# EDUCATIONAL METADATA

## Bela MARKUS

Key words: CPD, distance education, training, virtual academy, metadata

## ABSTRACT

A fundamental transformation continues to occur in higher education. Information and telecommunication technologies are having a profound effect on teaching and learning and are eroding traditional geographical boundaries resulting in greater competition and opportunity. Students are also changing. Part-time, adult learners know that learning is no longer confined to University campuses, and are seeking flexible ways to ' time shift' their learning to meet personal, academic or employment objectives.

The author gives a review of the results, present status and future activities of the projects of CSLM in building a virtual academy and the research study made in designing and implementing an educational metadata base.

## BACKGROUND

The criteria for entry to the EU established at the Copenhagen European Council of June 1993 include the existence of a **functioning market economy** as well as the ability to cope with competitive pressures and market forces within the Union. Fundamental to this is the existence of a sound and flexible system of Land Administration, to apply the *acquis,* and identified elements essential in the land administration. In Hungary the education and professional development of this system has since 1973 been supported by the University of West Hungary, College of Surveying and Land Management (CSLM). Under the aegis of CSLM considerable advances in technology and in corresponding education support have been made since 1989. However, there remain weaknesses in Hungary and in other **CEE countries** that are critical for meeting the EU criterion and for addressing the issue of Institution Building. These were identified in the United Nations Meetings of Officials in Land Administration (**UN MOLA**) Workshop on Land Market in 1998 in Budapest: Lack of education in the management, legal, economic, human and ethical aspects of land administration, General lack of user oriented approach to education, Lack of continuity in education from universities to professions and appropriate linkage between the two.

In the last four years CSLM has been involved in four European Union funded distance learning material development projects targeting professionals in land administration and land / geographic information management and leading at present EU projects on continuing professional developments. For the staff of Land Offices and Surveying / GIS companies distance learning offers a particularly flexible and effective way of training, eliminating most of the barriers, providing much better accessibility than traditional education.

There are three common strategic elements in these projects. First, the creation of a **knowledge centre** at CSLM which will form the base for programmes of education for continuing professional development for Land Administration in Hungary utilising existing programmes developed under the TEMPUS OLLO Project and the existing Land Administration infrastructure. To these will be added new programmes for higher management and a set of programmes for all levels in Land Administration focussing on matters pertinent to EU entry. Second, a **delivery system** for continuing professional development based on adapted curriculum structures, management tools and education technology, both CD and web, with a comprehensive **credit** system. Third, the creation of a **network of EU centres** and education providers with the objective of participating fully in EU activities in Land Administration and the EU professional community.

While in the first projects (OLLO and DLG) we have successfully introduced new approaches to teaching (open, distance learning) and CSLM significantly strengthened as a centre in professionally oriented land information management programmes, in the later stages the globalisation became more and more important (network-based education). The new SDiLA TEMPUS project is widening the target area and looking forward the needs arise from the opportunities, which now exist or under preparation in Hungary in connection with the EU accession. Principal amongst these requirements for the future is the need to disseminate a broad understanding of the ways in which land related data can be used, particularly in preparation for accession to the EU. Such degrees of understanding and competency must also be spread through relevant areas of the civil service and must not be isolated in pockets of the land administration sector.
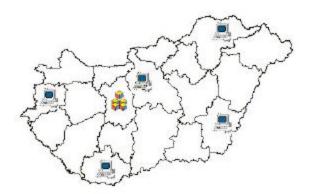


Fig. 1 – The SDiLA course is developed by CSLM and will be delivered by five sites

Thus, while OLLO sought to target the professional land officer, such defined posts are not targetable in the civil service at large and thus attempts are made in SDiLA to reach an audience covering a range of job functions and at differing levels of responsibility. Thus, where OLLO has developed a firm educational foundation for professional staff, SDiLA now seeks to build towards a more flexible programme of short cycle staff development activities. This differs from OLLO in that study will not be part of an overall academic programme, it can thus be more flexible and will be targeted in different ways towards differing levels of ability and differing staff requirements. SDiLA will be able to call upon the network resources of TAKARNET and will thus be able to take full benefit of CBT techniques.

The course material development is based on knowledge base approach and the course delivery on a distributed environment. The CSLM acts as a knowledge centre developer, land office study centres and high schools are dealing with course delivery. This necessitates the extended use of metadata on the educational resources.



Fig. 2 – TAKARNET is connecting Land Offices

Metadata will be crucial in implementing these strategic elements. Whilst learning units itself forms the building blocks of networked and inter-connected environment meta data is required to bind the units together and allow them to interoperate. Metadata is required to describe what learning units look like, how to build from them a learning route, what, if any refinements or value adding operations have been carried out on a unit. And in a networked environment what services can a tutor/learner request from a server and what parameters should the teacher/student send to the server to request the service.

**DATA ABOUT DATA**

Metadata can be called "data about data". The term "metadata" refers to **background information** about something. The concept of metadata is familiar to most people who deal with spatial issues. A map legend is pure metadata. The legend contains information about the publisher of the map, the publication date, the type of map, a description of the map, spatial references, the map's scale and its accuracy etc. Metadata are simply that type of descriptive information applied to a geospatial object. They are a common set of terms and definitions to use when documenting geospatial data. Metadata helps people who use data, find the data they need and determine how best to use it. Metadata benefit the data producing organization as well. As personnel change in an organization, undocumented data may lose their value. Later workers may have little understanding of the contents and uses for a digital data base and may find they can not trust results generated from these data. Lack of knowledge about data produced by other organizations can lead to duplication of effort. In the broad field of information management, metadata is that "stuff" that helps a person locate and then understand data, whether that data is in the form of a simple list, a spreadsheet, a database, a CAD drawing, or a GIS "map". Especially when data is computerized, it can become impossible to understand its essential details without appropriate background information. After selecting, ordering, and receiving external data, an organization

needs metadata to guide proper use of that data. In this way, metadata supports effective data sharing.

Metadata in education is a description of learning objects (like courses, subjects, learning materials, learning units, documents or educational services), which may contain data about their form and content. The best known metadata to us are the catalogue records for printed publications. Metadata is widely used outside libraries as well, e.g. in GIS data warehouses. ISO/TC 211 identifies four roles which metadata should play:

- availability - data needed to determine the sets of data that exist for a geographic location.
- fitness for use - data needed to determine if a set of data meets a specific need.
- access - data needed to acquire an identified set of data.
- transfer - data needed to process and use a set of data.

Educational metadata should provide the minimum requirements to define the attributes required to fully/adequately describe the educational objects. The reason for creating metadata, from the provider perspective, is to improve the possibilities of retrieval as well as to support control and management of learning objects. As the volume of digital learning resources increase keeping track of, and identifying resources which is often maintained by different organisations becomes increasingly difficult. A cross domain method of describing learning resources is required. Metadata will help potential learners 'discover' what information is available and will help them assess the suitability of that data for a given task. Distance learning materials and services with their abundance of different formats and control measures might not always be usable directly by everyone: the format might be unfamiliar or unreadable, the content might be encrypted, otherwise prohibited or only permitted after payment, the resource might be large, difficult or time consuming to access etc. In all those cases, metadata could support the learning process.

The costs of metadata creation, especially when manually generated, have lead to investigations of the possibility, advantages and disadvantages of using author, course developer, tutor or mentor provided metadata. Schemes for metadata creation by non-specialists have to be very simple, short, easy to understand and to use. They have to be built into editors and authoring packages and need supporting help and manuals. The ever increasing number of digital documents makes it necessary to use developer provided metadata for a large part of the learning resources. Specialist "cataloguers" will probably only contribute with a minor percentage of resource descriptions and will be likely to concentrate on quality- and long lasting documents. Whoever provides the metadata, it has to be interoperable and machine usable, because a richer metadata provision is the most important step needed to improve the quality and precision in automatic resource discovery and retrieval tools.

The aim is to develop a metadata model which assists:

- management
- discovery (including discovery of multimedia)
- commercial aspects
- evaluation
- security

- privacy
- quality control
- verification and
- authentication

There are three main levels of metadata. Collection level metadata provides the user with a quick look at the data. The user will be able to gain an overview of the contents and scope of the data set. Collection level metadata forms are often fairly short. They can be filled out easily by hand. Data set level metadata is more detailed than collection level. It provides a fuller picture of what a data set will contain, describing the types of features encoded and possibly the lineage (history) of the data set. Data set level metadata formats are typically verbose, running to several pages of printed text and can be boring and laborious to complete. Feature level descriptions provide very detailed descriptions of a data set. Such detailed records are used to describe the behaviour of individual objects in a data set. This description is required to allow the objects themselves to be transferred and manipulated in an open networked environment. The metadata model and metadata creation tool must also be user friendly.

## THE DUBLIN CORE

The Dublin Core was developed at the Dublin Workshop held in 1995 in Dublin, Ohio. Metadata records created from the Dublin Core are intended to mediate the above mentioned extremes, affording a simple structured record that may be enhanced or mapped to more complex records as called for, either by direct extension or by a link to a more elaborate record. The goal of the Dublin Workshop was to define a small, universally understood set of metadata elements that would allow authors and information providers to describe their work and to facilitate interoperability among resource discovery tools. The syntax was deliberately left unspecified as an implementation detail. The semantics of these elements was intended to be clear enough to be understood by a wide range of users.

Dublin Core - Element Description
- **Subject:** The topic addressed by the work
- **Title:** The name of the object
- **Author:** The person(s) primarily responsible for the intellectual content of the object
- **Publisher:** The agent or agency responsible for making the object available
- **OtherAgent:** The person(s), such as editors and transcribers, who have made other significant intellectual contributions to the work
- **Date:** The date of publication
- **ObjectType:** The genre of the object, such as novel, poem, or dictionary
- **Form:** The data representation of the object, such as Postscript file or Windows executable file
- **Identifier:** String or number used to uniquely identify the object
- **Relation:** Relationship to other objects
- **Source:** Objects, either print or electronic, from which this object is derived, if applicable
- **Language:** Language of the intellectual content

- **Coverage:** The spatial locations and temporal durations characteristic of the object

As an example, the following list is a Dublin Core metadata record given for this paper:
- **Subject** = Metadata
- **Subject** = Dublin core set
- **Subject** = Internet-based distance learning
- **Subject** = Learning resources storage and retrieval systems
- **Title** = Educational metadata
- **Author** = Markus, Bela
- **Publisher** = FIG
- **Date** = May 24, 2000
- **ObjectType** = computer file
- **Form** = text/html
- **Identifier** = http://www.fig2000.cz
- **Relation** (type=sibling) = http://www.ddl.org/figtree/office/prague.htm
- **Language** = English

After the first investigations became clear the Dublin Core will be valuable for many reasons. The Dublin Core became a standard, metadata records could be understood across user communities. The metadata record created with the Dublin Core could serve as the basis for a more detailed description if the need arises. The Dublin Core is flexible enough to modify with the experience of other researches. Many of the current efforts in metadata specification and standards are based upon the firm foundation of the Dublin Core, providing extensions where the core is not specific enough.

However the Core is also weak, as a result of its simplicity. Unlike more complex formats the Core doesn't have any formal syntax, has only vague semantics and there is no controlled vocabulary defined. Such a loosely defined format is of limited value and open to widespread misuse (UNIPHORM, 1998).

**IMPLEMENTATIONS**

The goal of the second workshop on metadata in Warwick was formulated "... to identify implementation strategies that will serve two main purposes:
- promote semantic interoperability across disciplines and languages, and
- define mechanisms for extensibility to support richer description and linkages to other description models" (Hakala et. al, 1996).

The Warwick framework provides an architecture for the interchange of distinct metadata packages. The framework itself is distinct from the syntax or semantics of any specific metadata dictionary, and promotes interoperability and extensibility by allowing tools and agents to selectively access and manipulate metadata packages as a whole. The most important result of this workshop was the proposal for a container record architecture, comprising many more and different types of metadata than a Dublin Core record. The elements of the Dublin Core were not changed, this standard should be kept crisp and the elements themselves should focus on the description of the documents form and content and not be extended anymore. The author should write

metadata into the HTML page or third parties could produce metadata on documents of others. A container, carrying other types of metadata, should be wrapped around this extracted or separate Dublin Core metadata information. The Warwick framework allows to combine good extensibility to provide elaborated schemes to certain communities with a simple interoperable Dublin Core description of form and content of the objects.

An other effort made by IEEE Learning Technology Standards Committee (LTSC) to define a Learning Object Metadata (LOM) Standard. This standard will specify the syntax and semantics of Learning Object metadata, defined as the attributes required to fully/adequately describe a Learning Object. Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology-supported learning (IEEE, 1998).
Purpose:
- To enable learners or instructors to search, evaluate, acquire, and utilize Learning Objects.
- To enable the sharing and exchange of Learning Objects across any technology supported learning systems.
- To enable the development of learning objects in units that can be combined and decomposed in meaningful ways.
- To enable computer agents to automatically and dynamically compose personalized lessons for an individual learner.
- To compliment the direct work on standards that are focused on enabling multiple Learning Objects to work together within open distributed learning environment.
- To enable, where desired, the documentation and recognition of the completion of existing or new learning & performance objectives associated with Learning Objects.
- To enable a strong and growing economy for Learning Objects that supports and sustains all forms of distribution; non-profit, not-for-profit and for profit.
- To enable education, training and learning organizations, both government, public and private, to express educational content and performance standards in a standardized format that is independent of the content itself.
- To provide researchers with standards that support the collection and sharing of comparable data concerning the applicability and effectiveness of Learning Objects.
- To define a standard that is simple yet extensible to multiple domains and jurisdictions so as to be most easily and broadly adopted and applied.
- To support necessary security and authentication for the distribution and use of Learning Objects.

Within the education arena, there is a move to define standards for metadata so that an appropriately constructed search engine could search for pages on the basis of a large number of criteria - not just author, publication date etc. but also specific categories of subject matter. A number of bodies have taken it upon themselves to be responsible for this standardisation and the process of making this a world standard rather than a local one has begun. The Gateway to Educational Materials (GEM) project has extended the Dublin Core with metadata to support the description of lesson plans, curriculum units, and other education resources. GEM is a special project of the Educational Resources

Information Center Clearinghouse on Information and Technology (ERIC/IT). The major players nowadays probably are: ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe) and the IMS (Instructional Management System) project.

**IMS and ARIADNE**

In 1997, The IMS Project, part of the non-profit EDUCOM consortium (now EDUCAUSE) of US institutions of higher education and their vendor partners established an effort to develop open, market-based standards for online learning, including specifications for learning content meta-data. Also in 1997, groups within the National Institute for Standards and Technology (NIST) and the IEEE P.1484 study group (now the IEEE Learning Technology Standards Committee - LTSC) began similar efforts. The NIST effort merged with the IMS effort, and the IMS began collaborating with the ARIADNE Project, a European Project with an active meta-data definition effort.

In 1998, IMS and ARIADNE submitted a joint proposal and specification to IEEE, which formed the basis for the current IEEE Learning Object Meta-data (LOM) base document, which is a classification for a pre-draft IEEE Specification. IMS publicized the IEEE work through the IMS community in the US, UK, Europe, Australia, and Singapore during 1999 and brought the resulting feedback into the ongoing specification development process.

The IEEE LOM Base Document defines a set of meta-data elements that can be used to describe learning resources. This includes the element names, definitions, data types, and field lengths. The specification also defines a conceptual structure for the meta-data. The specification includes conformance statements for how meta-data documents must be organized and how applications must behave in order to be considered IEEE-conforming. The IEEE LOM Base Document is intended to support consistent definition of meta-data elements across multiple implementations, but does not (at the time of this writing) include information on how to represent meta-data in a machine-readable format, necessary for exchanging meta-data. The number of items defined within the IEEE LOM Base Document was large and many participating organizations within the IMS community recommended that a select Core of elements must be identified to simplify initial implementation efforts. The IMS developed a representation of the meta-data in XML (eXtensible Markup Language) and surveyed its member institutions around the world to identify the Core elements.

The IMS will continue to offer guidance and support documents related to the IEEE meta-data efforts. The IMS community will continue to present the IEEE community with reference binding and implementation documents for a variety of learning resource needs such as enterprise interoperability, content packaging, and learning management. It is hoped that such reference documents may be helpful in the development of IEEE sanctioned binding and implementation guidelines.

The IEEE conceptual model for meta-data definitions is a hierarchy. At the top of the hierarchy is the "root" element. The root element contains many sub-elements. If a sub-element itself contains additional sub-elements it is called a "branch." Sub-elements that

do not contain any sub-elements are called "leaves." This entire hierarchical model is called the "tree structure" of a document.

ARIADNE is supported by the Commission of the European Union in the framework of the education and training program of the Telematics Applications Program. ARIADNE's primary goal is to foster the share and reuse of electronic pedagogical material, both by universities and corporations. In addition, ARIADNE is building a large Europe-wide repository for those pedagogical documents, that has been called the Knowledge Pool System. One of the key features of the Knowledge Pool System is the underlying metadata specification, which is being revised currently in view of the results of our extensive experimentations.

ARIADNE does not intend to develop metadata to describe the human actors involved in the process of education and training, to characterize or to record their educational performances. Neither is the intention of the present document to define the representation format for the metadata sets. For the latter purpose, use could be made of SGML, XML, RDF, a DBMS, etc. (The current Knowledge Pool System relies on a combination of SGML and a DBMS).
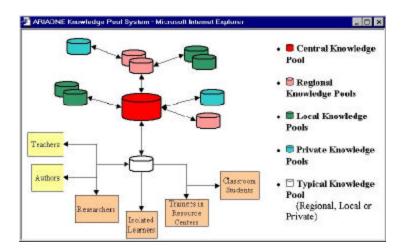


Fig. 3 – Knowledge Pool System defined by ARIADNE

ARIADNE aims to solve two practical problems that will arise in the case of the widespread use of any metadata system:
- indexation work (i.e. the creation of the metadata itself by human persons) should be as easy as possible;
- the exploitation of the metadata by users looking for relevant pedagogical material should be as easy and efficient as possible.

In the context of Europe, but also more generally of wide international cooperation, it is mandatory that the metadata system works in a multilingual and multicultural environment, and thus be neutral with regards to both the language of the original document being indexed and the language used to create the metadata. Mechanisms that would ensure such a multilingual interoperability are not easy to design and implement. We will not describe them here as they are really implementation issues. However, the

metadata set must capture the necessary information to render such mechanisms feasible.

Metadata can naturally be regrouped into families that will make them easier to provide. This translates into a structure set for the metadata descriptors. There is a small number of these categories of descriptors. We propose that most of them be mandatory, to ensure the stability of the scheme (even though those mandatory groups may contain optional fields).

It is recommended that the following categories of descriptors, which are presented in a logical order, be mandatory for any educational resource being indexed.
- general information on the resource itself,
- semantics of the resource,
- pedagogical attributes,
- technical characteristics,
- conditions for use,
- meta-metadata information.

Examples of optional categories are as follows:
- annotations.
- physical data of the represented educational resource.

**CUBER**

The project CUBER will develop a broker service that provides access to the vast collection of courses from European distance universities and helps the citizen to find a best match of vocational demands, academic offerings, and individual learning conditions. By integrating the universities' course offers through a common broker middleware that standardizes meta data, i.e., data about courses, a federated virtual university of Europe comes to existence.

This broker middleware will comprise three technical components:

- a user-centered customizable search engine;
- a knowledge base of standardized course descriptions learner stereotypes, and adaptable search patterns; and
- a forms- and menu-based authoring interface through which course providers enter and maintain their course meta data.

The search and inference engine will operate on top of the knowledge base to lead the individual learner through an iterative problem solving dialogue. Each dialog will end up with a package of courses or a complete study program meeting the learners' specific needs including qualification objectives, personal interests, difficulty level, and learning conditions. Course descriptions will be standardized through a set of learning objects meta data and a lexical database of subject-related technical terms. Special emphasis will be put on the inclusion of meta data that reflect the ECTS standard but add descriptors that take into account course quality and difficulty level. The knowledge base will include compatibility rules referring to course data such as teaching objectives, subject domains and qualification levels. The authoring interface will provide plausibility and consistency checks to maintain a high quality knowledge base.
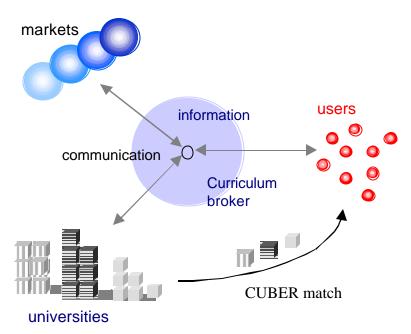
Fig. 4 - Broker and Gateway to a federated virtual university of Europe

The broker will provide a single-system view of the totality of courses offered despite their distribution across multiple European distance teaching universities and the disparity of languages in which these courses are taught. The broker helps to manage the dynamics of the European education market by regularly polling the content providers' electronic course catalogues for changes that may affect the CUBER knowledge base.

**CONCLUSIONS**

Metadata can be used to improve the search process, to build user-specific, guided paths, and to maintain relationships among disparate educational resources. Several metadata projects are under development to target and standardize the instructional qualities that are most useful in describing educational resources. The efforts of these projects will provide a metadata foundation that can be leveraged by future instructional applications. We suppose that the continued success of the Internet is contingent upon automated tools that efficiently guide the information gatherer toward relevant and appropriate material. In this paper, we have described our use of educational metadata in building applications for personalized navigation and search in the context of learning. Ultimately, we believe that the Web will be a viable environment for providing individualized instruction applications for learners in Land Administration.

The SDiLA Project would greatly improve the knowledge transfer from the above mentioned projects to Hungary and allow us to really participate in the international educational development arena, rather than to become late users of systems and regulations invented by others. SDiLA aims to increase collaboration between EU and Hungarian institutions and sharing of learning resources. Co-operation will support specialization, improve quality, increase choice, and lead to a better fit with changing vocational demands in Land Administration. The job market in general will become much more dynamic, complex and heterogeneous. The increased complexity will

increase the difficulty of optimizing job offers and job demands. SDiLA assists to avoid these problems and to develop more market oriented curricula. Since the strategic aim of SDiLA is directly support a European accession, the project will improve interaction beyond national boundaries and will facilitate the development of standards. Potential clients of SDiLA include not only Land Office staff, but professionals in land surveying, local governments, regional offices etc. The openness of SDiLA will also invite partners from Eastern European countries to participate in the dissemination.

**References:**
1. Anderson, T. McArthur, D. Griffin, S. Wason, T. 1999: IMS Meta-data Best Practice and Implementation Guide, Version 1.0, EDUCAUSE. http://www.imsproject.org/metadata/mdbest01.doc
2. ARIADNE, 1999: Educational Metadata Recommendation, Version 3.0
3. Brown, G. Markus, B. and Petch, J., 2000: SDiLA project document, London - Manchester - Salzburg - Szekesfehervar - Vienna.
4. Communication from the European Commission, 1993: White Paper on growth, competitiveness, and employment: The challenges and ways forward into the 21st century.
5. Communication from the European Commission, 1997: Towards a Europe of knowledge.
6. Hakala, J. Husby, O. and Koch, T., 1996: Report from the Metadata Workshop II, Warwick, UK.
7. IEEE Learning Technology Standards Committee (LTSC), 1998: Learning Object Metadata (LOM), Draft Document v2.1.
8. Krämer, B. 2000: CUBER - Personalized Curriculum Builder in the Federated Virtual University of the Europe of Regions, Hagen.
9. Márkus, B., 1998: Building a knowledge pool for land administration, UN ECE MOLA Meeting on Land Market, Budapest.
10. Medyckyj-Scott, D. Newman, I., Ruggles, C. and Walker, D., 1991: Metadata in the geosciences, Group D Publications.
11. Niklasz, L., 1999: Strategy of Staff Development in Hungarian Land Offices, FIG Commission 3 Annual Meeting and Seminar, Budapest.
12. UNIPHORM, 1998: Open GIS course materials, CSLM.

**Biographical Notes:**
Prof. Dr. Béla MÁRKUS is the head of Department of Geoinformatics at the College of Surveying and Land Management, The University of West Hungary. He is a surveyor with research interests in applications of digital elevation modelling and error handling aspects of GIS use. Prof Markus is the editor of the NCGIA Core Curriculum Hungarian version, Dictionary of GIS Terms and over sixty published papers on various aspects of using GIS, actively involved in many national and international academic programmes, chairman of the Hungarian UNIGIS Course Board, chairman of Educational Affairs of the Scientific Committee in Geodesy at the Hungarian Academy of Sciences, chairman of Working Group on Geoinformatics of Association of Hungarian Surveyors and Cartographers and the national representative of FIG Commission 2.

**Contact:**

*Prof. Bela Markus*

*Department of Geoinformatics, College of Surveying and Land Management, The University of West Hungary, Pirosalma 1-3, 8000 Szekesfehervar, Hungary*

*Tel.: +36 22 348 271, Fax.:+36 22 327 697, E-mail: mb@cslm.hu*

*http://geoinfo.cslm.hu*