

New search method in digital library image collections: A theoretical inquiry

Béla Lóránt Kovács¹, Margit Takács²

¹Affiliation: University of Debrecen, Faculty of Informatics

²Affiliation: University of Debrecen, Faculty of Informatics

Abstract: It is a challenge for today's library practice to make the entities in digital images retrievable according to the users demands. In our paper we wish to introduce a searching method which is able to seek certain elements and their environment in the images by applying a search within them in natural language. One key is to this method is extending the Dublin Core metadata system by a new element or by a classifier of the description element which help to enlist the natural language denominations of the elements in the images and the elements' positions within the image. Following this we are able to find certain elements then calculate and grade the information value of their environment with the help of a modified version of Shannon's entropy formula. The application of this method enables us to perform further value calculations for certain details within a particular image which make further, more efficient searches possible.

Key words: image collections, searching by image, metadata, Dublin Core, data structure, information, intelligence value, entropy

1. Introduction

Az elmúlt évtizedekben az elektronikus könyvtárak elterjedésével egyre nagyobb számban jelentek meg a képek a könyvtárakban (Eakins, 1999.). Ez egyaránt jelentette a könyvtárak szöveges adatbázisaiban felbukkanó képeket és a könyvtárak saját különálló képi gyűjteményeinek gyarapodását (Carson, 1996.). Mindkét esetben hasonló problémákkal kellett a könyvtárosoknak szembenézniük (Goodrum, 2000; Matusiak, 2006). Az első probléma az, hogy a képeken belül nehéz hagyományos könyvtári módszerekkel keresni. A második probléma az, hogy nehéz megfelelő adatszerkezetet találni a képi gyűjteményeknek. A harmadik probléma pedig az, hogy nehéz a hagyományos könyvtári módszerekkel leírni a képeket a gyűjteményeken belül (Yee et al., 2003.). Ez a három probléma valójában nagyon szoros összefüggésben áll egymással. Ahhoz ugyanis, hogy könyvtárak képi gyűjteményeiben keresni tudjunk először a megfelelő adatszerkezetre van szükségünk, az adatszerkezetet pedig nem hozhatjuk létre a megfelelő metaadatok hiányában.

Miért nehéz képi gyűjteményekben alkalmazni a könyvtári metaadat-rendszereket? Elsősorban azért, mert ezek a metaadatok elsősorban a dokumentumok szerzőjét, címét,

keletkezését és egyéb körülményeit írják le. Ezekkel azonban csak korlátozottan tudjuk leírni a dokumentum tartalmát. Annyira korlátozottan, hogy ez az említett metaadat-rendszereket alkalmatlanná teszi arra, hogy a képi gyűjteményeket természetes nyelvű kifejezések segítségével kereshetővé tegye. Valójában nincs is arra mód, hogy ezekben a rendszerekben jelenlegi formájukban a képek tartalmához természetes nyelvű kereső kifejezéseket rendeljünk (Li, 2010.). A könyvtárakban használt metaadat-rendszerek ugyanakkor kis módosításokkal alkalmassá tehetők arra, hogy a képi gyűjtemények is kereshetők legyenek.

Dolgozatunkban éppen ezért arra a kérdésre keressük a választ, hogy lehetséges-e természetes nyelvű keresést folytatni képi gyűjteményekben. A kérdés megválaszolásához pedig három lépésben kívánunk közelebb kerülni. Először megvizsgáljuk a szükséges metaadatokat, másodszor a szükséges adatszerkezetet és végül a kereséshez szükséges képleteket.

2. Metadata

In order to make it possible, first we have to examine metadata systems used in libraries. Among them Dublin Core (DC) (ISO 2009) is the most widespread and well-known in library world. The first 13 elements of DC were created in 1995. They were intended for describing textual sources available on internet. The DC has been changing since that time, newer and newer elements have been added to it. The DC element set was complemented with qualifiers which specify the meaning of each element (e.g. the date element can be qualified with date of production, publication and modification). Due to this DC can be better used for the cataloguing of library collections also. Later on DC was extended to non-textual, mainly image documents.

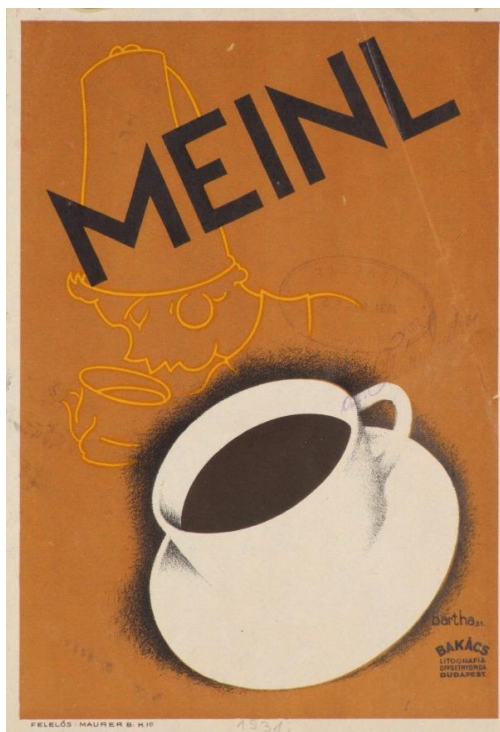
From the more significant 15 elements ISO standard was constructed which has already appeared in several countries, among them in Hungary too. The standard entitled *Information and documentation – Dublin Core metadata element set* became a description standard of information sources which extended beyond special fields. The electronic documents are typical information sources for Dublin Core applications. The standard determines only that element set which in general is used in relation to some task or application. Requirements and aspects applied at a given place and in a certain community can make additional restrictions, regulations and interpretations necessary. It is not objective of the standard to determine detailed criteria which the element set is used with for special tasks or applications (ISO 2009).

A special application like this retrieves entities which can be seen in traditional and digital images according to user needs in some collections. Therefore we have to apply a search method which finds each element in the images and applies a natural language search in its context. For this purpose those metadata is needed to be used which list elements in the images with their natural language names and positions.

Concerning Dublin Core we have three possibilities for subject description. We can describe topic of the information source with keywords, subject headings and classification numbers by using Subject element. Description element provides subject description which can be an abstract, a citation, table of contents, a reference to a graphical presentation of the content or a freely formulated description of the content. Third possibility for presenting the content is the Coverage element that means content or field of application of information source in space and time (extension). In the case of digital images the elements which are available in the standard are suitable for classifying each image into the appropriate category and describing events and happenings in the images and placing the content of the image or the image itself in space and time. Although all the three data can mean a lot in finding an image, but none of them can be used for searching each element in the images or searching its context.

To list entities in a given image the Description element can be best used among from the three above mentioned metadata. It is true that with the help of this element we can provide a brief content description, but with the introduction of the '*elements*' qualifier it is becoming suitable for listing entities in the image too. We can also give the position of certain entities by '*position*' qualifier. Due to the two qualifiers newly introduced the natural language search is already becoming possible.

A képi elemekre és információs értékükre történő keresés bemutatásához meg kell határozni a képeknek azt a csoportját, amelyben a keresés végbemegy. A képek jelen esetben a Debreceni Egyetem Elektronikus Archívumában (DEA) szereplő plakátok közül kerültek ki. Mivel egy adott tárggyal kapcsolatban végezzük a képi elemek keresését és vizsgáljuk az elemek információs értékét, ezért a plakátok kiválasztása a tárgyszó alapján történt. Ezek alapján az adatbázisunkban azok a plakátok kerültek be a DEA-ból, amelyeknek a metadatai között a *dc.subject* elemnél a kávé, mint tárgyszó volt megadva. Egy ilyen kép és a hozzá tartozó metaadatok a következőképpen szerepelnek az adatbázisunkban:



1. kép: Bartha: Meinl

DEA-ból származó DC metaadat

dc.date.accessioned
dc.date.available
dc.date.issued
dc.identifier.uri
dc.description.statementofresponsibility
dc.format.extent
dc.language
dc.publisher
dc.subject
dc.subject
dc.subject
dc.subject
dc.title
dc.subject.name
dc.subject.lcshhun

DEA-ból származó érték

2009-11-06T10:26:07Z
2009-11-06T10:26:07Z
1931
<http://hdl.handle.net/2437/89982>
Bartha
24 x 17 cm
Hun
Maurer B.
kereskedelmi plakát kereskedelem
Kávé
Reklám
plakát—grafikus
Meinl
Meinl, Julius (1824-1914)
Plakátok—Magyarország

dc.subject.lcshhun	Posters—Hungary
dc.publisher.place	Budapest
dc.format.extentpage	1 lap
dc.identifier.o	http://corvina.lib.unideb.hu:8082/WebPac/CorvinaWeb?action=onelong&showtype=longlong&idn
pac	o=bibDEK00709675
dc.identifier.bibid	bibDEK00709675
dc.format.color	grafikus, színes
dc.description.image	Fekete kávéval töltött fehér mokaálcscsészét látunk a képen. A háttér barna színű, rajta a sárga vonalakkal megrajzolt Meinel kávé reklámfigura, a fezt viselő szerecsen fiú körvonalai látszanak.
dc.subject.corporateName	Julius Meinel (Budapest)
dc.publisher.printing	Bakács Litográfia Offsetnyomda

Hozzáadott metaadatok és értékeik

Description.elements	sapka, szerecsen, fülbevaló, csésze, csészealj, kávé
Description.elements.positionx	sapka. (0,153; 0,46)
Description.elements.positiony	sapka. (0,05; 0,4)
Description.elements.positionx	szerecsen. (0,125; 0,7)
Description.elements.positiony	szerecsen. (0,37; 0,6375)
Description.elements.positionx	fülbevaló. (0,346; 0,4125)
Description.elements.positiony	fülbevaló. (0,391; 0,44)
Description.elements.positionx	csésze. (0,145; 0,326) (0,292; 0,805)
Description.elements.positiony	csésze. (0,497; 0,604) (0,5125; 0,8)
Description.elements.positionx	csészealj. (0,2875; 0,9)
Description.elements.positiony	csészealj. (0,578; 0,94)
Description.elements.positionx	kávé. (0,18; 0,32) (0,35; 0,686)
Description.elements.positiony	kávé. (0,524; 0,55) (0,5623; 0,725)

1. táblázat: Meinel plakát metaadatai

Adatbázisunkban minden plakátnál szerepel a DEA-ban megadott metaadat és a hozzátartozó érték, továbbá ezek a leírások kiegészülnek az általunk adott metaadatokkal is. A DEA-ban látható metaadatok az OPAC-ban lévő plakátleírásokból konvertálódnak. Azok az

adatok jelennek meg itt, amelyek a Debreceni Egyetem Egyetemi és Nemzeti Könyvtár (DEENK) online katalógusában megtalálhatóak. Természetesen vannak olyan metaadatok, amelyek csak a DEA-ban láthatóak. Ilyen a Dátum elemnél a csatlakozás és az elérhetőség ideje (*dc.date.accessioned*, *dc.date.available*), melyek elsősorban azt mutatják meg, hogy a plakátok digitalizált változata mely időponttól érhető el az adatbázisban. Ebből a szempontból fontos adat még az azonosító, ami mind a DEA-ban, mind az OPAC-ban lévő bibliográfiai leírásokat azonosítja (*dc.identifier.uri*, *dc.identifier.opac*, *dc.identifier.bibid*). Az URI a DEA által kiosztott azonosító, ami a digitalizált dokumentum és a hozzátartozó metaadatok azonosítója. Ezzel szemben az OPAC és a bibid azonosító az OPAC-ban található, az adott dokumentumról készült bibliográfiai rekord url címe és egyedi azonosítója.

A többi metaadat magáról a plakátról ad információt. Más dokumentumtípusokhoz hasonlóan a plakátok katalogizálásánál is az egyik legfontosabb adatelem a cím (*dc.title*), hiszen az ábrázolt képen kívül ennek az adatnak a segítségével válik azonosíthatóvá a mű. Ugyanakkor ez az egyik legkérdésesebb adat is, mivel a falragaszoknak nincs külön címük. Ezt megoldandó, általános eljárássá vált, hogy a bibliográfiai leírásnál a plakáton olvasható szöveget írják le címként. Ugyancsak fontos szerepe lehet a szerzőségi közlésnek is, bár a plakátokon csak elvéve szerepel szerzőségi közlésként leírható adat. A szerzőségi funkció meghatározásánál a könyvekhez hasonló módon járhatunk el. A mű készítője a szellemi tartalom létrehozója. Esetünkben szerzőként a Bartha családnév van megadva. Mivel ez a név az online katalógusban nem lett külön kiemelve és egységesítve (csak a cím és szerzőségi adatsorozatban szerepel), ezért a DEA-ban sem a *dc.author* metaadat azonosítja, hanem a *dc.description.statementsofresponsibility*. Ez utóbbi metaadat azt mutatja, hogy az OPAC-ban a 245. mező \$c almezőjében milyen nevek szerepelnek.

Képek esetében lényeges metaadat lehet a leírás (*dc.description.image*), amely a dokumentum tartalmának összegző leírása szabadszöveges formában. Nem szöveges dokumentumok esetében a méret is döntő lehet, melyet megadhatunk valamilyen mértékegységben vagy esetleg lapszámban (*dc.format.extent*, *dc.format.extentpage*). A DEA adatbázisban a Format elem egy minősítője mutatja (*dc.format.color*), hogy a plakát színes-e és hogy milyen az előállítás technikája.

Minden dokumentum esetén nagy szerepet kapnak a tárgyszavak, hiszen ezek nélkül a keresés nem lehetne eredményes. A *dc.subject* metaadathoz megadott tárgyszavak elsősorban a

plakátok típusaira és nagy témacsoportokra vonatkoznak. A DEENK által használt tárgyszavak egy része azonban a Library of Congress tárgyszórendszeréhez igazodik, melyek magyarul és angolul is szerepelnek a bibliográfiai leírásokban. A *dc.subject.lcshhun* metaadat tehát Library of Congress tárgyszórendszeréhez igazodó a dokumentum tartalmát leíró tárgyi kifejezés. Tárgyszóként jelennek meg azoknak a személyeknek és testületeknek az egységesített névalakja is, amelyek a dokumentum tartalmával kapcsolatba hozhatóak (*dc.subject.name*, *dc.subject.corporateName*).

Ezek a metaadatokon túl, a képi elemekre és információk értékükre történő kereséshez meg kell adni a képeken szereplő elemeket és azok pozícióit. A dokumentumon látható entitások megnevezésére a *dc.description.elements* metaadatot használhatjuk, ahol a képen ábrázolt elemek felsorolás formájában kerülnek megnevezésre. Itt felmerülnek olyan problémák, mint szinonimáké, a rész-egész viszonyoké stb., de ezek egy következő tanulmány keretében kerülnek megoldásra.

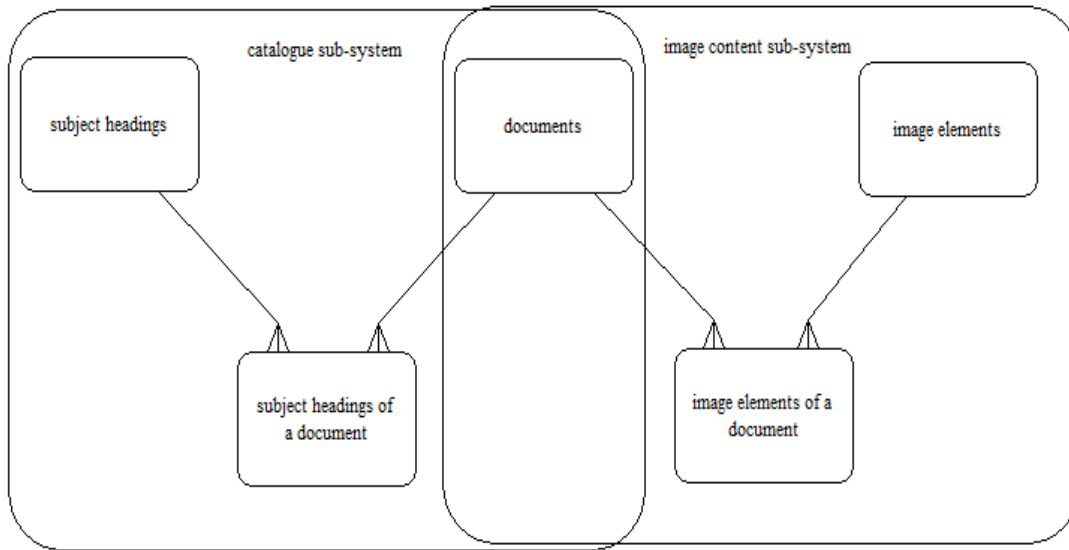
Ahhoz, hogy az elemek információk értékére tudjunk keresni, szükségessé válik az elemek pozíciójának megadása, melyet a *dc.description.elements.positionx* és *dc.description.elements.positiony* metaadatok segítségével valósíthatunk meg. Az első metaadat jelöli az adott elem x koordináta-tengely menti helyzetét, míg a második az y koordináta-tengely menti pozícióját adja meg. Mind a két tengely mentén a képek egységnyi hosszúságúak, a képen ábrázolt elemek helye az egységen belül került meghatározásra. Egy-egy elem helyét a koordináta-tengelyek mentén számpárok formájában adtuk meg, ahol az első érték az elem kezdő pozícióját, a második érték pedig záró pozícióját adja meg. Mivel nem határozhatunk meg metaadatot minden képen látható elemhez, ezért a számpár előtt kerül megnevezésre az az entitás, amihez hozzárendeltük az adott számpárt. Before presenting search for image elements and their information value it is also necessary to define the more important concepts that we will use during our presentation.

3. Data structures

At data storage and retrieval we will apply the following concepts: image document, image collection, image sub-collection, subject heading and image element. The *image document* is a physical carrier of information which users can perceive as a view with appropriate technical

devices. The *image collection* is a structured database which consists of image documents and their attributes (metadata). The image sub-collection is a part of image collection where image documents are available with at least one same subject heading. The *subject heading* is an attribute which is used for thematic grouping of image documents. The image element is a part of the image document. It can carry some kind of information for users and which we can supply with a natural language attribute. The newly introduced qualifiers of Dublin Core, i.e. the *dc.description.elements* and the *dc.description.position* can be applied for describing image elements and their positions. When we develop database structure of the search system, we will use these concepts.

We considered the relational data model (Codd 1969, Ullman 2008) to be the best one for the development of data structure of a search system. Our data model has three main components such as the catalogue sub-system, the image content sub-system and the image details information sub-system. We will obtain the *catalogue sub-system* with the conversion of OPAC data which includes the bibliographic descriptions of the documents. In the data structure of the sub-system the *subject headings* are very important. In data received from the OPAC the subject heading is a repeatable metadata. Since it is a significant attribute from a semantic aspect as well which after taking over need to be complemented if required. For this reason it is practical to place subject headings in a separate table. We place also the group code of the image sub-systems in this table. In the *image content sub-system* it is important to support giving a name to the image details and to provide authority control of names. Image details are connected directly to the documents table and they contain *dc.description.elements* or *dc.description.position* metadata which are repeatable attributes with regard to one document, for this reason we store them in a separate table. A librarian who is good at cataloguing has to supply image documents with subject headings or image elements with the appropriate metadata. (Jelenleg kezdetleges állapotban van a képek automatikus osztályozása (Müller, 2001.; Kapoor, 2010.; Tuytelaars, 2010.), sőt valójában szövegek esetében is hasonló a helyzet mind a könyvtárakban, mind a webes környezetben (Tóth, 2002.).) We illustrate the relationship of the catalogue sub-system and the image content sub-system in the following figure:



1. ábra: A katalógus alrendszer és a kepi elemek alrendszer kapcsolata

Besides the above mentioned two sub-systems we need to complement the data structure with a third sub-system, i.e. the *image details information sub-system*. It is developed for search interface of the database. It contains index files for browsing and search. The most important services of the sub-system: browsing (in the most significant attributes of the document, in subject-headings and in image details), search (basic, advanced, according to information content and intelligence value). After all these we have to find the answer only to this question: on the basis of what formulas we can calculate information value of the image documents.

4. Formulas

To calculate the information value of image elements we use Norbert Wiener's formula (Wiener 1948). In this formula I indicates the information value of element x which we calculate in the following way:

$$I(x) = -\log_2 p(x). \quad (1.)$$

Here $p(x)$ means the probability of occurrence of element x within the image sub-system. We obtain the probability of occurrence of the image element in the sub-system by taking the ratio of all occurrences of all elements there (number of all cases) and all occurrences of element x

(number of favoured cases) in the following way: . Adatbázisunkban képi elemként szerepelhet például egy *csésze*, aminek az információs értéke *3,51 bit*. Ennek az az oka, hogy ez az elem hat alkalommal bukkan föl a képeken előforduló ötvenhét képi elem között. A $p(\text{csésze})$ értéke tehát . Információs értékét ezek után a következő módon számoljuk ki: $I(\text{csésze}) = -\log_2 = 3,51 \text{ bit}$.

In an image sub-system we can give values to image documents and sort them with the help of two formulas. The first one computes *information value* of documents, the second one calculates the *intelligence value* of documents within the sub-system. However we wish to define the concept of intelligence value a bit later. We can receive the information value carried by image documents $I(\text{ID})$ if we simply add together information value of elements in the image according to the following formula:

$$I(\text{ID}) = -\log_2 p(x_i). \quad (2.)$$

The $p(x_i)$ is the probability of occurrence of one element in the image within the sub-system, n is the number of elements in the image. According to this, the image document will be identical with the elements which appear in it ($\text{ID} = x_1, x_2 \dots x_n$). During a search by means of this formula we can sort documents according to which one has the highest or the lowest information value in a collection.

Tesztadatbázisunkban az 1930-as évek kávéreklámai között a legtöbb információt a következő kép hordozta:



2. kép: Mallász Gitta: Nem véletlen...

A gyűjteményen belül a kép által hordozott információ mennyisége 50,15 bit. Ennek két oka van: egyrészt az, hogy ezen a képen található a legtöbb elem, másrészt az, hogy a képen található legtöbb elem információs értéke a gyűjteményen belül magas, mivel a többi képen csak ritkán fordulnak elő. Azon a képen azonban, amelyik a legkevesebb információt hordozza nem található más, csak egy kávésdoboz. Mivel a kávésdoboz a gyűjteményben máshol is előfordul, így annak információ értéke átlagosnak tekinthető. A legkevesebb kepi információt tartalmazó dokumentum éppen ezért mindössze 4,25 bitet hordoz.

We count *informativity* of the image documents with the following formula. We mean by informativity that on average how much information is carried by one element of the image. Informativity is often much more revealing than information value in the case of one document in a sub-system. It shows how new or usual the image is in a certain field. The informativity of an image document $H(ID)$ can be calculated with the following formula:

$$H(ID) = -\log_2 p(x_i). \quad (3.)$$

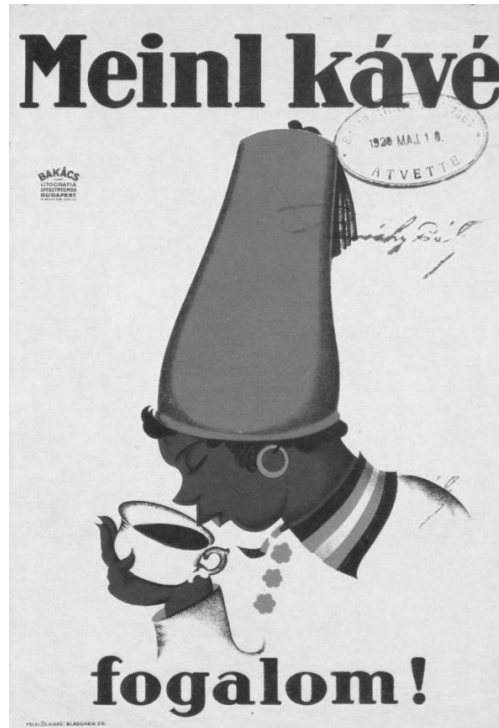
The $p(x_i)$ is the probability of occurrence of one element in the image within the sub-system, n is the number of elements in the image. Measure of the intelligence value is bit/element. Its formula shows a remarkable similarity with Shannon's entropy formula (Shannon 1948), for this reason it received a traditional sign of the entropy, H .

Tesztadatbázisunkban a leginkább informative – vagyis szokatlan – kép a következő volt:



3. kép: Meinl kávé V. keverék

A informativitása abból fakadt a gyűjteményen belül, hogy mindössze egyetlen domináns képi elem szerepel rajta: a címer. Minthogy máshol ez nem fordult elő és egyéb olyan elem nem szerepel a képen, amely rontaná az átlagos információs értékét a képnek, így ez a legszokatlanabb plakát a gyűjteményben. Átlagos és teljes információs értéke is 5,83 bit. A legkevésbé informatív ugyanakkor a következő kép:



4. kép: Meinl kávé fogalom!

Ez a plakát azért tekinthető szokványosnak, mert ebben az időben a Meinl plakátjain sűrűn fordult elő a kávé kortyolgató sapkás, fülbevalós szerecsen fiú alakja. A képen lévő összes elem információs értéke alacsony: sapka (4,25 bit), szerecsen (3,83 bit), fülbevaló (3,83 bit), csésze (3,51 bit), kávé (3,25 bit). Éppen ezért a kép informativitása mindössze 3,73 bit.

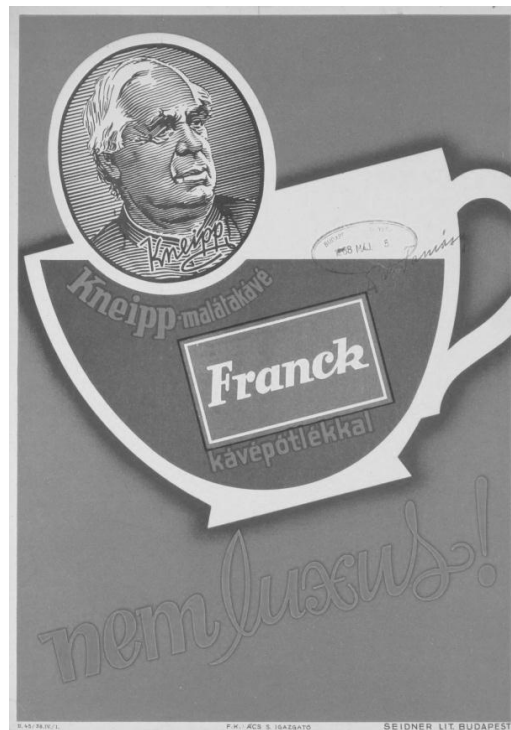
In addition to the above mentioned we can also count what information value is associated with a certain image element within a given document of an image sub-system. Here we can calculate two things as well. In the first case we can compute how much information value the context of an image element has in relation to the element. The $I(Cx_1)$ is information value of context of the image element x_1 that we obtain with the following formula:

	$I(Cx_1) = -\log_2 p(x_i).$	(4.)
--	-----------------------------	------

Here the $p(x_i)$ is probability of occurrence of the element x_i in context of element x_1 within the sub-system, n is the number of elements in the image. We obtain the $p(x_i)$ value in the sub-system by taking the ratio of all occurrences of all the other elements in the context of x_1 element

(number of all cases) and all occurrences of element x_i (number of favoured cases) in the following way: . (Of course element x_1 can occur several times as well in an image, so it also has a probability that this element emerges twice, three times or several times in an image. More times it appears less probability it has. Therefore we have to introduce $p(x_{1,1})$, $p(x_{1,2})$... $p(x_{1,z})$ signs which indicate the probability of the repeated occurrences within the image context).

Érthető okokból a kávé elem bukkant fel leggyakrabban adatbázisunkban. Amikor arra voltunk kíváncsiak, hogy a kávéval kapcsolatban melyik kép hordozza a legtöbb információt, akkor a Nem véletlen... című plakát (2. plakát) let az első. A kávéhoz kapcsolódó kepi elemek itt 45,16 bit információt hordoznak. A legkevesebb információt ugyanakkor a következő kép hordozta:



5. kép: Kneipp-malátakávé Franck kávépótlékkal nem luxus!

A kép összes információ tartalma 6,94 bit. Ennek oka az, hogy férfi fej és csész sűrűn fordul elő a kávékat tartalmazó képeken, így ezek együttes információs értéke sem túl magas.

In the second case we compute that how much informativity value a context of an image element has, i.e. on average how much information is carried by the other image elements in

connection with it. The $H(Cx_1)$ is intelligence value of context of the image element x_1 that we count in the following way:

	$H(Cx_1) = -\log_2 p(x_i).$	(5.)
--	-----------------------------	------

Here the $p(x_i)$ is probability of occurrence of the element x_i in context of element x_1 within the sub-system, n is the number of elements in the image. From our aspect perhaps this value is the most interesting because it shows how new or usual an image or an image detail is from the aspect of the given element. If we conduct a search for a phrase then we can sort documents in the sub-system according to the intelligence value of context of an image element specified by the query. The document with the highest value will reveal the most news about the image element, while the one with the lowest value will depict it in the most usual way.

A kávéval kapcsolatban a legszokatlanabb plakát a gyűjteményben a *Nem véletlen...* című (2. plakát). Átlagos információs értéke 5,02 bit. A magas informativitásnak az az oka, hogy a képen szereplő legtöbb elem csak itt bukkan föl. Ez a plakát tehát a gyűjteményen belül egyszerre hordozza a legtöbb információt és a leginformatívabb is a kávéval kapcsolatban. A legszokványosabb ugyanakkor itt is a *Meinl kávé fogalom!* című plakát (4. plakát). A kép informativitása a kávéval kapcsolatban mindössze 3,36 bit. Mindezeket az értékeket egyszerű táblázatkezelővel számítottuk ki, és nem vettük figyelembe a képeken található elemek geometriai pozíciójából származó információs értékeket. Ennek oka az, hogy jelenleg kutatásaink elején tartunk. Néhány kijelentést azonban már most is tudunk tenni az adatbázisról és a benne való keresésről.

The database management system can operate formulas with its own functions. When a new document enters the database, database management system will recalculate various probabilities or information values of occurrence of each image element every time with its functions. The user can make use of these values for different aims. On the search interface with the application of these values he can e.g. browse and search in index files. Not only a user but a graphical software can also use these values. The task of this software can be to make calculations with vector values related to the image elements. We stored these vectors by means of *dc.description.position* metadata in the image content sub-system of our database. The software counts what probabilities an image element can occur with at one position of one

document in an image sub-collection. It can also calculate for us what information values an image element can have at a given position. Information values arising from the frequency of occurrences and image positions can be added together. All this can make possible further search modes for the users, but on a later occasion we wish to give a talk on it.

5. Conclusions

On the basis of the above mentioned we can conclude that it is possible to conduct a natural language search in image collections of digital libraries. First it was necessary to construct the appropriate metadata system that we achieved with modification of the Dublin Core (ISO 2009). Secondly an appropriate data structure was required that we accomplished with the use of a relational data model (Codd 1969, Ullman 2008). Finally we had to determine those formulas which the functions of a search system used for sorting image documents by their information values or intelligence values. These functions are based on mathematical theory model of communication, first of all on Norbert Wiener (Wiener 1948) and Claude E. Shannon's research (Shannon 1948). During a search we had to introduce a new concept of intelligence value. It indicated average information content of all the other elements which appeared in context of an element in one of the images, or it meant average information content of all the other elements which appeared in context of all elements in one image. The measure of intelligence value is bit/element. Each image element was specified with natural language phrases, so we could use these phrases during a search. The search method can be applied in the most diverse image collections and after further developments – on the basis of our current research - it can be connected with the search systems of full-text databases as well.

References

- E.F. Codd, (1969), Derivability, Redundancy, and Consistency of Relations Stored in Large Data Banks, *ACM SIGMOD Record*, Volume 38 Issue 1, 17-36.
- John Eakins – Margaret Graham, (1999), Content-based Image Retrieval, *JISC Technology Applications*, 4.
- Abby A. Goodrum, (2000), Image Information Retrieval: An Overview of Current Research, *Informing Science*, 3, 63-66.
- ISO 15836, (2009), Information and documentation - The Dublin Core metadata element set
- Krystyna K. Matusiak, (2006), Towards user-centered indexing in digital image collections, *OCLC Systems & Services*, 22, 283-298.

- Jeffrey D. Ullman – Jenifer Widom, (2008), *A First Course in Database Systems*, Pearson Education Inc, Pearson Prentice Hall
- Claude E. Shannon, (1948), A Mathematical Theory of Communication, *The Bell System Technical Journal*, 27, 379-423. 623-656.
- Norbert Wiener, (1948), *Cybernetics: or Control and Communication in the Animal and the Machine*, The MIT Press, Cambridge
- Ka-Ping Yee – Kirsten Swearingen – Kevin Li – Marti Hearst , (2003), Faceted metadata for image search and browsing, *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, 401-408.
- Li-Jia Li – Li Fei-Fei, (2010), OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning, *International Journal of Computer Vision*, 88, 147-168.
- Henning Müller, (2001), Performance evaluation in content-based image retrieval: overview and proposals, *Pattern Recognition Letters*, 22, 593–601.
- Erzsébet Tóth, (2002), Innovative Solutions in Automatic Classification: A Brief Summary, *Libri*, 52, 48–53.
- Dublin Core homepage. <http://dublincore.org/>
- Debreceni Egyetem Elektronikus Archívum. <http://ganyemedes.lib.unideb.hu:8080/dea/>
- Chad Carson – Virginia E. Ogle, (1996), Storage and Retrieval of Feature Data for a Very Large Online Image Collection, *Bulletin of the Technical Committee on Data Engineering*, 19. (1996), 19-27.
- Ashish Kapoor – Kristen Grauman – Raquel Urtasun – Trevor Darrell, Gaussian Processes for Object Categorization, *International Journal of Computer Vision*, 88, 169–188.
- Tinne Tuytelaars – Christoph H. Lampert – Matthew B. Blaschko – Wray Buntine, (2010), Unsupervised Object Discovery: A Comparison, *International Journal of Computer Vision*, 88, 284-302.