the public arena

what annotation functionality is available to users. Moving from the physical to digital should always make use of the opportunity to enhance functionality.

Annotations created in the physical world are often done so in a private manner and are not appropriate for public use: the style of annotation may not lend itself to being easily understandable by anyone but the original author. This is the case for annotations such 'No!' or 'Must think on this'. These styles of annotation are not self-explanatory and can also suffer from 'crises of intelligibility' [52]. When annotations move from a private to a public context, they can lose their meaning due to their inherent qualities; they are not designed to be of any use to persons other than the annotator. This is not always the case, however, and it has been shown that in some cases private annotation can indeed be of use to the public [50,74].

This is an important point as the resources of a DL may be used by large numbers of users who may not have common information needs. The arbitrary display of all annotations may lead to frustration and information overload. These problems can be alleviated by access/viewing permissions. Examples of this within current online sources include IMDB[7] where users are required to have written a sufficiently large number of useful entries (reviews, comments and posts) before these entries are presented to the community as a whole, or Wikipedia[8] where specific editors with higher privileges may remove disruptive or unhelpful content from the site. Wilensky looks at the issues of annotation storage and display from the specific viewpoint of DLs [82] and provides some guidelines to their usage. His solution to access and preservation is to include multivalent documents that are 'highly open, highly distributed, and

[7] http://www.imdb.com.

[8] http://www.wikipedia.com.

highly extensible . . . a set of (possibly distributed) layers and behaviours, denoting a document's contents and functionality, respectively.' Through the use of these different 'layers' it is possible to incorporate different users, technologies, and extensions without having to change the original document.

Within the digital document paradigm, public and private annotations are created through the use of access rights. There is also the possibility to share annotation with only those users with suitable access rights. This has a disadvantage in that the access rights are not going to decay with time, unlike in the physical case where a private document may be traded or sold and so all annotations contained within become shared/public. This does, however, reduce the opportunity to create experiments similar to those of [50].

Through the continued growth of social networks and sharing platforms, the purpose of any single annotation can become user-centric. While annotations may be made to organise or aid navigation by one person, those same annotations may then be used by others to provide context or create the starting point for further discussion [37]. This in itself can lead to some of the concerns raised previously regarding access rights and sharing privileges [14].

3.2 Annotations as queries

Annotations focus the attention of a reader, allowing him/her to see what previous readers found of interest and importance. Besides this, annotations may be used to provide additional benefit to the current reader. Using annotations and annotated text as a means to query a DL (or the web as a whole), we may find other documents that are of interest to the reader in their current context [71]. Annotations have been shown to provide better results than automatically selected text for relevance feedback [30], annotations being of smaller size than the entire document taken as context for traditional relevance feedback [69]. Annotations more accurately reflect the intentions of the reader as opposed to traditional relevance feedback approaches, which while being statistically appropriate may not fully capture a user's intent in annotating [31].

3.3 Annotations as hyperlinks

Annotations may be used as explicit queries to a search engine or in a query-less manner as described above, creating hyperlinks between documents and enabling the reader to move between documents due to the annotations they have made. While the query-less use of annotations provides a means of linking documents, the use of annotations as hyperlinks themselves is not the same. The explicit creation of hyperlinks through annotations means that users can deliberately connect different documents. Importantly, these links are not created automatically, but manually by the annotator. These links can help to further clarify the information within an in-context annotation, or may be essential in connecting an annotation that is stored separately to the source document. Again the use of annotations helps convey what readers of the source document believe to be important rather than simply what the writer regards as important. It is also possible for information not available at the time of writing to be added in this manner, again echoing the ideas of fluid document transformations through interactions.

3.4 Grouping annotations

It is preferable to group annotations together into bundles of related annotations [63,88]. This helps us to alleviate the difficulty of searching annotations caused by annotation length which can be quite short. Grouped annotations may be thought of as similar in some way, and can be combined to form a pseudo-document used to retrieve relevant information [2]. Groups may also be created using automatic filtering techniques, allowing for temporal, length, user-specific, etc. filtering of annotations. In this way, a person can review all annotations that have happened since last viewing a document, or even just the annotations of a particular person. This grouping and filtering is easy with digital annotations whilst in the physical domain, it relies on such visual cues as handwriting, colour-coding and style of annotation (underlining etc.) to differentiate and group annotations [50,51].

The value of annotations as information in their own right has been discussed in [5], the annotations being autonomous from the source that they annotate but retaining a link. While annotations help to enrich a document by providing a focus to readers, they may also serve as a springboard for new ideas and other information. By collecting a document's annotations together, one may be able to construct a new document based on the annotations. In this way, the annotations retain a link to the original source document, whilst becoming a document in their own right [15].

4 Video digital library systems

Past research on DLs has shown that one of the many advantages of digital media is their inherent ability to represent information in different ways, allowing for access and visualisation dependent on need. This also means that different aspects of a corpus may be explored in different ways, from access patterns to article attributes. Large collections can be accessed and used by very large groups of users simultaneously and for very different purposes. In this section, we look at video DLs because of their richness and importance.

The Físchlár system [43] was designed to take advantage of the digital nature of video, providing its more than 1,500 users with viewing and recording capabilities for broadcast

TV. Besides personalised recording and programming, the system had on average 400 h of community relevant video content from news, sports and entertainment programming chosen by users. Later iterations of the system built on this allowed for cross-media querying and retrieval [22,25].

Físchlár's main purpose was to support different ways of accessing and displaying video content from within its digital corpus. Like many other video DL implementations, this means that the main focus was on content display and archiving, rather than the creation of new content: all content within in the system was stored broadcast TV video. Users were allowed to rate content so as to create a more personalised experience when content was recommended to them, but users could not annotate, upload, or collate recordings.

The Físchlár digital video library, like its contemporary Informedia [21], was a closed world in which users were able to interact with content, but not provide any meaningful contributions of their own. This is also true of many early content-based DLs or archives. While it is possible to select and filter material based on featured content (actors, scenes etc.) these filtering options are pre-defined by the creators of the archive and not adaptable to the user's needs. In contrast, the systems we describe in this article allow filtering of content that did not necessarily exist when the system was first implemented but instead has been created by the user community itself.

The Open Video Digital Library (OVDL) at the University of North Carolina also provides similar services for its users, though this time for the viewing and downloading of video [47,49]. Its main usage domains are the fields of education and research where it is used by thousands of users from around the globe. Similar to Físchlár, much of its success as a DL comes from the innovative user interface design, again providing its users many different methods of interaction with its stored content.

This ability to interact in ways that most suit the user was seen as highly important by the creators of the OVDL since they too recognised the opportunities that digital information storage and analysis could provide. They base their study on the concept of 'Agile views' [48], allowing a user to swap between different representations of the information before them with fluidity and ease. This is achieved through the use of 'visual surrogates' and distinct views. The surrogates may be thought of as abstractions of the content itself, presenting an overview or key into the associated video. The OVDL uses a selection of surrogates to create an outline of each video within its collection: a storyboard akin to a filmstrip; a flick-book style fast-forward/skim representation aimed at showing users the most pertinent elements of the video; and a 7-s excerpt containing both visual and audio contents.

The AgileViews metaphor makes an important distinction between navigation and understanding of underlying content. At its core lies the assumption that it is more useful to allow users to find an appropriate representation of the media they are engaging with than to provide simple navigational cues that may be of little use. The ever-increasing richness and depth of stored media means that new and more effective methods for information access are required. This is also very much akin to the ideas we implemented in our own systems where keyframes for events are chosen to give the user more direct and intelligent access to important and interesting content within the stored media.

Shared views and history views allow users not only to rate items but also add reviews. This type of user-generated content is something that is lacking in either of the Físchlár, Informedia or OVDL systems mentioned above, but one system, however, that does go further in its realisation of the concept is VANNOTEA [72]. VANNOTEA makes full use of the interactions of its users with video content stored in the system's database, even allowing for recording and playback of user interactions with content within a session. While VANNOTEA is not envisaged as a DL, it does provide many of the functional requirements, though perhaps not on the storage, categorisation and cataloguing side. The system was designed to provide synchronous/asynchronous annotation and video-streaming capabilities. These collaborative activities provide its users with the opportunity for content creation, but the overall system falls some way short of providing a full DL architecture. It is important to highlight that, however, VANNOTEA was one of the first few to offer collaborative, real-time, synchronous annotation of high quality MPEG-2 video content.

In more recent years, through the unrelenting creation and storage of user-generated content (UGC), large repositories of digital material have been created that may now be thought of as new forms of DLs. The functionalities that these repositories provide also cover a lot of the requirements for a video DL; however, there is no curation of the content save by the uploaders/creators of video objects.

It is clear that as the nature of a DL is changing, it should no longer be necessary to create silos of information accessible only to members or patrons of the silo's owner [19]. Resources such as YouTube show that people are willing to create and discuss content. This is brought to the extreme by [20] who show that the annotation and sharing of content can be extended to pre-existing media such as television. The ability to personalise and augment content through annotation is also shown, and making these opinions the focus of attention within the user community echoes the opinions expressed in [46] that 'content begets content'. In the video DL systems already mentioned, UGC happens as an aside, whereas in this study, we actively promote such content generation and allow it to become and to aid the focus of interest. The social nature of information gathering and dissemination is becoming central to the way in which people communicate and learn. This study, focuses on this social aspect of

information finding and sharing, allowing users to both find and share the opinions of other users.

5 A sports-focussed digital video library

In this study, we examine the usefulness and potential of user-generated content to aid in the video retrieval and browsing process. In order to do this, we focus on two high-profile sporting events and the digital media that surrounds them. We build up a digital multimedia library around these events, creating mechanisms for both the browsing and annotation of the stored content.

The proliferation of UGC and professionally created content across the web means that many different interpretations of the same event are available, and sports events are an excellent example of this. These same events are also reported post-event on the back pages of national and international newspapers as well as on other sources of real-time information sources such as Twitter. Our systems are a step to creating DLs around sports, or other events. While many media representations are already available to sports enthusiasts, these are independent and distributed without aggregation or portals or archives. Each offers the distinct option of reading professionally created reports, watching the associated video, or creating their own interpretations of events. Extending this further we have the concept of being able to share this through commentary in a public or private form. Allowing users to consume content, interpret it, and then create new interpretations is at the core of our study and a fundamental characteristic of the democratic nature of the internet. We weave together the separate strands of sports media into a complete representation that allows for better understanding and archiving of those events from a DL perspective. Our systems have focussed on the sports domain, but could be applied to such fields as news, entertainment, education or health.

By allowing users to create annotations on the content contained within our DL, we provide a more organic collection that will continue to grow by more than the additions of a library curator, and thus it will become the society's archive. Each annotation created from within the community provides new information, leveraging the viewpoints, time and expertise of all users. In the context of sport, it is easier to provide a catalyst for this kind of content generation as although newspaper articles are generally considered to be subjective, they will always create some inherent bias [78,81].

We now present, in the next section, the design of our systems. The systems differ from current commercial website as they make the annotation and discussion of content within the library a focal/intricate activity. Unlike in YouTube, the scope and domain of the associated media is limited and users share a common focus or interest. There are also explicit links created within the multimedia so as to enable navigation and focus within the different multimedia. While this focussing of domain does create limitations and frailties within the system design as mentioned later, it has also allowed us to concentrate on conversations and content created by a community that is a more obvious extension to that of a traditionally curated DL.

6 Video digital libraries

The two video library systems we developed focus on high-profile sporting tournaments: the FIFA World Cup 2006 (SportsAnno) and Rugby World Cup 2007 (Annoby). We recorded TV coverage of all games in each competition. The systems were developed and deployed University-wide, and in between the two tournaments, we carried out a series of interviews and focus group meetings and improved the interface to Annoby, learning from SportsAnno. Media articles covering the games from the tournaments were sourced from newspaper reports downloaded from the websites of newspapers. The sources for each of the newspaper reports were selected deliberately to provide differing interpretations of the games; in the Annoby system, we selected a local publication due to the presence of our own national team, as well as a publication from the traditional national rivals who were also present.

Figures 4 and 5 show the main interfaces to the two systems with several user interface (UI) differences—these are the results of both interviews with the users of our first system, and interim research between deployments that suggested UI improvements.

Both interfaces provide a list of all video content available within the DL on the left-hand side of the interface allowing for easy navigation between specific media objects. Included in this index is a count of the number of new comments made within each video since last accessing the system. All comments are placed in-context within the text of the original articles, generating a more focussed style of annotation [16]. In addition, we include the facility to reorganise and re-order the index based on annotations, match date or alphabetically. This creates something akin to the 'AgileViews' of the OVDL [47,49] where users can see the index in the context of community activity based on what generates the most annotations, or they may order the index by currency and see the newest additions to the DL.

After examining the annotations from our first system, SportsAnno (Fig. 4), it became apparent that the inclusion of two written reports was sufficient to provide alternative viewpoints on the events within the associated match video. This observation is, however, specific to sports reporting since the reports may be seen essentially as re-worded descriptions of the same events, the scope of which is quite limited. In a larger DL where the associated documents may

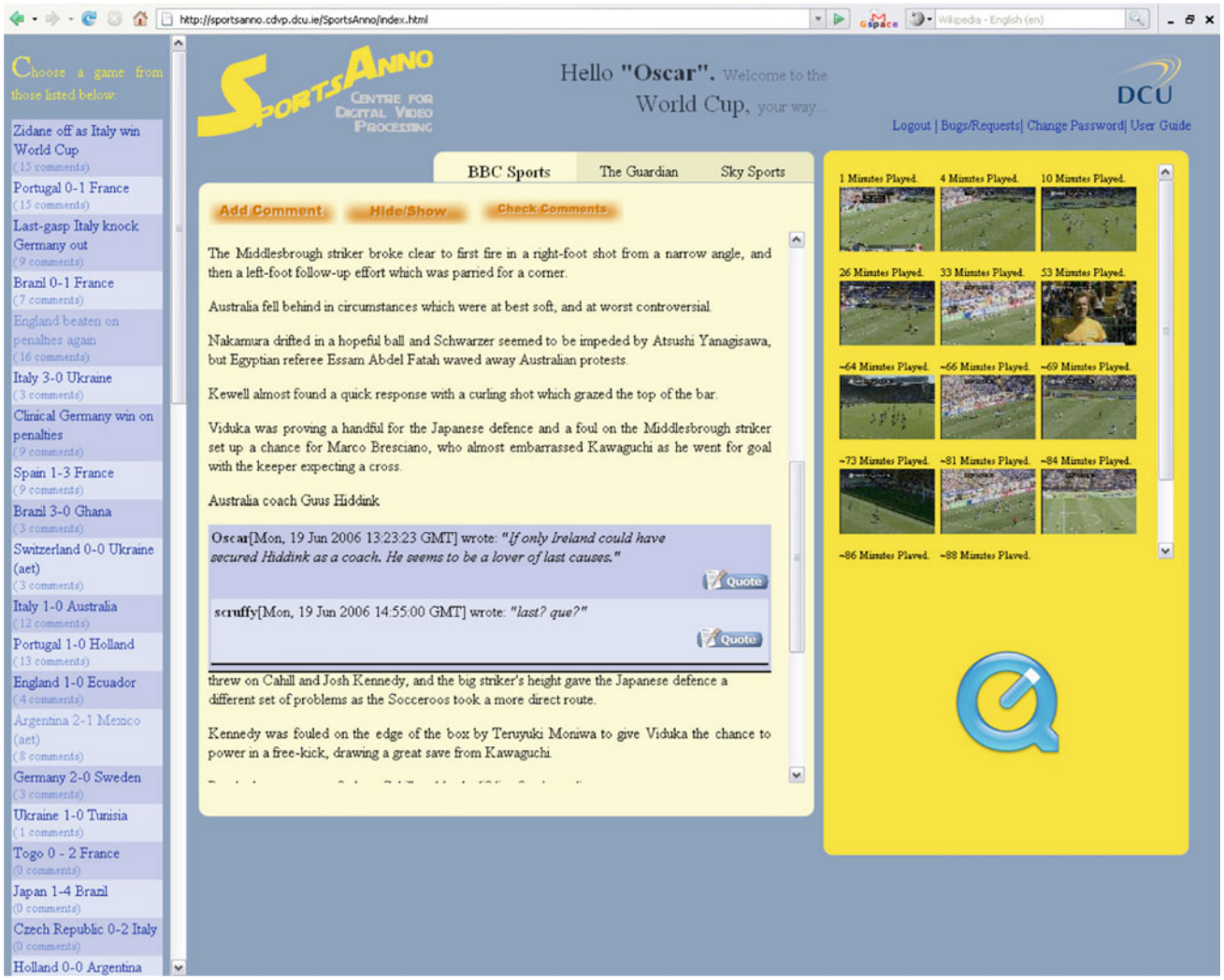


Fig. 4 The SPORTSANNO main interface

discuss completely different aspects of a piece of visual media (e.g. the camera work and cinematography of a film versus the social commentary present), two reports would not be enough.

The list of all commentators/annotators for the current video is to the right of the screen. Within ANNOBY (Fig. 5), this list is sorted by activity, contributions, and alphabetically, and shows the number of comments per annotator, as well as the number of threads that the annotator is involved in. The information that may be gained here is the level of interaction that each annotator has with the community. Commentators with many annotations in many threads may be involved in less in-depth conversations than commentators with similar numbers of annotations but far fewer threads. This ability to see at a glance those who have commented on a report was something that users of our first system indicated would be of particular interest to them. Below this list is the video player allowing playback of associated video events.

Another feature that was added to our second system ANNOBY, was display of metadata about social activity around the video. This allows for a quick understanding of all community activity centered on the object, and not solely the report as explained later. By providing an individual thread count as well as the number of annotations and commentators, an immediate gauge of the depth of conversation is shown. A game with several comments and few threads helps to show that conversation has been focussed on a few key points. It is also more likely that this conversation will yield interesting information that was not within the original reports. This feature again is something that was specifically mentioned by users of the first system as desirable (Fig. 6).

The most important enhancement made to the user interface in the second system was the combination of news reports and representative event keyframes. These keyframes represent important events (how these events are chosen in discussed below) and the reports pane at the centre of

Fig. 5 The ANNOBY main interface

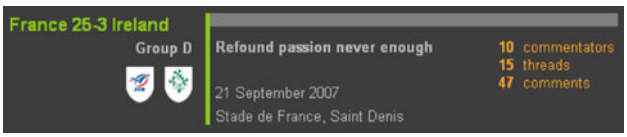


Fig. 6 Match and social metadata overview

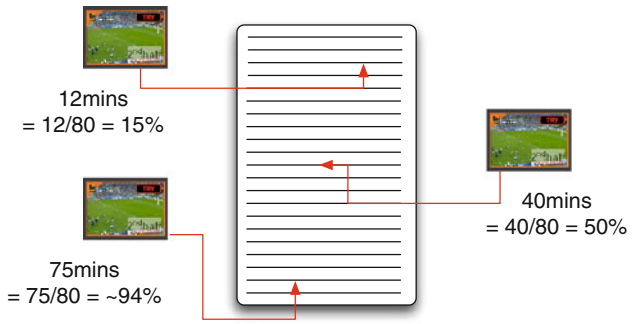


Fig. 7 Keyframes are inserted at the same offset into the written reports as they appear within the video

the screen provides a clear focal point, allowing users to read reports on the selected match, create in-context annotations, and view video highlights of key events. The presentation of keyframes and annotation in-line within the reports again illustrates the strong connection between these three information sources.

The position of keyframes in the text alternates between left and right, each keyframe being presented at approximately the same offset into the report as the percentage time into the match (Fig. 7). Sports reports are theoretically written with the first paragraph summarising the game, and major events then presented in chronological order providing an outline of the match as a whole [8]. Using this structure and without any semantic analysis, we can present video events in the region of their corresponding text descriptions. This provides a navigational link between the text reports and the video. Placing the keyframes intelligently also provides contextual information for the keyframe.

This approach is admittedly specific to sports writing and therefore could not be used within other domains. One could, however, imagine the use of text-analysis techniques such as text-tiling approaches to separate out the different topics from within a text [34]. Used in combination with speech recognition, images could then be inserted at relevant points.

Keyframes are used to both index important events within the video, and also create placeholders for annotation conversations centered directly on the video of an event. By clicking on the keyframe, playback is initiated at that event and users can also click on the marker on the top left of the keyframe to create and read annotations about the event. Before annotations are created, a grey border (as seen at the bottom of

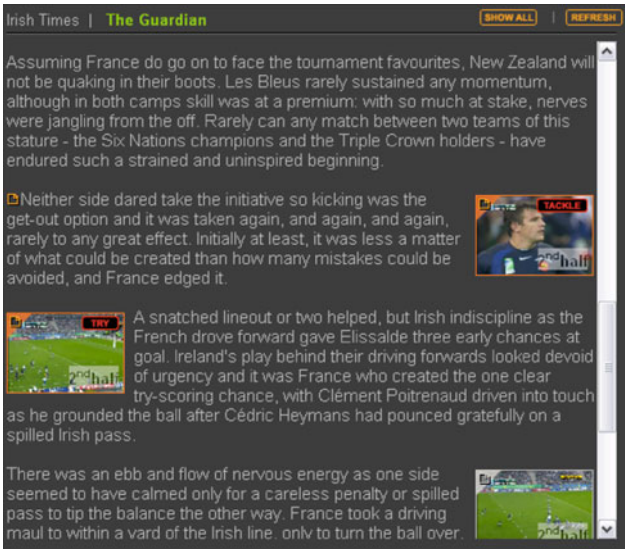


Fig. 8 The Reports Panel with in-context annotations. Various buttons along the top allow for annotations to be hidden/revealed, as well as viewing of each of the different reports

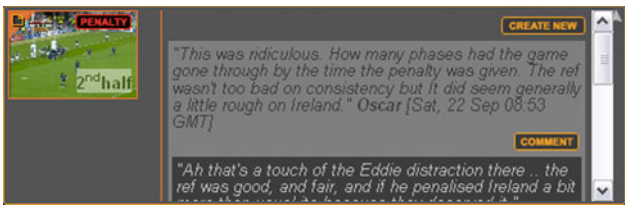


Fig. 9 Comments may be made directly on the keyframes representing an event within the video

Fig. 8) is used to signify the absence of annotations, orange borders are used to signify annotations exist. Although this is not strictly direct annotation of the video itself, the annotation of keyframes does allow for annotations centered on a particular event within the video rather than on a point raised within the written report.

As stated in Sect. 2.2, tags provide an important mechanism for generation of a user-centric folksonomy that is more focussed on the contextual needs of users. Owing to the size of our collections, only a couple of hundred hours of video, the provision of user-generated tags was seen as excessive, though we do not doubt the necessity of this feature within the context of larger DLs. Instead of free-form tag generation, we provided an ontology of domain-specific tags that had already been applied to the video events within the collection. An example of this may be seen in Fig. 9.

In both SportsAnno and Annoby, it was of great importance to make annotating as easy and intuitive as possible. Annotations are created by highlighting a portion of text within the reports creating a pop-up window to insert a comment box into the report. Users may create annotations anchored on any sentence or paragraph within the original reports. Replies to annotations are considered to be focussed the entire text of their parent annotation. While the option to show all annotations is still present via the 'Show All' button, single annotations may be shown by clicking on the small orange annotation symbol, an example of which may be seen before the second paragraph in Fig. 8.

Once the annotation thread is revealed, replies to an annotation may be added by clicking the 'comment' button, or the comment box is hidden by clicking on the small grey arrow in the top corner (Fig. 9). The alternating background colour for each post is used to signify the depth of the comment, but if two replies are posted to the same parent, the same colour background is used.

It is a recognised weakness within the approach taken for both our systems that as the number of annotations grows, the possibility that the reports panel may become illegible also grows. For this reason, we see the current approach of placing every comment within the text, and with its own marker, as a temporary design. If a system similar to this were to be deployed on web-scale, then decisions would have to be taken as to how to display all comments, or indeed how to filter those being shown. The authority or influence as discussed in Sect. 7.3 is one possible solution to deciding which comments provide the most information or stimulate community interaction.

Our systems are designed around the concept of a DL that is not limited just to search or retrieval, but also content creation and discussion. They facilitate more interactive experiences for their users focussed on sports events. The detection and representation of events within the associated video stream builds on the study of [67].

Post-processing of recorded match video removed all non-game footage, e.g. studio discussions, advertising etc. In this way, the analysed video begins just before the initial whistle is blown and ends just after the final whistle. The video length itself is a minimum of 80 or 90 min, this being the length of a rugby or soccer match. The systems use an XML backbone so as to enable easy integration of existing standards and easy extensibility. All data required by the system are stored within an XML database.

In order to establish segment boundaries, we use a shot boundary detection algorithm [62] to find abrupt shot cuts, and then the study of Sadlier and O'Connor [67][9] to combine these shots into segments that are viewable/watchable for our users. For a more detailed description of the system architecture the reader is refered to [41]. After finalising the segment boundaries, an MPEG-7 representation contains all

[9] The detection approach used is multimodal and relies on both audio and visual information streams to determine confidence levels. Six classifiers are used to detect the presence of player close-ups, crowd shots, scoreboard changes, increased audio activity, playing field boundaries and increased visual activity.

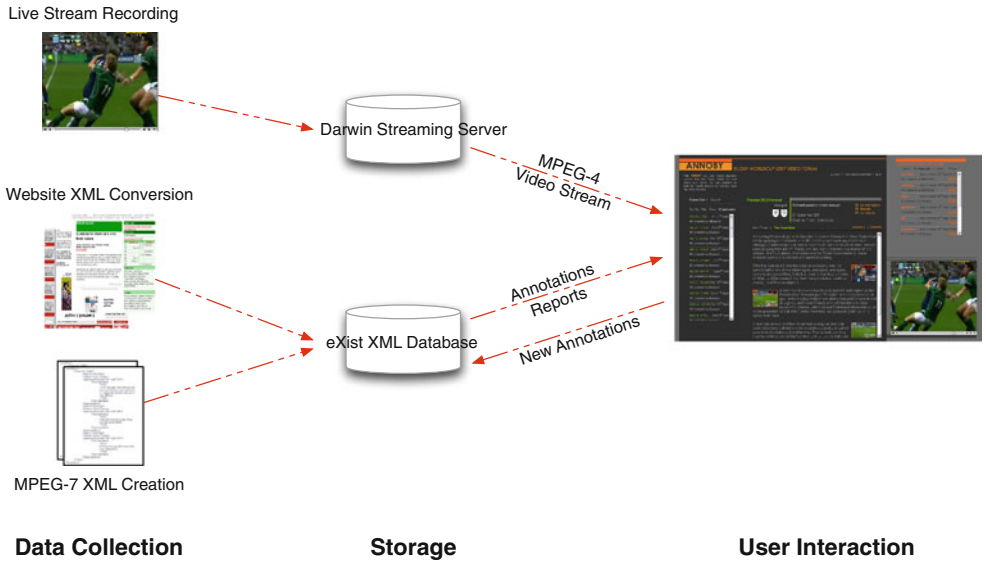


Fig. 10 ANNOBY system architecture

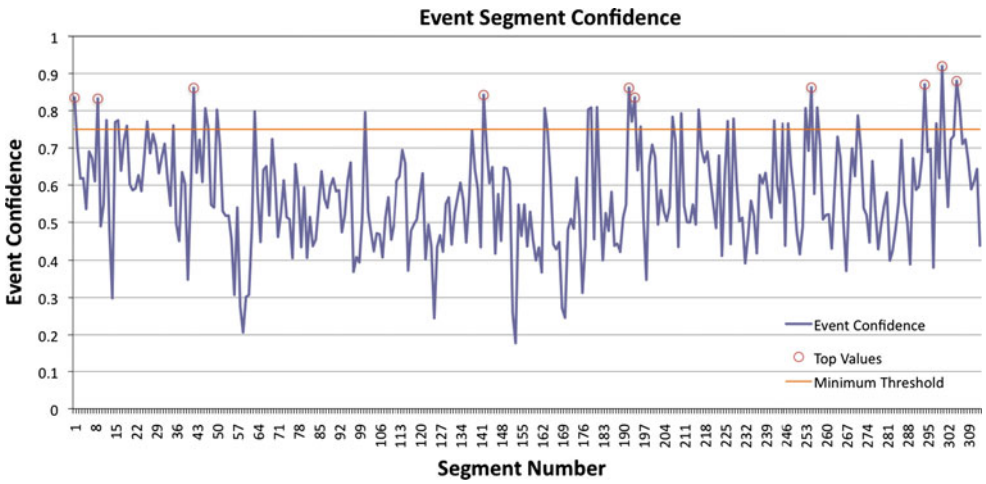


Fig. 11 Only the top ten most exciting events a confidence higher than the minimum threshold are returned

shot information including duration, start point and confidence for an event occurring during each segment.

These segments are created by looking for the highest confidence values returned by the event detection algorithm, and then taking a 20-s window around these values. We limited the number of events to the top ten only.

6.1 Usage study

These systems were deployed in real-world scenarios and not in a closed laboratory environment. We deployed both SPORTSANNO and ANNOBY during the respective sports tournaments, the soccer and the rugby world cups. We broadcast information about its availability and made it available within the university community, requiring login and authentication to access. This is something that hampered our efforts to generate a large user-base. During their deployment, we were able to collect the reactions and conversation (annotations) created by a population of interested users. The system was used in a catch-up or review mode, typically the morning following a match, which was contrary to what we had initially hypothesised (Figs. 10 and 11).

User demographics

Owing to constrains beyond the control of the authors, both SPORTSANNO and ANNOBY were closed in nature, and so the user base consisted of people either directly known by the authors or know by a direct colleague of the authors. Registration periods for both systems ranged from before the start of their respective competitions to any time during them. All games were made available to all users, and so even those

who registered late could browse and comment on any match including those played before registration.

Seventy people registered with the SPORTANNO system, and of these 25 were researchers within the group who had experience of annotation systems. A further 12 came from associated research institutes who would again have had experience with annotation. The rest of the user community was made up from friends of registered users. Of the 70 people who registered, 24 made no further visit to the system. Of the remaining 46, 24 were active browsers viewing the comments left by others but not contributing themselves. 22 users made comments and took part in discussions about particular events within each match.

ANNOBY gained a slightly larger userbase of 89 people: Again 25 were researchers within the group who have had had experience of annotation systems. 14 of the registered users for ANNOBY had also registered and actively used the SPORTSANNO system. Of the 89 people who registered, 50 made no further visit to the system. Of the remaining 39, 20 were active browsers returning to the system at least twice and viewing the comments left by others, but not contributing themselves. The remaining 19 users took part in active discussion, creating the annotation corpus for the ANNOBY system. These users were not solely those with past experience of annotation systems.

Corpus information

Across the two systems, over 130 h of video data were recorded. They consisted of 102 matches, accompanied by 258 newspaper reports. These matches and reports were fully parsed and indexed for use in the system. The total number of annotations made on the ANNOBY system was 411, slightly higher than the 338 made in the SPORTSANNO system[10] The distributions of these comments across the reports, however, are roughly similar.

The time between first posting the match to the website and the last comments on a game being made was also recorded. Owing to the type of data being presented, it is not altogether surprising that the number of comments made on a match fell dramatically 3 days after its first posting. Some games proved exceptional, mainly those involving teams that stayed in the competition for longer period. Users did post comments on earlier performances involving teams such as Germany (the hosts) and France (the current champions of the World Cup), but in general, comments were of a more immediate and transient nature.

[10] The effect of the National team's presence in the Rugby World Cup 2007 should not be ignored, these games making up 152 annotations. The most highly commented games in SPORTSANNO was the first game of the national rival with 29 annotations. This is not a surprising result as the hype surrounding these teams within the media of both countries generated lots of talking points.

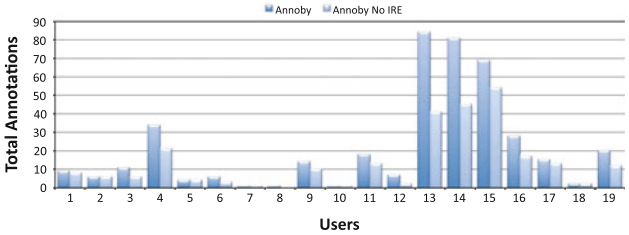


Fig. 12 Number of annotations per user of ANNOBY with/without the inclusion of National Team's matches

The number of days after which the game was commented on was affected most strongly by the advent of weekends (during which very few comments were made) and the amount of time between the recording of the matches and when they were made available on the SPORTSANNO website. This time varied from same day to 2 days after the recording date. It is also thought that lack of a notification system prevented discussion from having an average life-span of greater than 3 days, as mentioned earlier.

In the analysis that follows, we present statistics for both ANNOBY and SPORTSANNO systems. While we are aware that there are several factors which need to be taken into account when making comparisons between the systems (user familiarity with annotation systems; incident levels within the respective sports; overall viewership figures for both sports), the underlying purpose of both systems is identical.

Conversation statistics

Before discussing our finding, we make the make the important observation of the presence of the authors' national team within the Rugby World Cup 2007. The effect of this cannot be underestimated since the number of annotations received by each of their games is far above the average number of annotations per game. In the Football World Cup of SPORTSANNO, the national 'rival' again focussed attention meaning that again the number of annotations were far above average per game. For these reasons, we have presented each of our results and analysis in two ways: we include all national team games, as well as the highly commented rival's first game in our analysis, and then exclude these five games.

The effect of removing these games can be seen in Figs. 12 and 13 where the total number of comments per user is displayed. By removing the games we see a dramatic decrease in the number of comments made by the most active users. The number of users who posted comments on each system is also reduced, showing that some users commented solely on these five games.

The ratio of commentators to comments shows that commenting is a useful way in which to generate discussion within a group. Figures 14 and 15 show the average number of replies received to annotations made by users of both

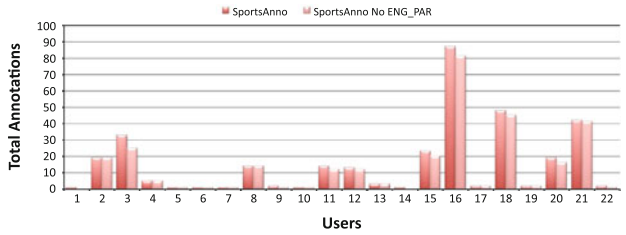


Fig. 13 Number of annotations per user of SPORTSANNO with/without 'Rival''s First Match

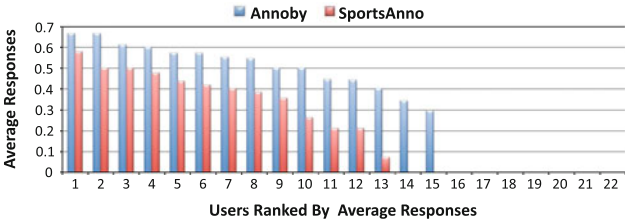


Fig. 14 Average responses per user including Natational/Rival matches

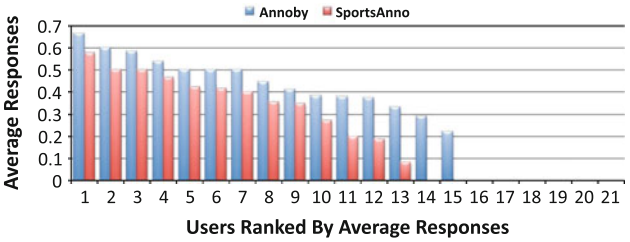


Fig. 15 Average responses per user excluding Natational/Rival matches

the ANNOBY and SPORTSANNO systems. It is clear that the number of replies received by postings in the ANNOBY system is on average higher. As noted above, the total number of annotations by the top ranked users (rank by annotation creation) is greatly reduced national/rival games are not considered. This leads us to believe that there is a great amount of conversation being held around these games, mostly by highly active users. The average responses to comments in ANNOBY, however, remain greater even with the exclusion of these games. If this were the only reason for users being more interactive and conversational, then we would not expect this to be the case. This fact, as well as the answers to a survey carried out, led us to believe that the ANNOBY system made conversation-building easier and more engaging.

The average number of text annotations per game in ANNOBY (6.146) was similar to that of SPORTSANNO (6.115); however, games in ANNOBY also received an average of 2.427 image annotations. The significant impact of national/rival games may be seen in the fact that without them, the average total annotations per game for SPORTSANNO (5.679) and ANNOBY (5.886) are almost identical. From these figures, it would appear that the introduction of keyframe (and implicitly video) annotation does not in fact increase the average

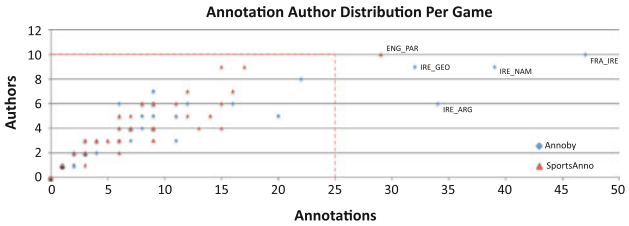


Fig. 16 Correlation: SPORTSANNO(0.900) ANNOBY(0.854)

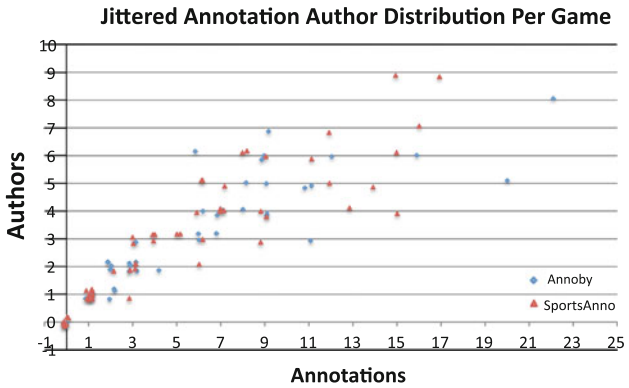


Fig. 17 Correlation: SPORTSANNO(0.999) ANNOBY(0.858)

number of user annotations, but instead replaces an equal number of text annotations.

We can see that there is almost perfect correlation between the number of users and number of annotations created in the SPORTSANNO system, especially with the removal of the rival's first game as in Fig. 17 (The data points in Figs. 16 and 17 have been *jittered*[11] so as to allow a clearer visualisation of the data.). While correlation is still strong in ANNOBY, the weakened correlation echoes the observation of Figs. 14 and 15; increased responses to comments made by users in the ANNOBY system show that while semi-direct video annotation does not increase the number of annotations created by users, it does seem to increase the conversation and interaction of users.

Observations and reactions

As with SPORTSANNO, after completion of the experiments we asked a random selection of eight users complete a questionnaire about their experiences with the ANNOBY system. The experiences of users with the new system (which took into account the suggestions made by users of the SPORTSANNO system) seem to, on the whole, have been good. The users surveyed ranged from the most active to those who

[11] Jittering is a process by which a small positive/negative number is added to the value of data points. By doing so, the distribution of values may be more easily observed. In the example in Fig. 17, we are able to see the two distinct plots.

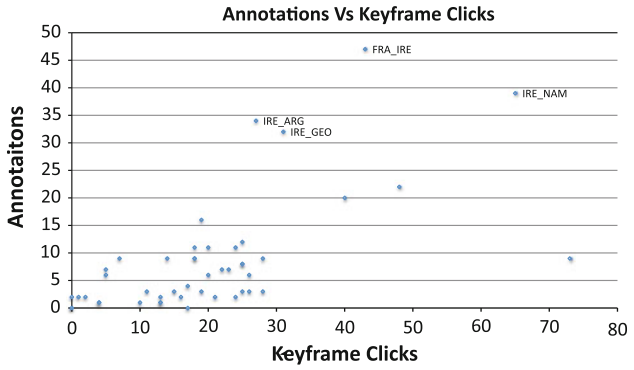


Fig. 18 The Number of Annotations Made versus the Number of Keyframe Clicks Per Game

spent more time browsing and reading the comments of others.

Initial assumptions regarding the usage of the system were disproved; it had been assumed that users would take advantage of the recorded matches to catch-up on and summarise the matches which they had not seen. In fact, those surveyed preferred to browse the games that they had already seen, rarely bothering with games that they had not watched live. Only 25% said they used the system primarily to watch highlights of games they had missed live. Far more important was the commenting aspect of the system with over 85% of users saying the primary reason for using the system was to share comments about the games they had watched live on TV, closely followed by browsing the opinions of other users. The viewing of highlights clearly played an important part, however, in reminding users of important event. We can see from Fig. 18 that keyframe clicks far outweigh the amount of annotations made per game. This is as a result of users who browse but do not annotate ($\sim$ 50% of active users), as well as those who made annotations after viewing the associated highlights.

The fact that users of the system preferred to comment on the games which they had seen live was interesting since only one of the surveyed user had seen more than half the games broadcast during the competition. Most had seen between 6 and 10 games. Usage of the system was typically between 5–15 min, with the majority of users stating they had used the system as a way to view highlights. This short usage time is seen as beneficial to the ideas used when creating the system. Both ANNOBY and SPORTSANNO were designed to facilitate summarisation of events within the original media, and community participation. The short usage times are seen as an indication that users did not see the need to spend large amounts of time browsing through games to find what they needed. Perhaps if some form of instant messaging service similar to those of social networking sites[12] was integrated into the interface, users may have spent longer on the system. The lack of instant feedback meant that users were more prone to checking the system for new information rather than browsing for extended periods.

All users surveyed said they followed sports regularly, watching highlights or live broadcasts on the television. Afterwards, users went to internet forums and websites to find more in-depth analysis and commentary. It was frequently stated that the ability to annotate and view matches in the same place was of great benefit.

Additional requested features for ANNOBY were similar to those of SPORTSANNO. The creation of threads that were not anchored to any specific point within a report, allowing for creation of general conversation was desirable. Our intention with these systems was to build a corpus of annotations focussed on specific moments within accompanying media, something which is less likely when general conversation is allowed [17]. Though it was not requested, the idea of external resources being added to the system (fan videos, or additional reports) is something worth pursuing. In our current context, however, it was beyond the scope of our initial goals.

Another important requirement is notification of replies (i.e. e-mail updates) to a user's comments. As mentioned with SPORTSANNO and previously [16,18,70] the lack of a notification system can reduce the amount of interaction undertaken with the system by users. It also requires additional effort on the part of the users to re-find their own comments and check for replies. Analogous to this, the ability to see all comments made by a specific user throughout the system was also requested, allowing for tracking of particular users by others and the creation of a more formal social network.

The two systems we have built have enabled us to explore the requirements and attractions of a contributive, multimedia DL. Through the creation of SPORTSANNO and subsequent refinement of ideas and presentation within ANNOBY, we have learned a lot about what features are most necessary to allow community interaction and participation. In the next section we describe some of the important issues raised in developing such systems and how these impact the view of what makes a DL.

7 Issues in modern (video) digital libraries

Building and deploying video management systems like SPORTSANNO and ANNOBY as well as our earlier study on the Físchlár systems, has revealed several issues which are important for future development of library systems which aim to archive the modern world of mobile and social content creation and sharing. Among these are the issues of data quality, trust and authorship, and measuring the value of authorship. We now examine each of these in turn.

[12] http://blog.facebook.com/blog.php?post=12811122130.

Table 1 Mapping attributes of Data Quality on the web by Zhu and Gauch [89] to Data Quality in contributive video DLs

Currency "How recently a web page has been updated, measured as the timestamp of the last modification of the document".
→ How long has an annotation been in the system ? What is the timespan of an entire thread?
Availability "The number of broken links on a web page divided by the total numbers of links it contains"
→ Not really applicable in this scenario but may be relevant if private annotations are considered.
Information-to-Noise "The proportion of useful information contained in a web page of a given size meaning the ratio of the total length of the tokens after preprocessing divided by the original size of the page" → This same definition holds.
Authority "The reputation of the organization that produced the web page based on a (then) reliable reviewing site which examined who had created the content and what is their influence and standing within the community of users?" → This is provided by AuthorRank.
Popularity "How many other web-pages point to this particular web-page, i.e. the number of inlinks"
→ How many replies has this annotation received, and who has been replying?
Cohesiveness "How closely related to each other are the major topics of a web page?"
→ Not really applicable in this scenario since annotations are in-context, but could be of relevance if the actual text of annotations is compared to the text of the annotated document

7.1 Data quality

In this study, we have attempted to look at the quality of the data or commentaries provided by users. One of the many reasons for creating automatic metrics that judge this quality/utility is the removal of any requirements on the users themselves to rate other's comments. The automation of metrics also provides a means of removing user bias from the measurement, these metrics being based on global information rather than any one user's knowledge. The approach of [89] fits nicely with our ideas for providing an automatic quality measure to the contributions/annotations of users. They give 6 attributes on which the quality of web pages (and implicitly the data within those web pages) may be judged. We adapt their idea of quality measures for web pages to the annotations provided to a web-page on sports events as in ANNOBY or SPORTSANNO.

7.2 Trust and authorship

Trust is a complex notion involving many different considerations. It has in the past been viewed in the context of environments such as recommender systems [59], economics [23], on-line [80] and social networks [27]. Trust was originally studied in the context of encryption and security but there has been a growth in research which views trust from a more interactive and societal stand-point. In the context of contributive DLs, it is relative because the sources of DL contributions and annotations have no specific authority and so they need to be validated.

Several of the models that are discussed by [66] use a propagated system of reputation where agents build up a level of trust in other agents by querying their established contacts. The idea here is to help cope with the sparsity problem which occurs in large online systems. With a very large network of users, most of whom are engaged in one-time interactions, it is hard for any one agent to build up a trust rating for every user. Instead we must rely on trust propagated through others. This system is known as a *'web of trust'*. Abdul-Rahman [1] implemented the first specific computational version, but it was originally proposed by Zimmermann in 1994 [90]. In the context of our own study and of trust and reputation in DLs, we implement a weighting scheme for author annotations based on the co-occurrence of annotations within both threads and web-pages (See Table 1).

7.3 Measuring and guiding browser interactions

In a contributive DL it is highly desirable to rate authors and their contributions, something that in turn can be used to filter out or highlight important insight and discussion around the documents in the DL. One way to weight users relative to each other is by measuring the quality of the information that each provides to the community. We developed two techniques using the theoretical basis provided by [89] that provide information on who are the most influential authors/participants based on their overall contributions to the topic being discussed. A complete explanation of these techniques is outside the scope of this article, but the main premise behind these is that importance flows from commentator to annotation to document [40].

By focussing on influential authors, and adjusting the system's responses to user behaviour accordingly, it is hoped that users may be provided with the most interesting and informative browsing experience. A similar idea is employed by [36] to aid in ranking pages tagged by a popular social-bookmarking site. If we then go one step further and focus on the conversations between the top ranked authors, we can

find documents that are both most likely to satisfy the user's needs, and are also most likely to serve as the anchor for informative and insightful annotations.

Like hubs [38], the more interesting an author is, the more conversation they promote. Similarly documents that contain many annotations may be considered to be more interesting or important by virtue of the fact that more people are interested in the conversation being had [24,86]. It may also however be a reflection on the material which is being annotated. This equally validates the assumption that longer threads have held the readers' attention for longer, and are therefore more interesting.

8 Conclusions

In this article, we have addressed some of the issues of how DLs can accomplish their traditional role of archiving the most important information in a society while at the still time accommodating the type of information which characterises the present time, namely multimedia, mobile, and where end users are both readers, and contributors. We examined some of the characteristics of the modern era of information generation, focusing on user annotation rather than tagging of existing content, and on digital video as the example of multimedia. To add some substance, we presented two case studies, two systems which had been developed to archive two major world sporting fixtures but which allowed user generated content to form an integrated component of the library archive. We described the two systems, how they worked and what the user interaction and experience was.

The description of these two systems, ANNOBY and SPORTSANNO, was included to highlight several issues which modern (video) DLs need to address, mostly revolving around the quality of data which is user-generated, the trust that the library and its readers might have to place in authors of such user-generated content given that there is no custodial role and how we could measure the value of such authorship. Our conclusions are that there are mechanisms to address each of these issues, and in this study, we have applied these to relatively small and self-contained (video) libraries, but these need to be scaled up and applied to much larger archives.

We also suggest that the quality metrics—for data such as annotations and for users who contribute such annotations—might form a useful part of a DL's search system. We believe it is possible to improve library searching, especially searching multimedia information, using annotations and threads. When searching for multimedia artefacts using their annotations, the scores for authors and for messages involved in discussion threads can be taken into account, going beyond the level of sophistication of links-based PageRank scoring: It is not just the links that are considered, but also the creators of these links. This distinction is becoming ever more important as we move away from the author-centric publishing paradigm of yesteryear and towards the prevalent democratisation of content creation, publishing, and sharing.

Acknowledgments We would like to thank the reviewers of this paper for their insightful and useful comments. This study was supported by Science Foundation Ireland under Grant No. 07/CE/I1147.

References

1. Abdul-Rahman, A.: The PGP trust model. J. Electron. Commer. **10**, 27–31 (1997)
2. Abel, F., Frank, M., Henze, N., Krause, D., Plappert, D., Siehndel, P.: GroupMe!—Where Semantic Web Meets Web 2.0. In: The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007+ ASWC 2007, p. 871, Busan, Korea. Springer-Verlag, New York (2007)
3. Adler, M.: How to Read a Book. Touchstone, Austin (1972)
4. Agosti, M., Bonfiglio-Dosio, G., Ferro, N.: An historical and contemporary study on annotations to derive key features for systems design. Int J Digit Libr **8**(1), 1–19 (2007)
5. Agosti, M., Ferro, N.: Annotations: enriching a digital library. In: Research and Advanced Technology for Digital Libraries, Proceedings of the 7th European Conference on Digital Libraries, Trondheim, 17–22 Aug (2003)
6. Agosti, M., Ferro, N.: Annotations on Digital Contents. Int J Digit Libr, **6**(2), 124–138 (2006)
7. Anderson, C.: The long tail: why the future of business is selling less of more. Hyperion, New York (2008)
8. Andrews, P.: Sports Journalism: A Practical Guide, chapter 5. pp. 48–50. Sage, London (2005)
9. Arms, W.Y.: Key concepts in the architecture of the digital library. D-Lib magazine **1**(1) (1995)
10. Aurnhammer, M., Hanappe, P., Steels, L.: Augmenting navigation for collaborative tagging with emergent semantics. In: ISWC 2006: The 5th International Semantic Web Conference, vol. 4273, p. 58. Springer, Athens (2006)
11. Bateman, S., Farzan, R., Brusilovsky, P., McCallan, G.: OATS: The open annotation and tagging system. In: I2LOR '06: Proceedings of 3rd Annual E-Learning Conference on Intelligent Interactive Learning Object Repositories, Montreal (2006)
12. Begelman, G., Keller, P., Smadja, F.: Automated tag clustering: improving search and exploration in the tag space. In: Collaborative Web Tagging Workshop at WWW2006, Edinburgh (2006)
13. Berners-Lee, T., Hendler, J., Lassila, O. et al.: The semantic web. Sci. Am. **284**(5), 28–37 (2001)
14. Besmer, A., Lipford, H.: Tagged photos: concerns, perceptions, and protections. In: CHI EA '09: Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '09, pages 4585–4590. ACM, Boston (2009)
15. Bottoni, P., Levialdi, S., Rizzo, P.: An analysis and case study of digital annotation. In: DNIS 2003: Databases in Networked Information Systems, Third International Workshop, vol. 2822/2003, pp. 216–230, Aizu (2003)
16. Brush, A., Bargeron, D., Grudin, J., Borning, A., Gupta, A.: Supporting interaction outside of class: anchored discussions vs. discussion boards. In: Proceedings of ACM CHI 2002, Minneapolis, 20–25 Apr (2002)
17. Brush, A.J.B., Bargeron, D., Gupta, A., Cadiz, J.J.: Robust Annotation Positioning in Digital Documents. In: CHI '01: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 285–292. ACM, Seattle, 31 Mar–5 Apr (2001)

18. Cadiz, J., Gupta, A., Grudin, J.: Using web annotations for asynchronous collaboration around documents. In: CSCW '00: Proceedings of the 2000 ACM Conference on Computer Supported Co-operative Work. Philadelphia, Dec (2000)
19. Casey, M., Savastinuk, L.: Library 2.0: service for the next-generation library. Libr. J. **131**(14), 3 (2006)
20. Cesar, P., Bulterman, D.C., Geerts, D., Jansen, J., Knoche, H., Seager, W.: Enhancing Social Sharing of Videos: Fragment, Annotate, Enrich, and Share. In: MM '08: Proceeding of the 16th ACM International Conference on Multimedia, pp. 11–20. ACM, Vancouver (2008)
21. Christel, M., Kanade, T., Mauldin, M., Reddy, R., Sirbu, M., Stevens, S., Wactlar, H.: Informedia digital video library. Commun. ACM **38**, 57–58 (1995)
22. Cooke, E., Ferguson, P., Gaughan, G., Gurrin, C., Jones, G., Le Borgne, H., Lee, H., Marlow, S., Donald, K.M., McHugh, M., Murphy, N., O'Connor, N., O'Hare, N., Rothwell, S., Smeaton, A.F., Wilkins, P.: TRECVID 2004 Experiments in Dublin City University. TRECVID 2004—Text REtrieval Conference TRECVID Workshop, Gaithersburg (2004)
23. Das-Gupta, P.: Trust as a commodity. In: Trust: Making and Breaking Cooperative Relations, pp. 49–72. Blackwell, New York (1988)
24. Fiore, A.T., Tiernan, S.L., Smith, M.A.: Observed behavior and perceived value of authors in usenet newsgroups: bridging the gap. In: CHI '02: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 323–330. ACM, New York (2002)
25. Foley, C., Gurrin, C., Jones, G., Lee, H., Givney, S.M., O'Connor, N., Sav, S., Smeaton, A.F., Wilkins, P.: TRECVid 2005 Experiments at Dublin City University. TRECVid 2005—Text REtrieval Conference TRECVID Workshop, Gaithersburg (2005)
26. Freyne, J., Farzan, R., Brusilovsky, P., Smyth, B., Coyle, M.: Collecting community wisdom: integrating social search & social navigation. In: Proceedings of the 12th International Conference on Intelligent User Interfaces, pp. 52–61. ACM Press, New York (2007)
27. Golbeck, J., Hendler, J.: Accuracy of metrics for inferring trust and reputation in semantic web-based social networks. In: The 14th International Conference on Engineering Knowledge in the Age of the Semantic Web, EKAW 2004, volume 3257/2004, pp. 116–131. Whittlebury Hall, Northamptonshire (2004)
28. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. Commun. ACM **35**(12), 61–70 (1992)
29. Golder, S., Huberman, B.: Usage patterns of collaborative tagging systems. J. Inf. Sci. **32**(2), 198–208 (2006)
30. Golovchinsky, G.: What the query told the link: the integration of hypertext and information retrieval. In: HYPERTEXT '97: Proceedings of the 8th ACM Conference on Hypertext, pp. 67–74. ACM , Southampton (1997)
31. Golovchinsky, G., Price, M.N., Schilit, B.N.: From reading to retrieval: free-form ink annotations as queries. In: SIGIR '99: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 19–25, Berkeley (1999)
32. Habib, M.: Toward Academic Library 2.0: Development and Application of a Library 2.0 Methodology. Master's thesis, School of Information and Library Science, University of North Carolina, Chapel Hill (2006)
33. Hannon, J., Bennett, M., Smyth, B.: Recommending twitter users to follow using content and collaborative filtering approaches. In: RecSys '10: Proceedings of the 4th ACM Conference on Recommender Systems, RecSys '10, pp. 199–206. ACM, Barcelona (2010)
34. Hearst, M.A.: TileBars: Visualization of Term Distribution Information in Full Text Information Access. In: CHI '95: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 59–66. ACM Press/Addison-Wesley Publishing Co., Denver (1995)
35. Hirsch, J.: An index to quantify an individual's scientific research output. Proc. Nat. Acad. Sci. **102**(46), 16569–16572 (2005)
36. Hotho, A., Jaschke, R., Schmitz, C., Stumme, G.: Information retrieval in folksonomies: search and ranking. In: ESWC '06: Proceedings of the 3rd Annual European Semantic Web Conference, vol. 4011, p. 411. Springer, Budva (2006)
37. Joinson, A.N.: Looking at, looking up or keeping up with people?: Motives and use of facebook. In: CHI '08: Proceeding of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems, CHI '08, pp. 1027–1036. ACM, Florence (2008)
38. Kleinberg, J.M.: Authoritative Sources In a Hyperlinked Environment. In: SODA '98: Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms, pp. 668–677. Society for Industrial and Applied Mathematics, Philadelphia (1998)
39. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., Riedl, J.: GroupLens: applying collaborative filtering to usenet news. Commun. ACM **40**(3), 77–87 (1997)
40. Lanagan, J.: Social impact retrieval: measuring author influence on information retrieval. PhD thesis, Dublin City University, Dublin (Sept 2009)
41. Lanagan, J., Smeaton, A.F.: Query independent measures of annotation and annotator impact. In: ESAIR '09: Proceedings of the WSDM '09 Workshop on Exploiting Semantic Annotations in Information Retrieval, pp. 35–38. ACM, Barcelona (2009)
42. Lanagan, J., Smeaton, A.F.: Using twitter to detect and tag important events in live sports. In: Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011), Barcelona (July 2011)
43. Lee, H., Smeaton, A.: Designing the user interface for the Físchlár digital video library. J. Digit. Inf. **2**(4) (2006)
44. Lerman, K.: Dynamics Of Collaborative Document Rating Systems. In: WebKDD/SNA-KDD '07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, pp. 46–55. ACM, New York (2007)
45. Levy, D.M., Marshall, C.C.: Going digital: a look at assumptions underlying digital libraries. Commun. ACM **38**(4), 77–84 (1995)
46. Marchionini, G.: Augmenting library services: toward the sharium. In: Proceedings of International Symposium on Digital Libraries, pp. 40–47, Tsukuba (1999)
47. Marchionini, G., Geisler, G.: The open video digital library. D-Lib Magazine **8**(12), 1082–9873 (2002)
48. Marchionini, G., Geisler, G., Brunk, B.: Agileviews: A Human-Centered Framework for Interfaces to Information Spaces. In: Proceedings of the Annual Conference of the American Society for Information Science, pp. 271–280. American Society for Information Science and Technology, Silver Spring (2000)
49. Marchionini, G., Wildemuth, B., Geisler, G.: The open video digital library: a Möbius strip of research and practice. J. Am. Soc. Inf. Sci. **57**(12), 1629–1643 (2006)
50. Marshall, C.C.: Annotation: From paper books to the digital library. In: DL '97: Proceedings of the 2nd ACM International Conference on Digital Libraries, pp. 131–140, Philadelphia (1997)
51. Marshall, C.C.: Toward an ecology of hypertext annotation. In: HYPERTEXT '98: Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia : Links, Objects, Time and Space—Structure in Hypermedia Systems, pp. 40–49, Pittsburgh (1998)
52. Marshall, C.C., Brush, A.B.: Exploring the relationship between personal and public annotations. Joint Conference on Digital Libraries **7**(11), 349–357 (2004)
53. McNally, K., O'Mahony, M.P., Smyth, B., Coyle, M., Briggs, P.: Towards a reputation-based model of social web search. In: IUI '10: Proceeding of the 14th International Conference on Intelligent User Interfaces, IUI '10, pp. 179–188. ACM, Hong Kong (2010)

54. Mika, P.: Ontologies are us: a unified model of social networks and semantics. Web Semant Sci Serv Agents World Wide Web **5**(1), 5–15 (2007)
55. Miller, P.: Coming together around Library 2.0. A focus for discussion and a call to arms. D-Lib Magazine **12**(4), 2006 (2006)
56. Mislove, A., Koppula, H.S., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Growth of the flickr social network. In: OSN '08: Proceedings of the First Workshop on Online Social Networks, WOSN '08, pp. 25–30. ACM, Seattle (2008)
57. Murley, D.: Technology for everyone... What's all the fuss about Library 2.0?. Law Libr. J. **100**(1), 197 (2008)
58. Nov, O., Naaman, M., Ye, C.: What drives content tagging: the case of photos on flickr. In: CHI '08: Proceeding of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems, pp. 1097–1100. ACM, New York (2008)
59. O'Donovan, J., Smyth, B.: Trust in recommender systems. In: Proceedings of the 10th International Conference on Intelligent User Interfaces, pp. 167–174, San Diego (2005)
60. O'Hara, K., Sellen, A.: A comparison of reading paper and on-line documents. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 335–342. ACM, New York (1997)
61. O'Mahony, M., Smyth, B.: A classification-based review recommender. Knowl.-Based Syst. **23**(4), 323–329 (2010)
62. O'Toole, C., Smeaton, A.F., Murphy, N., Marlow, S.: Evaluation of automatic shot boundary detection on a large video test suite. In: CIR'99 - The Challenge of Image Retrieval: 2nd UK Conference on Image Retrieval, Manchester (1999)
63. Ovsiannikov, I., Arbib, M., McNeill, T.: Annotation technology. Int. J. Hum.-Comput. Stud. **50**(4), 329–362 (1999)
64. Paolillo, J., Penumarthy, S.: The social structure of tagging internet video on del.icio.us. In: IEEE International Conference on System Sciences, vol. 40, p. 1414. IEEE, Hawaii (2007)
65. Pomerantz, J., Abbas, J., Mostafa, J.: Teaching digital library concepts using digital library applications. Int. J. Digit. Libr. **10**(1), 1–13 (2009)
66. Sabater, J., Sierra, C.: Review on computational trust and reputation models. Artif. Intell. Rev. **24**(1), 33–60 (2005)
67. Sadlier, D., O'Connor, N.: Event detection in field sports video using audio-visual features and a support vector machine. IEEE Trans Circuits Syst Video Technol **15**(10), 1225 (2005)
68. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors. In: WWW '10: Proceeding of the 19th International Conference on World Wide Web. Raleigh (2010)
69. Salton, G., McGill, M.: Introduction to Modern Information Retrieval. McGraw-Hill Inc., New York (1986)
70. Sannomiya, T., Amagasa, T., Yoshikawa, M., Uemura, S.: A Framework for Sharing Personal Annotations on Web Resources Using XML. In: Proceeding of the Workshop on Information Technology for Virtual Enterprises ITVE 2001. pp. 40–48, Gold Coast (Dec 2000)
71. Schilit, B.N., Golovchinsky, G., Price, M.N.: Beyond paper: supporting active reading with free-form digital ink annotations. In: CHI '98: Proceedings of the 16th SIGCHI Conference on Human Factors in Computing Systems in Computing Systems, pp. 249–256. Los Angeles (1998)
72. Schroeter, R., Hunter, J., Kosovic, D.: Vannotea-A collaborative video indexing, annotation and discussion system for broadband networks. In: Knowledge Markup and Semantic Annotation Workshop, K-CAP 2003. Sanibel Island (2003)
73. Shamma, D.A., Kennedy, L., Churchill, E.F.: Tweet the debates: understanding community annotation of uncollected sources. In: WSM '09: Proceedings of the First SIGMM Workshop on Social Media, pp. 3–10. ACM, Beijing (2009)
74. Shipman, F., Price, M., Marshall, C., Golovchinsky, G., Schilit, B.: Identifying useful passages in documents based on annotation patterns. In: Proceedings of the 7th European Conference on Digital Libraries, vol. 2769, pp. 101–112. Trondheim, 17-22 Aug (2003)
75. Smeaton, A.F., Murphy, N., O'Connor, N.E., Marlow, S., Lee, H., McDonald, K., Browne, P., Ye, J.: The Físchlár digital video system: a digital library of broadcast TV programmes. In: JCDL '01: Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '01, pp. 312–313. ACM, Roanoke (2001)
76. Smyth, B., Freyne, J., Coyle, M., Briggs, P., Balfe, E.: I-SPY-anonymous, community-based personalization by collaborative metasearch. In: Proceedings of the 23rd SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp. 367–380. Springer, Cambridge (2003)
77. Talis. Project whisper (Feb 2010)
78. Tannenbaum, P., Noah, J.: Sportugese: a study of sports page communication. Journal. Q. **36**(2), 163–170 (1959)
79. Terveen, L., Hill, W., Amento, B., McDonald, D., Creter, J.: PHOAKS: A system for sharing recommendations. Commun. ACM **40**(3), 59–62 (1997)
80. Van House, N.: Trust and epistemic communities in biodiversity data sharing. In: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 231–239. Portland (July 2002)
81. Wann, D., Metcalf, L., Adcock, M., Choi, C., Dallas, M., Slaton, E.: Language of Sport Fans: Sportugese Revisited. Percept. Mot. Skills **85**(3 Pt 1), 1107–1110 (1997)
82. Wilensky, R.: Digital Library Resources as a Basis for Collaborative Work. J. Am. Soc. Inf. Sci. **51**(3), 228–245 (2000)
83. Windley, P.J., Daley, D., Cutler, B., Tew, K.: Using Reputation to Augment Explicit Authorization. In: DIM '07: Proceedings of the 2007 ACM Workshop on Digital Identity Management, pp. 72–81. Fairfax (2007)
84. Wolfe, J.L.: Effects of Annotations on Student Readers and Writers. In: DL '00: Proceedings of the 5th ACM Conference on Digital Libraries, pp. 19–26, San Antonio (2000)
85. Wu, X., Zhang, L., Yu, Y.: Exploring social annotations for the semantic web. In: WWW '06: Proceedings of the 15th International Conference on World Wide Web, pp. 417–426. ACM, New York (2006)
86. Xi, W., Lind, J., Brill, E.: Learning Effective Ranking Functions For Newsgroup Search. In: Proceedings of the 27th Annual International Conference on Research and Developement in Information Retrieval, pp. 394–401. ACM Press, New York (2004)
87. Zacharia, G., Moukas, A., Maes, P.: Collaborative Reputation Mechanisms for Electronic Marketplaces. Decis. Support Syst. **29**(4), 371–388 (2000)
88. Zheng, Q., Booth, K., McGrenere, J.: Co-Authoring with structured annotations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 131–140, Montréal (2006)
89. Zhu, X., Gauch, S.: Incorporating quality metrics in centralized/distributed information retrieval on the world wide web. In: SIGIR '00: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 288–295, Athens (2000)
90. Zimmermann, P.: PGP User's Guide, Volume I: Essential Topics. Available on the WWW via ftp://ftp.pgpi.org/pub/pgp/2.x/doc/pgpdoc1.txt (Oct 1994)