International Conference
FROM DELC TO VELSPACE
Plovdiv, 26–28 March 2014

Plenary talk

UNIVERSAL DIGITAL LIBRARY

Awatef Rahuma, Hussein Zedan

1. INTRODUCTION

Learning is the acquisition of knowledge or skill by either instruction and/or self-study. Most commonly accepted definition of learning is that it refers to a change in behaviour that is due to experience. This is a very basic functional definition of learning where learning is seen as a function that maps experience onto behaviour, i.e. the effect of experience gained on behaviour(s).

A crucial dimension to learning is *regularity* in the learner's environment. The notion of regularity encompasses all states in the learner's environment that entail more than the presence of a single stimulus or behaviour at a single moment in time. It can thus refer to the presence of a single stimulus or behaviour at multiple moments in time and the presence of multiple stimuli or behaviours at a single moment in time (as in one-trial learning).

Libraries have long been one of the primary sources used by instructors and learners to search and obtain learning resources. Traditionally, the main mission of any library has been to provide the infrastructure aimed at supporting the creation, assimilation and leverage of knowledge.

The notion of a Digital Library (DL) started as a system for providing access to digitized books and other learning resources. Now a Digital Library is thought of as a tool at the centre of intellectual activity having no logical, conceptual, physical, temporal, or personal borders or barriers to information. It has shifted from a *content-centric* system that merely supports the organization and provision of access to particular collections of data and information, to a *person-centric* system that delivers innovative, evolving, and personalized services to users.

The way for conventional and/or digital libraries to support learning includes the improvement of reading rooms, the quality of reference services and the general availability of books themselves.

Any library should therefore

- identify usable learning resources and services, and
- provide learning materials and seamless, integrated access to a range of resources across boundaries of media. This provision is done in a *controlled* fashion.

For either types of libraries, there are fundamental questions that need to be answered:

- What learning resources and services should be covered?
- How to organise and present learning resources and services?
- How can the library accommodate learning theories?

This paper explores the evolution of libraries and argues that with the advent of the Internet, traditional and digital libraries are no longer fit-for-purpose and that a *Universal* library, with an ever increasing learning resources, has emerged. However this Universal library, whilst its massive popularity and usefulness for both learner and instructors, faces many challenges. We focus on one of these challenges, namely, how to improve the accuracy of information retrieval under the assumption that an *object* (or a resource) in the Universal Library is *continually* tagged:

- with 0 or more tags (in some system, up to 75 tags are allowed per an object),
- by more than one *subject* (user)
- using different languages or notations (ie abbreviations – *lol, omg, etc* and
- a tag can be removed at any time by its creator.

2. DIGITAL (E–) LIBRARY EVOLUTION

Digital Libraries represent the linkage of many disciplines and fields, including *data management*, *information retrieval*, *library sciences*, *document management*, *information systems*, *the web*, *image processing*, *artificial intelligence*, *human-computer interaction*, and *digital curation*.

Conceptions of the role of Digital Libraries have shifted from static storage and retrieval of information to facilitation of communication, collaboration, and other forms of dynamic interaction among scientists, researchers, or the general public on themes that are pertinent to the information stored in the Digital Library. Moreover, expectations of the capabilities of Digital Libraries have evolved from handling mostly centrally located text to synthesizing distributed multimedia document collections, sensor data, mobile information, and pervasive computing services.

The variety of conceptions of what a Digital Library is has had a substantive impact on attempts to define and bound the term 'Digital Library'. Since 2006 the term has been generally used to refer to systems that are heterogeneous in scope and provide diverse types of functionality.

This overloading of the term results in Digital Library services and systems that do not deliver interoperability and reuse of content and technologies.

A Digital Library is an evolving organization that comes to existence through a series of development steps that bring together all necessary constituents.

In Figure 1, this process is presented, where

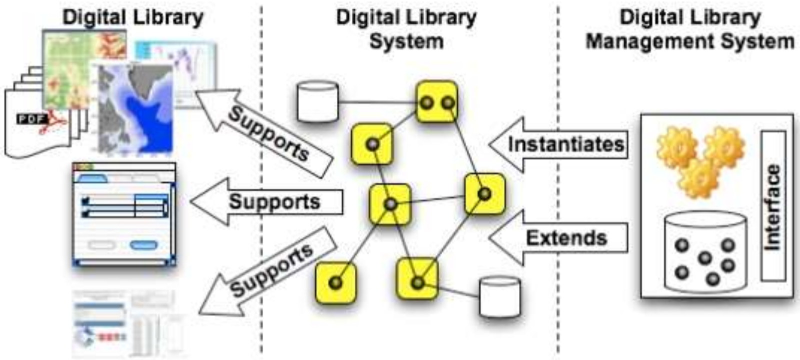

Figure 1. Digital Library Structure

- Digital Library – A possibly virtual organization that comprehensively collects, manages, and preserves for the long term rich digital content, and offers to its user communities specialized functionality on that content, of measurable quality and according to codified policies.
- Digital Library System – A software system that is based on a defined (possibly distributed) architecture and provides all functionality required by a particular Digital Library. Users interact with a Digital Library through the corresponding Digital Library System.
- Digital Library Management System – A generic software system that provides the appropriate software infrastructure both (i) to produce and administer a Digital Library System incorporating the suite of functionality considered foundational for Digital Libraries and (ii) to integrate additional software offering more refined, specialized, or advanced functionality. It belongs to one of the following three types:
  - Extensible Digital Library System – A complete Digital Library System that is fully operational with respect to a defined core suite of functionality. DLs are constructed by instantiating the DLMS and thus obtaining the DLS.
  - Digital Library System Warehouse – A collection of software components that encapsulate the core suite of DL functionality and a set of tools that can be used to combine these components in a variety of ways to create Digital Library Systems offering a tailored integration of functionalities. New software components can easily be incorporated into the Warehouse for subsequent combination with those already there.
  - Digital Library System Generator – A highly parameterized software system that encapsulates templates covering a broad range of functionalities, including a defined core suite of DL functionality as well as any advanced functionality that has been deemed appropriate to meet the needs of the specific application domain. Through an initialization session, the appropriate parameters are set and configured; at the end of

that session, an application is automatically generated, and this constitutes the Digital Library System ready for installation and deployment.

Although the concept of Digital Library is intended to capture an abstract system that consists of both physical and virtual components, the Digital Library System and the Digital Library Management System capture concrete software systems. For every Digital Library, there is a unique Digital Library System in operation (possibly consisting of many interconnected smaller Digital Library Systems), whereas all Digital Library Systems are based on a handful of Digital Library Management Systems.

Despite the seeming richness and diversity of existing digital libraries, there is only a small number of core concepts defined by all systems. These concepts are identifiable in nearly every Digital Library currently in use. There are six core concepts provide a foundation for Digital Libraries. Five of them appear in the definition of Digital Library: Content, User, Functionality, Quality, and Policy; the sixth one emerges in the definition of Digital Library System: Architecture. All six concepts influence the Digital Library framework.

The organization of resources and the way they are identified and retrieved follow exactly the same mechanisms and policies as in a typical traditional library. This is done under the *Content* concept which encompasses the data and information that the Digital Library handles and makes available to its users. It is composed of a set of information objects organised in collections. Content is an umbrella concept used to aggregate all forms of information objects that a Digital Library collects, manages, and delivers, and includes primary objects, annotations, and metadata. For example, metadata have a central role in the handling and use of information objects, as they provide information critical to its syntactic, semantic, and contextual interpretation.

3. BEYOND DIGITAL LIBRARY: THE UNIVERSAL LIBRARY

The availability of internet-enabled devices, such as mobile phones, desktops, laptops, tablets, mini tablets, PDAs, set-top boxes and smart TVs, has facilitated the access to upload and modify content, images, video and audio media. Here are some statistics which show

- An estimated 7 Petabytes is the amount of photo content added to Facebook every day;
  - By the middle of 2011, an estimated 100 billion – photos were hosted;
  - During 2012, 300 million – new photos were added every day;
- In 2012, 4.5 million – photos uploaded to Flickr each day;
  - In 2011, 6 billion – photos were hosted by them;
- 5 billion – is the total number of photos uploaded to Instagram since its start Sept 2012
  - 58 – photos are uploaded every second.

Similar statistics can be found for other types of media – audio and textual. Therefore, the Internet has become a universal source of rich learning resources. We call this Universal Library that almost renders obsolete the traditional libraries (albeit physical or virtual).

Nonetheless, we should observe the following

- Whilst we can imposed structure on traditional libraries, the Universal Library is, on the whole, unstructured and un-organised;
- Each object in the Universal Library is tagged (including empty tags).
- The un-structured nature of the Universal Library gives its popularity and should be maintained and preserved.
- The growth of a Traditional library is limited, bounded and finite in nature, the Universal Library is almost unbounded and unlimited.
- We can argue that, from an educational theorist view point, the globalisation and the scale of the Universal Library should greatly benefit the learning process.

The inherent structure in the traditional libraries, gives it the advantages of precise searching. The inherent difficulty of the Universal Library is in its imprecise search.

Therefore, the key challenge is *how can the accuracy and efficiency of search and retrieval be improved without too many restrictions*? In what follows, we will attempt to shed some light on such a challenge.

4. TAGS AND THEIR CHALLENGES

A tag is a word chosen by the subject to identify the object. Such presents many challenges, including:

The influence of the users' culture: Ethnicity and cultural differences guide perception and cognition differently. For example, an analysis of image tags created by European American and Chinese participants concluded that whereas Westerners focus more on foreground objects, the Easterners have a more holistic way of viewing images early on. This was discovered through the analysis of tag assignment order. For Easterners, the specificity of tags increased from holistic scene description to individual objects. On the other hand, the tags given by the Westerners focused on individual objects first and then on overall scene content.

The influence of Motivation: Motivations, probably, forms a major influence on the usability of tags for all purposes. Tags that arise from the need of future retrieval and contribution, particularly for the benefit of external audience, are likely to be visually more relevant compared to tags used for personal references. Images that are annotated and shared within special interest groups are very likely to be specifically annotated and heavily monitored. They would also be heavily influenced by the motivation of the interest group.

The Users' Domain knowledge: Some users who tag their images with non-understandable words, characters, personal references or numeric symbols, can only be thought of as doing so because they have a praticular knowledge about the domain that caused them to save and annotate the images in the first place. Such tags have no use or meaning to the wider audience, and should be filtered out, so as not to affect usage statistics.

The issue of Semantic loss: An annotator in folksonomies is not obliged to associate all relevant tags with an image, leading to semantic loss in the textual descriptions. The batch-tag option provided by most photo sharing sites adds to this problem by allowing users to annotate an entire collection of photos with a set of common tags. Even if such tags are potentially useful to provide a broad personal context, they cannot be used to identify image-level differences, thus leading to semantic loss. One consequence of this fact is that the absence of a tag from an image description cannot be used to confirm the absence of the concept in that image. Hence, such images cannot be directly used as negative examples for training.

The issue of Vocabulary: The spontaneous choice of words to describe the same content varies among different people, and the probability of two users using the same term is very little. Known as the vocabulary problem, this issue is often cited as a common characteristic of folksonomic annotations. The different word choices introduce problems of polysemy (one word with multiple meanings), synonymy (different words with similar meanings) and basic level variation (use of general versus specialized terms to refer to the same concept).

Operational Challenges: In addition to these challenges, *collaborative tagging* imposes its own addition challenge. An *object* (or a resource) in a collaborative tagging environment, is *continually* tagged:

- with 0 or more tags (in some system, up to 75 tags are allowed per an object),
- by more than one *subject* (user),
- using different languages or notations, i.e. abbreviations – *lol, omg, etc.*, and that
- a tag can be removed at any time by its creator.

5. SIMILARITY MEASURES

The similarity between two images can be characterised as follows [6].

- The similarity between two images A and B is related to their commonality. The more commonality of attributes they share, the more similar the two images are.
- The similarity between two images A and B is related to the differences between them. The more differences their attributes have, the less similar the two images are.
- The maximum similarity between two images A and B can only be reached when A and B are identical, no matter how much commonality they share.

There are many similarity measures that can be employed to deal with the many challenges that collaborative tagging presents. Jaccard coefficient [2] can be used to work out the similarity between images based on user's tags. The coefficients are used to normalise the co-occurrence between two tags. The Jaccard coefficient, sometimes referred to as the "Jaccard similarity coefficient", can be defined as a *statistic* used for comparing the similarity and diversity of sample sets. That is, given two objects, X_1 and X_2, each with n binary attributes, the Jaccard coefficient is a useful measure of the overlap that X_1 and X_2 share with their attributes. Each attribute of X_1 and X_2 can either be 0 or 1.

$$\sigma(X_1, X_2) = \frac{|X_1 \cap X_2|}{|X_1 \cup X_2|}$$

For example if we consider the following attributes for a fruit: *Sphere, sweet, sour* and *crunchy*. Then, an *Apple* (X_1) and a *Banana* is represented as

$$Apple = \{1,1,1,1\} \text{ and } |Apple| = 4,$$

$$Banana = \{0,1,0,0\} \text{ and } |Banana| = 4.$$

Here we have

$$Apple \cup Banana = \{1,0\} \quad \text{with } |Apple \cup Banana| = 2, \text{ and}$$

$$Apple \cap Banana = \{1\} \quad \text{with } |Apple \cap Banana| = 1.$$

$$\sigma(Apple, Banana) = \frac{|Apple \cap Banana|}{|Apple \cup Banana|} = 0.5.$$

The resulting similarity between two images ranges from –1 meaning the images are exactly opposite, to 1 meaning the two images are exactly the same, with 0 usually indicating independence, and in-between values indicating intermediate similarity or dissimilarity.

For example, if we search for a photo of a Barking Dog using Flickr, Figure 5:

Figure 2. Inaccurate search of barking dog

We can add a 3-eld structure to the tag, namely, object, action and background. This would be then compared with similar words stored in the Word-Net database. By applying Jaccard similarity measure over the returned list of words, will give back a similarity score between 0 and 1.

Applying this to our barking dog yields a much better results, in Figure 3

Figure 3. Examples of Accurate search for a barking dog

6. THE FRAMEWORK

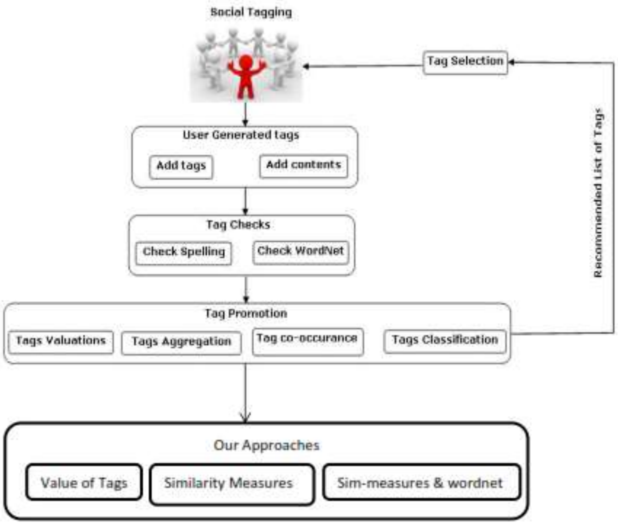

Figure 4. General Approach

Our model of image ranking [6] can further be detailed by the following chart, Figure 4, where web users tag images with semantically related words, such as “Jelly Bean” together with “Android”. Within a large photo sharing social website containing numerous independent users, such as Flickr, the semantic relationship can be captured and utilised. However, this method alone is not sufficient to all the relationships between the tags such as “window” in the photo of a “house”. The photos containing both “house” and particular style of “window” may be tagged as "house" only. Such an issue can be solved by applying tag visual correlation to measure the tags visual similarity. These two methods of correlations only use the relation between tags, which can be combined in the Rankboost framework [8, 1], which in turn uses the order of instances rather than the absolute distance.

The tag recommendation process can be explained by an example, where a selected photo with user-defined tags and an ordered list of candidate tags is derived for each of the user-defined tags, based on tag co-occurrence. The lists of candidate tags are then used as input for tag aggregation and ranking, which ultimately produces the ranked list of recommended tags. For example, the photo of Sacré-Coeur Figure 5 may have two user-defined tags, namely Sacré-Coeur and Paris. Using Tag Co-occurrence, a list of co-occurring tags (church, architecture, montmarte, seine, Europe, travel and night) is derived Figure 5. They have some tags in common, such as France and Paris. After aggregation and ranking four tags are recommended: *Paris*, *Church*, *Architecture* and *France*. The actual number of tags being recommended should, of course, depend on the relevancy of the tags, as we will see in the example case of using the 'value of tags'.

Tag co-occurrence is the pillar that the tag recommendation approach is built upon, and as a consequence, only works reliably when a large quantity of supporting data can be captured and accessed [7]. Fortunately, the amount of user-generated content that is created by Flickr users, satisfies this demand and provides the collective knowledge base that is needed to make tag recommendation systems work in practice. There exists various methods to calculate co-occurrence coefficients between two tags. The co-occurrence between two tags is defined as the number of photos, in our collection, where both tags are used in the same annotation.

Using the raw tag co-occurrence for computing the quality of the relationship between two tags is not very meaningful, as these values do not take the frequency of the individual tags into account. Therefore it is common to normalise the co-occurrence count with the overall frequency of the tags. There are essentially two different normalisation methods: symmetric and asymmetric.

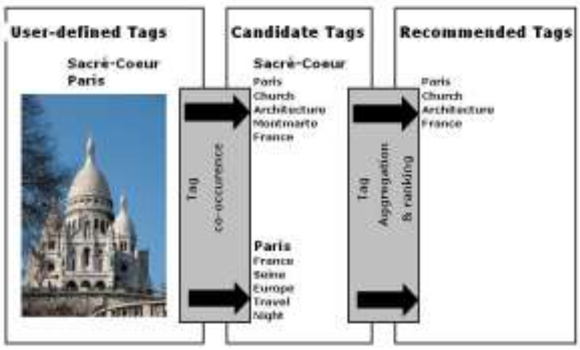

Figure 5. The tag recommendation process

Symmetric measures:

We use the Jaccard coefficient, introduced in section 2, to normalise the co-occurrence of two tags t_i and t_j by calculating:

$$J(t_i, t_j) := \frac{|t_i \cap t_j|}{|t_i \cup t_j|}$$

The coefficient takes the number of intersections between the two tags, divided by the union of the two tags. The Jaccard coefficient is known to be useful to measure the similarity between two objects or sets. In general, we can use symmetric measures, like Jaccard, to deduce whether two tags have a similar meaning.

Alternatively, tag co-occurrence can be normalised using the frequency of one of the tags. We can use the equation:

$$P(t_j | t_j) := \frac{|t_i \cap t_j|}{|t_j|}.$$

The equation captures how often the tag t_t co-occurs with tag t_j normalised by the total frequency of tag t_i. This can be interpreted as the probability of a photo being annotated with tag t_j given that it was annotated with tag t_i. Many other variations of asymmetric co-occurrence measure have been proposed in the literature before to build tag (or term) hierarchies.

To illustrate the difference between symmetric and asymmetric co-occurrence measures consider the tag Eiffel Tower. For the symmetric measure we find that the most co-occurring tags are (in order): Tour Eiffel, Eiffel, Seine, La Tour Eiffel and Paris. When using the asymmetric measure the most co-occurring tags are (in order): Paris, France, Tour Eiffel, Eiffel and Europe.

It shows that the Jaccard symmetric coefficient is good at identifying equivalent tags, like Tour Eiffel, Eiffel, and La Tour Eiffel, or picking up a close by landmark such as the Seine. Based on this observation, it is more likely that asymmetric tag co-occurrence will provide a more suitable diversity of candidate tags than its symmetric opponent.

The next step in the process of tag aggregation is to merge the known lists of candidate tags for each of the user-defined tags, into a single ranking. There are two aggregation methods, based on voting (a strategy that computes a score for each candidate tag) and summing (a strategy that takes the union of all candidate tag lists) [7] that can be used along with a re-ranking procedure (where tags are arranged in their order of high relatedness [3]) that promotes candidate tags containing certain properties and significance values.

To achieve this, we use three different types of tags:

- User-defined tags *U* refer to the set of tags that the user assigned to a photo.

- Candidate tags C_u is the ranked list with the top most co-occurring tags, for a user-defined tag $u \varepsilon U$. We denote *C* to refer to the union of all candidate tags for each user-defined tag $u \varepsilon U$.
- Recommended tags R is the ranked list of the most relevant tags produced by the tag recommendation system.

For a given set of candidate tags (C) a tag aggregation step is needed to produce the final list of recommended tags (R), whenever there is more than one user-defined tag. In this section, we define two aggregation strategies. One strategy is based on voting (a strategy that computes a score for each candidate tag), and does not take the co-occurrence values of the candidate tags into account, while the summing strategy (which takes the union of all candidate tag lists) [7] uses the co-occurrence values to produce the final ranking. In both cases, we apply the strategy to the top co-occurring (or highly related) tags in the list.

Another method of increasing the accuracy of image tags is to expand the three fields used to four fields (or parameters), namely 'primary object', 'secondary object', 'action' and 'primary colour'. However, in this case, each potential tag information received will first be assessed for its value. This is also referred to as Value of Information or Value of tags.

As an example of the implementation of the 4 fields method, consider the search for a photo of a red sky at a lake. In normal circumstances, such a search may return the non-relevant image Figure 6, which shows a lake with red flowers but without the red sky at dusk.

Figure 6. Lake with red flowers

However, our enhanced method of tagging would allow users to enter extra object names to further identify the tag. In this case, the primary object would be 'lake', the secondary object would be 'sky', the action would be 'dusk', or more precisely, 'sunset', and finally the colour would be 'red'.

Figure 7. Lake with red sky at dusk

In this method, before the WordNet database is queried to check if the specified words stored in the tags and returned by the search do exist, each tag returned is 'valued' against a set of pre-defined criteria. Examples of this criteria are:

- Popularity: What is the size of the tag on the Flickr tag cloud, i.e. how many times has the tag been voted for?
- Topicality: Is the tag suitable for the search topic? As an example, consider a search for an image of the city of London. The returned tags may represent London City or the novelist Jack London. In this case, the results are compared to the categories on WordNet, where London city belongs to 'noun.location'. This category is ranked higher (as it has more tags per photo) than the London Novelist category 'noun.person'
- Uniqueness: Is the tag of the photo unique and unambiguous? For example, a photo of a 'car' which is also tagged 'car' is unique and can only refer to a car, irrespective of its type.
- Redundancy: Are there too many irrelevant and redundant tags? For example, a search for a photo of a cat that returns 'cat', 'feline', 'tabby', 'fluff', 'jinx' (for a photo of a black cat) and 'cuddles' is, obviously, plagued by too many redundant tags, when 'cat' or 'feline' would suffice.
- Simplicity: How simple is a photo tag? For example, a photo of a Teapot that is tagged 'Teapot for brewing Darjeeling tea' may be too complex for search engines, as well as tag rankings algorithms (and the word Darjeeling may also be classified as spam). Ideally, the photo should be tagged as, simply, 'Teapot'.
- Spelling: Misspelled tags should, obviously, be excluded from the list of returned tags.
- Recency: For this assessment, tags are ranked by age, such that an image that has several possible tags, which were created over a long period of time, would rank the most recent tags higher than the oldest ones.

The returned list of tags is deemed to be much more accurate in terms of the search query, and this can be used to more accurately return the image in Figure 7, which represents exactly the criteria being searched for i.e. a lake with a red sky.

There are many other tag criteria that can be used to assess returned tag values. However, the criteria of topicality and relevance is of more importance as it answers the question “What are users tagging?” This criteria is mapped to WordNet categories, which are used to bind tags to the category with the highest ranking. Figure 8 shows the distribution of Flickr tags over the most common WordNet categories, which can be used to assess and classify tags. When focussing on the set of classified tags, we find that locations are tagged most frequent (28 %); followed by artefacts or objects (16 %), people or groups (13 %), actions or events (9 %), and, finally, time (7 %).

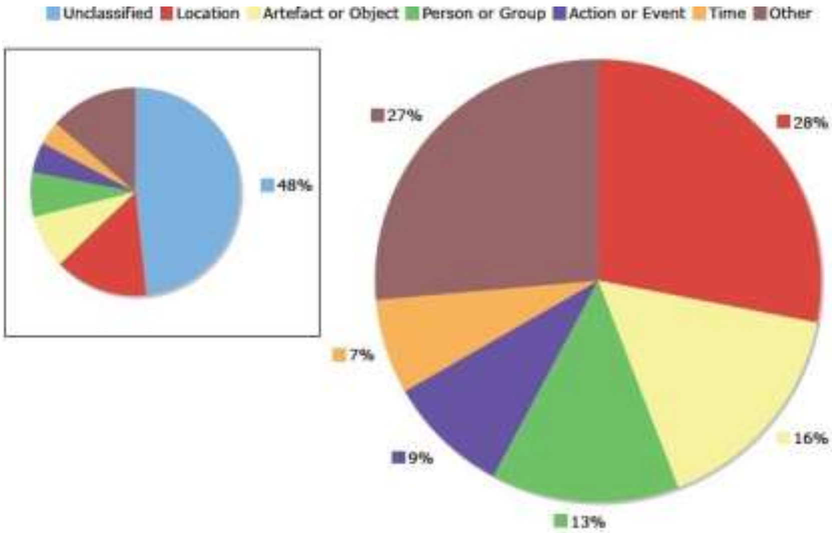

Figure 8. Flickr’s tags Most frequent WordNet categories

From this information, we can conclude that users do not only tag the visual contents of the photo, but to a large extent provide a broader context in which the photo was taken, such as, location, time, and actions.

Another criteria that would be used to rank photo tags, is the classification of tags as defined in Table 1, which looks at classes of photos with one tag, photos with 2–3 tags, 4–6 tags, and more than 6 tags, respectively. The table can be used to compare voting strategies (i.e. photos with a high number of user tags) against summation strategies (photos with aggregated tags). Experiments have shown that the increases in the accuracy of tag ranking is proportional to the number of a photo’s user-defined tags [5]. This indicated that only 13 % of all tagged photos have a higher degree of accuracy as they contain more than six tags. The high number of tags will serve as an input into the Jaccard measure of co-occurrence, as we will discuss in the Example section.

	Tags per Photo	Photo %
Class I	1	31 %
Class II	2 – 3	33 %
Class III	4 – 6	23 %
Class IV	>6	13 %

Table 1. Definition of photo-tag classes and the percentage of photos in each class}

Finally, below we will introduce an example that details how to increase the accuracy of tagging by employing n-dimension of resources, i.e. as many of the above methods as possible.

Summary

Our research work revolves around the improvements of image tagging, and for this reason, we have opted to combine many of the methods discussed in the previous sections. Users will be able to enter tags based on two searchable objects, as well as the photos action and background. This will significantly enhance the value added to the photo tags.

Once the user defined tags are saved with the photos, the returned list of tags, from a search query, will be enhanced by comparing it against a set pre-defined values (or criteria). Such an action would serve to filter out many irrelevant results. The returned list would be further enhanced by promoting the tags via the use of tag classes that utilise voting strategies.

The final filtered list of tags would then be used as recommended tags for users to choose from, as this would reduce the introduction of irrelevant tags that can be entered due to misspellings, inaccurate descriptions and attempted spamming. Users would then select one or more tags from this pre-defined list, without the ability to enter free text.

Once we get a large number of photos that have been tagged by a promoted and recommended set of tags, the set of results returned by a search query would be highly accurate. This would allow us to accurately compare similarity measures between photos using the Jaccard coefficient.

7. EXAMPLE

The example we will use is a photo of Big Ben's tower. Initially, we allow users to enter their tags into the four fields described in the previous sections; namely primary object, secondary object, action and colour. However, before the tags can be added, we use the Jaccard method to calculate co-occurrence coefficients. Both normalisation methods; symmetric and asymmetric will be used for the calculations.

For the primary object, we use the symmetric measure to find the most co-occurring tags which returns (in order): *Big Ben, Big Ben Tower, Westminster, Thames, London and England*. These recommendations will be offered to the users to populate the primary object field. Next, we use the asymmetric measure to calculate the most co-occurring tags for the secondary object which returns (in order): *London, England, Clock, Tower, Westminster, Architecture and Europe*. It is more likely that asymmetric tag co-occurrence will provide a more suitable diversity of candidate tags than its symmetric opponent. Therefore, it is more useful for returning the secondary object's recommended list. Similarly, the co-occurring tags for action would return: *Travel, Tour, Visit and Book*. Finally, the colours returned are: *Blue, Black and White*.

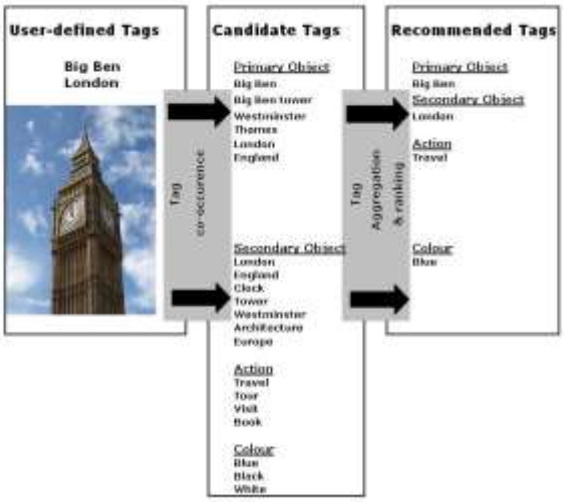

Figure 9. Big Ben's Tag Recommendation

In Figure 9, the list of tags produced by the symmetric and asymmetric measures for each of the four fields are further aggregated to produce the final list of recommended tags. We use two aggregation strategies. One strategy is based on voting, and does not take the co-occurrence values of the candidate tags into account, while the summing strategy uses the co-occurrence values to produce the final ranking. In both cases, we applied the strategy to the top co-occurring tags in the list.

The voting strategy computes a score for each candidate tag, where a vote for that candidate is cast. A list of recommended tags is obtained by sorting the candidate tags on the number of votes. The summing strategy also takes the union of all candidate tag lists, and sums over the co-occurrence values of the tags.

Figure 8 showed that users do not only tag the visual contents of the photo, but to a large extent provide a broader context in which the photo was taken, such as, location, time, and actions. The tags being recommended, by our above strategy, and accepted by our users can now be analysed based on vote aggregation (summing) and promotion (voting). At first, we can see that the largest (most frequent) category in the Figure is 'Unclassified' at 48 %.

WordNet	Acceptance ratio %
Unclassified	39 %
Location	71 %
Artifact or Object	61 %
Person or Group	33 %
Action or Event	51 %
Time	46%
Other	53 %

Table 2. Acceptance ratio of tags of different WordNet categories

However, when voting is taken into account, where users select one or more of the tags recommended by our strategies, we can deduce that there exists a gap between user-defined and accepted tags for those tags which can not be classified using WordNet.

Table 2 shows the acceptance ratio for different WordNet categories. In the Table we can see that locations, artifacts, and objects have a relatively high acceptance ratio. However, people, groups and unclassified tags (tags that do not appear in WordNet) have relatively low acceptance ratio. We conclude that our system is particularly good at recommending additional location, artifact, and object tags.

REFERENCES

[1] Iyer, Y., R. Schapire; Freund and Y. Singer, An efficient boosting algorithm for combining preferences, *Proc. of the 15th Intl. Conference on Machine Learning*, 1998.

[2] Jaccard, P., Otude comparative de la distribuition orale dans une portion des alpes et des jura.

[3] Seongjae, L. and C. Soosun, Web image retrieval re-ranking with wikipedia semantics, *International Journal of Multimedia and Ubiquitous Engineering*, 2012.

[4] Pagano, P., C. Thanos, Y. Ioannidis, G. Koutrika, S. Ross, H. Schek, H. Schuldt, L. Candela and D. Castelli, Setting the foundations of digital libraries: The delos manifesto, *D-Lib Magazine*, Vol. 13, No. 3/4, 2007.

[5] Angus, E., M. Thelwall and D. Stuart, General patterns of tag usage among university groups in ickr, *Online Information Review*, 2008.

[6] Rahuma, A., *Semantically-enhanced image tagging system*, Ph.D. Thesis, De Montfort University, Leicester, UK, 2013.

[7] Brkur Sigurbjrnsson and Roelof van Zwol, Flickr tag recommendation based on collective knowledge, *Yahoo! Research*, 2008.

[8] Wu, L., L. Yang, N. Yu and X. Hua, Learning to tag, *Proceeding of ACM International World Wide Web Conference*, 2009.