

Doktori (PhD) értekezés tézisei

**Okos egyetemi szolgáltatásokat
megalapozó kapcsolt adathalmazok**

Szász Barnabás

Témavezető: Dr. Micsik András



DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2019

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Smart Campus based on Linked Data

by Barnabás Szász

Supervisor: Dr. András Micsik



UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF INFORMATICS

DEBRECEN, 2019

Table of Content

1	Background and motivation.....	1
2	New scientific results and theses	3
2.1	Auto-generating URIs for Linked Data resources	3
2.2	Holistic Linked Data publishing for the university domain... 6	
2.3	Ontology for semantic indoor space descriptions	8
2.4	Indoor navigation based on preferences, abilities, and permissions	10
2.5	Modeling indoor spaces with the iLOC ontology	11
3	Bevezetés és motiváció	14
4	Kutatási eredmények.....	16
4.1	Automatikus URI generálás Kapcsolt Adathalmazokhoz....	16
4.2	Holisztikus Kapcsolt Adat publikálás a felsőoktatásban	20
4.3	Ontológia szemantikus beltérek leírásához.....	22
4.4	Beltéri navigáció személyes preferenciák, képességek és jogosultságok alapján.....	24
4.5	Beltérek modellezése az iLOC ontológia segítségével.....	26
	References.....	29

1 Background and motivation

Universities around the world are facing nowadays new questions and challenges, while they are working on their strategy. These include: How universities should change to satisfy new demands and needs? What role new technologies play in how students learn, and academics teach? And how are university campuses going to adapt to new technologies whilst maintaining their position as incubators for exploration and innovation? In this new competitive landscape of services, the university campus needs to remain a relevant place of engagement. The focus, therefore, is on making the university a more sustainable place for human exchange and interaction. As in every aspect of modern life, technology is playing a major role in this transition too.

Nowadays, information sharing, and consuming is happening online, using different services, such as email and messaging, file transfer, chat, and video calls and many other, but still the most popular service is the World Wide Web. The traditional use cases of the WWW were focusing on human users as direct consumers and producers of the content. The data in most Web pages is at best semi-structured [1] and a large percentage of the pages are “hidden” behind form-based

interfaces designed for human users [6], resulting in a new set of best practices for publishing and connecting structured data on the WWW. The resulting form of the published data is commonly referred to as Linked Data¹ and has gained tremendous momentum in recent years.

One aspect of the Smart University or Smart Campus concept is to improve the teaching and learning environment using modern data fusion and data consumption techniques. A campus has a large number of people with a substantial set of common information needs [8]. Therefore, it has great importance in this area to establish a common data model which enables the interconnection of fragmented data from heterogeneous data sources.

The overall goal of publishing Linked Data about universities is to open up and connect the rich data sources, both concerning administrative and research data.

¹ <https://www.w3.org/standards/semanticweb/data>

2 New scientific results and theses

This thesis contributes to the “Smart Campus” topic, researching the opportunities and benefits of publishing university related open linked data in a machine consumable way. In this section, a short overview is provided for each research objective followed by the main contributions.

2.1 *Auto-generating URIs for Linked Data resources*

Linked (Open) Data (LOD) resources are identified by Uniform Resource Identifiers (URIs). Publishing Linked Data on Linked Data Platform [12] demands certain best practices, e.g. to provide meaningful URIs to identify entities, so it is an important step to define the conventions for URI assignments. Although it is widely accepted, that URIs should not carry any semantic meaning, in the practice it provides value, if the URIs are human-readable, so a human being can also easily identify a specific resource in a list of URIs.

In some cases, resources already have their natural identifiers to be used as the local part of the resources’ URI, or they can be inherited from other databases. In other cases, resources might not have an inverse-functional property providing a single identifier for their URI, only combination of multiple keys can identify such resources. Typical

examples are recurring events without specific titles or other identifiers for each recurrence. These events can be identified by their exact time and location. RDF (Resource Description Framework)² provides tools to completely omit identifiers and use Blank Nodes, but these have further issues as such resources cannot be referenced directly.

For human usage – especially for manual entry – it is important to encode some semantics in the URI structure. This can cause redundancy in the data, as these entities should be linked to other entities while this information may present in the URI as well. If this redundancy cannot be avoided, some automated mechanism should take care of the maintenance or verification.

In the context of Linked Data processing, to support correct queries on the data, it is often crucial to have a complete description of the entities as one might expect certain properties for certain entity types. The Self-Unfolding Semantic URIs were proposed for this purpose: these are a pair of URIs following a specific pattern and a template, which describes the structure the entity should have. Let's suppose a LOD dataset is given and there is a mechanism (e.g. a trigger) monitoring the data that in case of a special type of entity appearance generates new LOD triples that are semantically derived from the old ones.

² <https://www.w3.org/RDF/>

In [12] four scenarios, i.e. four problem types, have been identified, where automatic data generation with self-unfolding URI could provide a solution. These are

1. entities with composite keys in URI,
2. Blank Node (Anonymous Resources) Candidates,
3. linking related entities based on URI and
4. reasoning on information encoded in the URI.

A key part of the self-unfolding URI process are the templates expressing rules, which determine the generation of the new information based on the entity's URI. Templates of a specific type of entity are supporting query consistency. The template provides a URI pattern and prescribes RDF triplets to be generated based on the values extracted from the URI.

A URI pattern has been proposed based on the Matrix URI pattern³ which can use multiple independent parameters of the entity. A URI following this pattern starts with a base part, followed by the type of the resource and then optionally each part of the composite key.

Thesis 1.

The Self-Unfolding URIs were proposed to support URI assignment for RDF resources without inverse functional properties. These are URIs following a pattern and a template with a new mechanism for automated triple generation based on

³ <http://www.w3.org/DesignIssues/MatrixURIs.html>

the structured information given in the URI. Self-Unfolding URIs support human RDF data manipulation by reducing complexity and improving consistency, as well as eliminate generated URI conflicts. The self-unfolding URI mechanism was applied successfully in the OLOUD server of Óbuda University, and examples were provided for several more application scenarios in [12].

2.2 Holistic Linked Data publishing for the university domain

In the Hungarian practice, university information systems are typically separate and somewhat isolated from each other. There are science management, e-learning, etc. services, but the integration of such services and data sources are still not fully utilized. Currently, there is no completely integrated database in the Hungarian higher education holding and serving all related data. The available data sources are usually acting as closed systems, and although the data itself is public, public access is not granted. Data about courses, students and lecturers are managed in separate systems and the connections between them are not always set explicitly. Study materials for courses and students are not easy to determine.

Publishing Linked Data requires using a (or multiple) schema(s) to define the meaning of the used terms. To access implicit relationships,

it is worthwhile to model the relevant concepts and their relationships in formal ontologies. Using ontologies to publish Linked Data adds more explicit semantics to the schema and enables inference on the data.

The Ontology for Linked Open University Data (OLOUD) is a practical approach to model the education activities of a typical Hungarian university. Although there are several ontologies to cover different parts of the university domain, a single ontology would be too complex to cover the whole use case. We summarize issues with the source ontologies, such as missing links, inconsistencies as well as many overlaps between them and provide a holistic model.

The OLOUD ontology aims to reuse and integrate concepts from several sources to support publishing and consuming personal timetables, navigation and other types of help for students and lecturers. The modeled domains include lectures, semesters and personnel data, but also buildings, events and learning related concepts. OLOUD acts as a glue for a selection of existing ontologies, and thus enables to formulate SPARQL⁴ queries for a wide range of practical questions of university students. The ontology enables the universities to publish data from multiple data sources in a consistent and easy to use manner.

Our papers, [2], [9] and [10] provide further details about our work.

⁴ <https://www.w3.org/TR/rdf-sparql-query/>

Thesis 2.

The new Ontology for Linked Open University Data (OLOUD) provides a holistic knowledge model to enable Linked Data publishing in the higher education domain [2]. Based on the analysis of available ontologies covering the domain, OLOUD bridges the gaps between existing ontologies and completes the semantic model for universities. Data from the Óbuda University has been published as Linked Open Data using the concepts of OLOUD.

2.3 Ontology for semantic indoor space descriptions

The objective of the iLOC ontology is to support the publishing of indoor location and navigation data as Linked Open Datasets and to support the development of applications using the data. The ontology facilitates location description and links to relevant other data. It is also able to describe how two places are connected including distance or constraints (e.g. a route section is not recommended for wheelchair users). It can provide routes to unknown places described by their semantic properties (e.g. nearest wheelchair accessible restroom or vending machine providing gluten-free sandwiches, etc.). Currently, we do not know about any indoor navigation services utilizing Linked Data, and there is also lack of online available vocabularies or ontologies for the description of building plans.

The suggested Linked Data model enables the integration of various datasets related to buildings and can be used not only in the university environment, but in museums, shopping malls, hospitals, campuses, stadiums, airports, etc. By integrating iLOC with other ontologies, navigation queries can be enriched with filtering closures defined by concepts from third-party ontologies. These can define room categories or accessibility options and route types. Possible use cases include the description of nearby places, planning routes or finding places with given features, even in emergency situations. iLOC supports three layers for a building plan, with increasing complexity and capabilities. According to the classification of Worboys [13], iLOC is a hybrid ontology that supports entity information, connectivity and distance as well.

The applied topologic description is the most accurate supporting step-by-step navigation, which was extended by semantic descriptions (e.g. type of rooms, access level, capability and preference of users, etc.).

The ontology is documented in our papers: [3], [4], [5] and [11].

Thesis 3.

The iLOC indoor ontology has been constructed for semantic indoor space description. We were the first to demonstrate with a mobile application developed at the Óbuda University, that using standard SPARQL queries and the iLOC ontology, one can get personalized indoor routes from Linked (Open) Data. Example SPARQL queries for route calculation were provided

and the performance has been measured with different SPARQL engine implementations. [3], [4], [5] and [11]

2.4 Indoor navigation based on preferences, abilities, and permissions

Outdoor navigation is widely available nowadays and helps people to find a place while driving or walking or using public transport. This type of navigation is usually based on coordinates provided by GPS or Glonass signals. Inside buildings, however, a navigation system must cope with more complex routes using floors, elevators, and staircases in the lack of GPS/Glonass signals. The users might be more limited regarding their abilities, permissions and their preferences could be more fine-grained. The current solutions usually require special and expensive hardware for exact positioning. Thus, opposed to the outdoor scenario, there is no single, accepted basic methodology for navigation.

For this purpose, we introduced the iLOC ontology to support indoor navigation with Linked Open Data. The ontology facilitates locations description and links to relevant data. It can also describe how two places are connected including distance or constraints (e.g. fulfilling wheelchair requirements or needing special badges, etc.).

iLOC supports three navigation methods, with different complexity and accuracy. While the first method only informs the user which

building, part of building and floor they should look for, the second is able to provide the actual steps it takes to navigate through the building and the third method provides the most accuracy by describing the dimensions of the steps as well as some of the restrictions one can face while being underway.

The iLOC Ontology supports more advanced and accurate navigation, providing tools to describe the route sections as restricted, one-way or with metrics to quantify the length or time duration between the start and end points. Furthermore, routes can be calculated with flexible constraints for users' goals and abilities.

In [3], [4], [5] and [11] we were the first to demonstrate, that using standard SPARQL queries and the iLOC ontology, one can get personalized indoor routes from Linked (Open) Data.

Thesis 4.

People with disabilities or special needs are facing challenges when navigating in public spaces. The iLOC ontology has been extended in [4] and [5] to enable personalized route planning aligned to permissions, capabilities, and needs.

2.5 Modeling indoor spaces with the iLOC ontology

When we ask someone the way to some places, people usually give route descriptions in an abstract and symbolic manner referring to different easy to identify landmarks or signs along the route. In order

to provide a good indoor navigation experience, the guidance needs to integrate human wayfinding principles in route instructions.

There are multiple approaches to model indoor spaces. These models are fundamentally different considering their use, cost of modeling and the use of the model and the concepts the models are using. Models can be structural, geometric and semantic, amalgamating the properties of structural, geometric and topologic models. Structural descriptions relate rooms and concepts to each other in a hierarchy, not providing the required level of details for indoor navigation, while geometric descriptions (e.g. floor plans) are too complex, processing and using them for indoor navigation can be troublesome.

The iLOC ontology provides concepts for the topological description of indoor spaces, which is the most important part of a semantic indoor description. In a nutshell, with the ontology one can describe a network of connected POIs (Point of Interest). Semantic indoor description can be derivated from their geometric (e.g. floorplan) and topologic description, omitting certain complexities and dividing up bigger spaces into smaller ones. During the process, the building should be marked with easy to identify POIs and the available direct routes between these. Indoor routes will always lead through multiple POIs marked in such a way.

Using the floorplan, one needs to systematically identify these points, making sure there is an ideal coverage within the rooms. As this process is cumbersome and error-prone, it is important to follow the

outlined methodology in [3]. It provides guidelines to classify the building, how to divide larger spaces, like hallways, up to smaller ones, how to identify the necessary POIs, how to deal with multi-floor layouts and how to structure the POI network. It also gives examples of special cases, like multi-floor rooms.

Smaller rooms with single entrance need to be treated differently than bigger rooms with multiple entrances. Also, the role of the room impacts the type and the number of points to be identified. Next to identifying the right points on the floorplan, there are other special cases, where passages might be one-way or restricted by other rules. Also, to consider the personal needs and abilities require further data to be collected about the passages. The quality of the resulted semantic map relies heavily on the process of the map's production. Linked Data can support indoor navigation if proper semantic models are built from the structural and geometric description of the involved buildings using concepts from the iLOC ontology.

Thesis 5.

Creating semantic models from complex layouts might have various challenges, which were analyzed, and a methodology and best practices have been introduced defining the steps of such semantic indoor space modeling [3].

3 Bevezetés és motiváció

A felsőoktatási intézmények manapság új kérdésekkel és kihívásokkal szembesülnek, amikor a hosszú távú stratégiájukat alakítják. Ezek többek között: Hogyan elégítheti ki egy egyetem az új követelményeket és szükségleteket? Milyen szerepe van a technológiának abban, ahogyan a hallgatók tanulnak és az oktatók oktatnak? Hogyan tudják az egyetemi kampuszok az új technológiákat alkalmazva megtartani az inkubátor szerepüket a kutatás fejlesztésben? Ebben az új helyzetben az egyetemeknek meg kell tudni őrizni versenyképességüket, megfelelően támogatva az oktatókat és a hallgatókat. A hangsúly tehát azon van, hogy az egyetemek fenntartható módon legyenek a tudományos érintkezés központjai. Mint az élet minden más területén, a technológia fontos szerepet játszik ebben az átalakulásban is.

Manapság az információ megosztása és fogyasztása is túlnyomóan online történik, különböző szolgáltatásokat használva (pl. email és üzenetküldés, chat, file küldés és fogadás, videokonferencia), de továbbra is a legnépszerűbb internetes szolgáltatás a World Wide Web. A világháló tradicionális felhasználása az emberi fogyasztók és a tartalomszolgáltatók összekapcsolásán alapul. Az adatok a legtöbb weboldalon a legjobb esetben is csak részben strukturáltak [1] és a weboldalak nagy százaléka különböző emberi felhasználásra szánt

felhasználói felületek által „el vannak rejtve” a gépi feldolgozás előtt [6]. Mindez új ajánlások létrehozását eredményezte strukturált adatok webes publikálására és összekapcsolására. Az ilyen formában létrejövő adatokat Kapcsolt Adatoknak⁵ hívjuk.

Az okos egyetemek egyik fontos vonatkozása a tanulás és tanítás támogatása modern technológiákkal. Az egyetemi kampuszokon megforduló emberek információs igénye jelentős átfedést mutat [8]. Ezért nagyon fontos létrehozni egy közös adatmodellt, ami lehetővé teszi a különböző heterogén forrásokból származó adatok összekapcsolását és ezáltal a felmerülő információs igények emelt szintű kiszolgálását az összekapcsolt adatokon alapuló új alkalmazások révén.

⁵ <https://www.w3.org/standards/semanticweb/data>

4 Kutatási eredmények

A disszertáció azt a kérdést vizsgálja, hogy a megfelelő információ megosztásával hogyan tudnak a 21. században az egyetemek lépést tartani az új igényekkel, egyfajta “okos egyetemként” működve. Ebben a fejezetben minden kutatási eredmény röviden bemutatásra kerül.

4.1 *Automatikus URI generálás Kapcsolt Adathalmazokhoz*

A Kapcsolt Adathalmazok erőforrásait egységes erőforrás azonosítók, URI-k segítségével azonosítjuk. Fontos lépés minden Kapcsolt Adathalmazokkal foglalkozó munka során definiálni az URI hozzárendelések szabályait. Ugyan széles körben elfogadott vélekedés, hogy az URI-k ne hordozzanak semmilyen jelentést, a gyakorlatban hasznosnak bizonyul, ha ember által is olvashatóak, így egy hosszabb listából az emberi szem is ki tudja választani a keresett erőforráshoz tartozó URI-t.

Ha az erőforrásoknak természetes módon adódik azonosítója, vagy ezen azonosítók már valamilyen úton adottak, akkor a kihívás az ismétlődő azonosítók elkerülése. Ha az erőforrások nem rendelkeznek egyértelmű inverz-funkcionális tulajdonsággal, ami egyszerű azonosítási lehetőséget adna, csak összetett kulccsal lehet az ilyen

jellegű erőforrásokat azonosítani. Ilyen erőforrásokra jó példa az ismétlődő események sorozata, ahol az egyes események nincsenek egymástól jól megkülönböztetve, így ezeket csak a pontos időpontjuk és helyszínük alapján lehet azonosítani. Az RDF (Resource Description Framework)⁶ lehetőséget ad ilyen esetekben az azonosítók teljes elhagyására Blank Node-ok használatával, de ez további problémákat okozhat, mivel a létrehozott erőforrások közvetlenül nem hivatkozhatók.

Emberi felhasználásra – különösen a kézi adatbevitel támogatására – kiemelkedően fontos, hogy valamilyen szemantikát adjunk az URI tartalmának. Ugyanakkor ez redundanciát okoz, hiszen az URI-ba kódolt információnak kapcsolódó entitások formájában is meg kell jelennie. Ha a redundancia nem elkerülhető, valamilyen automatikus mechanizmusnak biztosítani kell az adatok ellenőrzését és/vagy karbantartását.

A Kapcsolt Adatok feldolgozásához és főként a megfelelő lekérdezések végrehajtásához szükséges lehet az erőforrások teljes szemantikus leírása. A Kibontható URI-k ezen esetek kezelésére javasoltak: az ilyen URI-k egy előre meghatározott mintát követnek, tartozik hozzájuk egy vagy több sablon, ami leírja a struktúrájukat.

Sok olyan eset létezik, amikor gyakori RDF hármashalmazokat kell létrehozni mindenféle egyszerű módon definiálható azonosítók nélkül.

⁶ <https://www.w3.org/RDF/>

Ilyen esetben az egyik megoldás egy véletlenszerűen hozzárendelt azonosító, míg az általunk publikált [12] másik módszer ezeknek az eseteknek a kezelését egyszerűsíti le, illetve további sokrétű lekérdezési és következtetési lehetőséget ad, standard SPARQL⁷ felhasználásával.

Tegyük fel, hogy adott egy Kapcsolt Adathalmaz és egy mechanizmus (pl. trigger), ami egy adott típusú új egyed megjelenése esetén további RDF hármast ad az adathalmazhoz, amelyek szemantikusan az eredetiből származnak, ahhoz kapcsolódnak. Ez a mechanizmus leegyszerűsíti az RDF tárolók karbantartását azok lényegi módosítása nélkül.

Négy típuspéldán keresztül demonstráltuk az automatikusan kibontható URI-k felhasználási lehetőségeit:

1. összetett kulccsal azonosítható egyedek
2. Blank Node (anonim erőforrás) jelöltek
3. URI alapján történő kapcsolódó egyedek összekapcsolása
4. következtetés URI-ban megjelenő információk segítségével

Fontos része a kibontható URI-k körüli folyamatoknak a különböző sablonok, amelyek azokat a szabályokat reprezentálják, amelyek eldöntik, hogy milyen új információ származtatható az adott egyed

⁷ <https://www.w3.org/TR/rdf-sparql-query/>

URI-jából. Ezek a sablonok az adatokon történő lekérdezések konzisztenciáját hivatottak biztosítani. Egy sablon egy URI mintát és a hozzá tartozó generált RDF leírás struktúráját adja meg.

A javasolt URI séma a Mátrix URI-n alapul, amely képes több egymástól független paraméter segítségével az erőforrást. Követve a Mátrix URI sémát⁸, az ilyen URI az alap résszel kezdődik, amelyet az erőforrás típusa követ, majd az összetett kulcsának elemei kerülnek felsorolásra.

1. Tézis

A Kibontható URI-k egy bizonyos mintát követő URI-k és egy hozzájuk tartozó sablon párosa, amelyek egy új mechanizmus segítségével automatikusan RDF hármassokat generálnak az URI-kban található strukturált információ segítségével. Támogatják a kézi RDF adatomódosítást azáltal, hogy csökkentik a komplexitást és növelik a konzisztenciát, valamint kiküszöbölik a generált URI-k összeakadását. A Kibontható URI-k sikeresen kerültek alkalmazásra az Óbudai Egyetem OLOUD szerverén, valamint példák segítségével további alkalmazási lehetőségek kerültek bemutatásra [12].

⁸ <http://www.w3.org/DesignIssues/MatrixURIs.html>

4.2 Holisztikus Kapcsolt Adat publikálás a felsőoktatásban

A magyar felsőoktatásban az egyetemi információs rendszerek általában egymástól valamennyire elkülönülten működnek. Léteznek kutatás és oktatás menedzsment, e-tanulás, illetve egyéb jellegű szolgáltatások, de az ezek közötti integráció nem teljes körű: jelenleg nincsen a magyar felsőoktatásban olyan integrált adatbázis, amely az összes kapcsolódó adatot tárolni és kiszolgálni tudná. A meglévő adatforrások pedig többnyire zárt rendszert alkotnak: az adat jellegét tekintve nyilvános volna, de a nyilvános hozzáférés nem biztosított. Adatok a kurzusokról, hallgatókról és oktatókról különböző egymástól elválasztott rendszerekben vannak kezelve, amelyek között a kapcsolat többnyire nem létezik explicit módon. Ahhoz, hogy adatokban lévő implicit kapcsolatokhoz hozzáférjünk, érdemes a fontosabb fogalmakat és kapcsolatait formális ontológiával modellezni. Kapcsolt Adatok publikálása igényli a felhasznált fogalmak egy (vagy több) sémában történő leírását. Ontológiák felhasználása további explicit szemantikát ad a sémához és lehetővé teszi az adatokon történő következtetést, így az implicit kapcsolatokat explicitté teszi.

Az Ontology for Linked Open University Data (OLOUD) egy gyakorlati megközelítése a kurzusokkal kapcsolatos információk modellezésének a magyar felsőoktatásban. A tárgyterület összetettsége miatt egyetlen ontológia túlságosan bonyolult lenne az

összes felhasználási eset lefedésére. Több ontológia is létezik a tárgyterület egyes részeinek a leírására, így kiválasztottuk azokat, amelyek együtt ki tudják elégíteni az összes követelményünket. Összefoglaltuk a különböző kihívásokat és a rájuk adott válaszokat, mint például hiányzó kapcsolatok fogalmak között, inkonzisztens vagy egymással redundáns, átfedésben lévő ontológiák. A munka eredményeként az ontológia felhasznál és integrál különböző forrásokból származó fogalmakat, támogatva a személyes időbeosztással, navigációval, illetve egyéb szükségletekkel kapcsolatos adatok fogyasztását és publikálását. Többek között a következő tárgyterületek kerültek modellezésre: tanterv, tantárgy, kurzus, szemeszter, személyzet, épületek és események.

Az OLOUD ontológia ragasztóként köt össze és egészít ki meglévő ontológiákat, így egy egységes modellt ad a magyar felsőoktatás fontos fogalmainak a leírására és lehetővé teszi a széles körű SPARQL lekérdezésekkel megfogalmazott egyetemi igények kielégítését Kapcsolt Adatokkal.

Cikkeinkben ([2], [9] és [10]) további részleteket közöltünk az ontológiáról.

2. Tézis

Az új Ontology for Linked Open University Data (OLOUD) egy holisztikus modellt szolgáltat Kapcsolt Adatok publikálásához a felsőoktatás szakterületén belül [2]. A rendelkezésre álló, a

tárgyterületet lefedő ontológiák elemzése alapján az OLOUD hiánypótló szerepet tölt be, kiegészítve a meglévő ontológiák hiányosságait, teljessé téve a felsőoktatás szemantikus modelljét. Az Óbudai Egyetem sikeresen publikált több Kapcsolt Adathalmazt felhasználva az OLOUD fogalmait.

4.3 Ontológia szemantikus belterek leírásához

Az iLOC ontológia belterek leírását és annak Kapcsolt Nyílt Adathalmazként történő publikálását teszi lehetővé. Ilyen adatok felhasználásával beltéri navigációs szolgáltatás készíthető, ha az ontológia fogalmait felhasználva az épületekről szemantikus modell készül azok strukturális és geometrikus leírása (pl. alaprajzuk) alapján. Lehetővé teszi azt is, hogy két pont között különböző megszorításoknak is megfelelő (pl. kerekesszékekkel járható) útvonalat, illetve annak hosszát is kiszámoljuk. Jelenleg nem ismert más beltéri navigációt biztosító szolgáltatás, amely Kapcsolt Adatok felhasználásával működik és nincsen olyan elérhető ontológia sem, amellyel ilyen módon leírhatók lennének épületek. Más, hasonló célt megvalósító megoldások alapvetően különböznek a modellezés költségében, illetve a felhasználás jellegében, valamint a felhasznált fogalmakban.

Az általam javasolt Kapcsolt Adatmodell lehetővé teszi többféle épületekhez kapcsolódó adathalmaz összekapcsolását és nem csak

egyetemi környezetben használható, hanem sikerrel alkalmazható más nyilvános helyszínek, mint múzeumok, bevásárló központok, kórházak, stadionok, valamint repülőterek leírására is. Az iLOC ontológiával megvalósított épületleírások más ontológiákkal történő integrálása a navigációs lekérdezések különböző kiegészítését teszi lehetővé ezen ontológiák fogalmainak felhasználásával. Ilyen kiegészítés lehet a helyiségek kategorizálása, illetve az útvonalak hozzáférési lehetőségeinek, illetve megszorításainak a leírása. A felhasználási lehetőségek között szerepel közeli helyek leírása és megtalálása azok tulajdonságai alapján, útvonal tervezés és navigálás akár vészhelyzeti körülmények között is. Az ontológia három különböző részletességű épület leírási lehetőséget biztosít, amelyek egyre növekvő részletességgel és pontossággal írják le egyre több részletét az épületnek. Worboys [13] osztályozása alapján az iLOC egy hibrid ontológia, ami támogatja a részek, a köztük lévő kapcsolatok és távolságuk leírását is. Az alkalmazott topológiai leírás a legpontosabb a lépésről-lépésre történő navigációhoz, amely kiegészíthető szemantikus leírással (pl. a helyiség típusa, a hozzáférhetőség módja).

Cikkeinkben ([3], [4], [5] és [11]) bemutattuk az iLOC ontológiát és egy általános módszert beltéri útkeresésre és helyszínek leírására.

3. Tézis

Az iLOC ontológiát beltérek szemantikus leírására hoztam létre.

Elsőként mutattam meg a gyakorlatban, egy az Óbudai

Egyetemen készült mobil alkalmazás és az iLOC ontológia segítségével, hogy standard SPARQL lekérdezések felhasználásával személyre szabható beltéri útvonalak számíthatók (nyílt) kapcsolt adatokból. Szerzőtársaimmal példa SPARQL lekérdezéseket és különböző SPARQL implementációkon végzett teljesítmény méréseket is publikáltunk: [3], [4], [5] és [11]

4.4 Beltéri navigáció személyes preferenciák, képességek és jogosultságok alapján

Napjainkban a kültéri navigáció széleskörben elterjedt szolgáltatás, ami nagy segítségünkre van autóvezetés, séta vagy akár a tömegközlekedés használata során is. Ez a fajta navigáció általában a GPS vagy a Glonass rendszer jeleit használva koordinátákban adja meg a pozíciót. Épületek belsejében ugyanakkor egy navigációs rendszernek sokkal bonyolultabb útvonalakkal kell dolgoznia, amelyek folyosókon, felvonókon, illetve lépcsőkön vezetnek keresztül, miközben általában a GPS vagy Glonass jelek nem érhetőek el. A felhasználók korlátozottak lehetnek képességeiket, illetve hozzáférési jogosultságaikat illetően, valamint személyes preferenciáik sokkal széleskörűbbek lehetnek. A jelenleg elérhető megoldások általában valamilyen speciális és drága hardveres megoldáson alapulnak, amely segítségével a pontos hely

megállapítható ugyan, de ellentétben a kültéri navigációval, nincsen egy elterjedt alapvető navigációs módszer.

Ebből az okból hoztuk létre az iLOC ontológiát, amely a belterekben történő navigációt támogatja. Az ontológia segítségével belterek szemantikus leírása lehetséges, összekapcsolva azt már meglévő más adatforrásokkal. Képes leírni azt, hogy két helyszín hogyan kapcsolódik egymáshoz, azok távolságával és a köztük lévő útvonal egyéb tulajdonságaival (pl. alkalmas-e kerekesszékekkel történő közlekedésre vagy szükséges-e bármilyen hozzáférést biztosító kártya). A követelmények része volt az alacsony implementációs költség, valamint a széleskörű felhasználási lehetőség, ami a speciális eszközigénnyel rendelkező szolgáltatások legfőbb problémája. A használatához csak egy átlagos okostelefonra vagy tabletre van szükség.

Az iLOC ontológia három egymásra épülő, növekvő bonyolultságot és pontosságot biztosító navigációs módot támogat. Amíg az első mód csupán tájékoztatja a felhasználót, hogy a keresett helyiség melyik épület melyik részén és szintjén található, addig a második képes az épületen keresztül vezető egyes lépéseket is megadni az adott helyszín eléréséhez. A harmadik, legpontosabb mód képes az egyes lépések hosszát és esetleges korlátozásait is megadni.

Az ontológia támogatja útvonalak részletes leírását, az egyes részekre vonatkozó korlátozásokat, egyirányúságot, valamint a köztes lépések

hosszára vonatkozó metrikákat. Az útvonalak így rugalmasan kereshetőek, figyelembe véve a felhasználó célját és képességeit.

Cikkeinkben ([3], [4], [5] és [11]) elsőként mutattuk meg, hogy standard SPARQL lekérdezések és az iLOC ontológia segítségével személyre szabott útvonalak számíthatók ki Kapcsolt (Nyílt) Adathalmazokon.

4. Tézis

Mozgásukban korlátozott, vagy más különleges igénnyel rendelkező emberek számára a nyilvános helyeken történő navigáció kihívásokat jelenthet. Az iLOC ontológia kiegészítése támogatja a személyes preferenciák, képességek és jogosultságok alapján történő beltéri útvonaltervezést. [4] és [5]

4.5 Belterek modellezése az iLOC ontológia segítségével

Ha valakitől útbaigazítást kérünk, általában a kapott leírás szimbolikus és absztrakt elemeket fog tartalmazni, különböző, könnyen azonosítható helyekre és jelekre hivatkozva. Ahhoz, hogy jól működő beltéri navigációs élményt biztosítson egy szolgáltatás, ezt az emberi tájékozódási módot be kell építeni az útvonalak leírásába.

Többféle megközelítés létezik beltérek modellezésére, az így létrejött modellek alapvetően különböznek egymástól felhasználásukban, a modellezés és a felhasználás költségében. A modellek adhatnak

strukturális, geometrikus, valamint topologikus leírást. A strukturális leírás a helyiségeket és egyéb fogalmakat egy hierarchiában kapcsolja össze, ezáltal nem adva megfelelő részletes leírást beltéri navigációhoz, ugyanakkor a geometrikus leírás túlságosan bonyolult, beltéri navigációhoz történő felhasználása problémás lehet.

Az iLOC ontológia beltérek topologikus leírásához ad fogalmakat, ami a szemantikus leírás legfontosabb része, az ontológia segítségével ismert, egymással hálózatba kapcsolt pontok halmazát lehet leírni. A szemantikus leírás a geometriai és a topológiai leírásból származtatható, azok bizonyos elemeit elhagyva. A beltéri navigációhoz szükséges a könnyen azonosítható pontokat (POI-kat) hálózatba szervezni. Az útvonalak ugyanis ilyen POI-kon keresztül fognak a célhoz elvezetni.

Az alaprajzot felhasználva, nagyon fontos, hogy ezek a pontok módszeresen azonosításra kerüljenek, biztosítva ezzel az ideális lefedettséget. Mivel ez a folyamat komplex és sok hibalehetőséggel járhat, érdemes a cikkünkben [3] közzétett módszert követni, amely javaslatot tesz az épület osztályozására, nagyobb terek (pl. folyosók, előterek) kisebbekre bontására, a szükséges POI-k azonosítására és hálózatba szervezésére, valamint a többszintes alaprajzok kihívásaira. Példákat hoz különleges esetekre, mint a több szinten elhelyezkedő termek. Kisebbs helyiségeket, amelyeknek csak egyetlen bejárata van, másképpen kell kezelni, mint a nagyobb tereket akár több bejárattal. A terem típusa is befolyásolja, hogy milyen fajta és számú pontot

érdeemes azonosítani. A pontok azonosításán túlmenően fontos a köztük lévő útvonalakat is megvizsgálni és minden speciális esetet (pl. egyirányú útvonalak mozgólépcsőkön, vagy másként korlátozott helyek), illetve az útvonal egyéb tulajdonságait rögzíteni. Az így létrejövő szemantikus térkép minősége erőteljesen függ a létrehozása során alkalmazott módszertől.

Cikkünkben [3] közzétettünk egy módszertant is olyan beltéri modellek kialakítására, amelyekkel beltérek ismert pontjai (POI) kapcsolhatók hálózatba és ezen pontok között optimális útvonalak számíthatók.

5. Tézis

Összettel alaprajzokból szemantikus modellt készíteni sok kihívással járhat, amelyek elemzése alapján egy módszertant alkottam, amely megmutatja a szemantikus beltér modellezés egyes lépéseit [3].

References

- [1] Abiteboul, Serge. "Querying semi-structured data." International Conference on Database Theory. Springer, Berlin, Heidelberg, 1997.
- [2] Fleiner, Rita, Barnabás Szász, and András Micsik. "OLOUD-An Ontology for Linked Open University Data." Acta Polytechnica Hungarica 14.4. 2017. pp. 63-82.
- [3] Fleiner, Rita, Barnabás Szász, and Péter Piros. "Indoor navigation Linked Data at Obuda University." Applied Computational Intelligence and Informatics (SACI), 2016 IEEE 11th International Symposium on. IEEE, 2016.
- [4] Fleiner, Rita, Barnabás Szász, Gabriella Simon-Nagy, and András Micsik. "Indoor navigation for motion disabled persons in medical facilities." Acta Polytechnica Hungarica 14, no. 1. 2017. pp. 111-128.
- [5] Fleiner, Rita, Gabriella Simon-Nagy, and Barnabás Szász. "Accessible indoor navigation based on linked data in hospitals." Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on, IEEE, 2016. pp. 2425-2430.

- [6] He, Bin, Mitesh Patel, Zhen Zhang, and Kevin Chen-Chuan Chang. "Accessing the deep web." *Communications of the ACM* 50.5. 2007. pp. 94-101.
- [7] Janowicz, Krzysztof, Hitzler, Pascal, Adams, Benjamin, Kolas, Dave, and Vardeman II, Charles. "Five stars of linked data vocabulary use." *Semantic Web* 5.3. 2014. pp. 173-176.
- [8] Rohs, Michael, and Jürgen Bohn. "Entry points into a smart campus environment-overview of the ETHOC system." *Distributed Computing Systems Workshops, 2003. Proceedings. 23rd International Conference on Distributed Computing Systems Workshops, 2003.* pp. 260-266. IEEE, 2003.
- [9] Szász, Barnabás, Rita Fleiner, and András Micsik. "Practical uses of location and event data as Linked Open University Data." *FutureRFID 2014. Eger, Eszterházy Károly Főiskola.* 2014. pp. 151-159.
- [10] Szász, Barnabás, Rita Fleiner, and András Micsik. "A case study on Linked Data for University Courses." *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems.* Springer, Cham, 2016.
- [11] Szász, Barnabás, Rita Fleiner, and András Micsik. "iLOC—Building In-door Navigation Services using Linked Data." *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems -*

SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESS'16) 12-15 September 2016 Leipzig, Germany. CEUR, 2016.

- [12] Szász, Barnabás, Rita Fleiner, and András Micsik. "Linked data enrichment with self-unfolding URIs." Applied Machine Intelligence and Informatics (SAMi), 2016 IEEE 14th International Symposium on, pp. 305-309. IEEE, 2016.
- [13] Worboys, Michael. "Modeling indoor space." Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness. ACM, 2011. pp. 1-6.



Registry number: DEENK/160/2019.PL
Subject: PhD Publikációs Lista

Candidate: Barnabás Szász
Neptun ID: T76821
Doctoral School: Doctoral School of Informatics
MTMT ID: 10060289

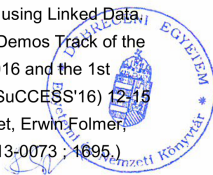
List of publications related to the dissertation

Foreign language Hungarian book chapters (1)

1. **Szász, B.**, Fleiner, R., Micsik, A.: Practical uses of location and event data as Linked Open University Data.
In: Proceedings of the 1st International Conference and Exhibition on Future RFID Technologies. Eds.: Emőd Kovács, Gábor Kusper, Tibor Juhász, Tibor Tómacs, [EKTF], [Eger], 151-159, 2014.

Foreign language international book chapters (5)

2. **Szász, B.**, Fleiner, R., Micsik, A.: A case study on Linked Data for university courses.
In: On the Move to Meaningful Internet Systems : OTM 2016 Workshops : Confederated International Workshops: EI2N, FBM, ICSP, Meta4eS, and OTMA 2016 Rhodes, Greece, October 24-28, 2016 Revised Selected Papers. Eds.: Ioana Ciuciu, Christophe Debruyne, Hervé Panetto, Georg Weichhart, Peter Bollen, Anna Fensel, Maria-Esther Vidal, Springer, Cham, 265-276, 2018, (Lecture Notes in Computer Science, ISSN 0302-9743 ; 10034.)
ISBN: 9783319559605
3. Fleiner, R., Simon-Nagy, G., **Szász, B.**: Accessible indoor navigation based on Linked Data in hospitals.
In: 2016 IEEE International Conference on Systems, Man, and Cybernetics : Proceedings, IEEE, Piscataway, 2425-2430, 2016. ISBN: 9781509018970
4. **Szász, B.**, Fleiner, R., Micsik, A.: iLOC: building indoor navigation services using Linked Data.
In: SEMANTICS Leipzig 2016 : Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTICS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) 12-15 September 2016 Leipzig, Germany. Eds.: Michael Martin, Marti Cuquet, Erwin Folmer, CEUR, Leipzig, [1-4], 2016, (CEUR Workshop Proceedings, ISSN 1613-0073 ; 1695.)
5. Fleiner, R., **Szász, B.**, Piros, P.: Indoor navigation Linked Data at Óbuda University.
In: SACI 2016 11th IEEE International Symposium on Applied Computational Intelligence and Informatics : Proceedings : May 12-14, 2016 Timișoara, Romania. Ed.: Anikó Szakál, IEEE Computer Society, Piscataway, 25-30, 2016. ISBN: 9781509023790





6. **Szász, B.**, Fleiner, R., Micsik, A.: Linked Data enrichment with self-unfolding URIs.

In: SAMI 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics : Proceedings : January 21-23, 2016, Herľany, Slovakia, IEEE Computer Society, Piscataway, 305-309, 2017. ISBN: 9781467387408

Foreign language scientific articles in Hungarian journals (2)

7. Fleiner, R., **Szász, B.**, Simon-Nagy, G., Micsik, A.: Indoor navigation for motion disabled persons in medical facilities.

Acta Polytech. Hung. 14 (1), 111-128, 2017. ISSN: 1785-8860.

DOI: <http://dx.doi.org/10.12700/APH.14.1.2017.1.8>

IF: 0.909

8. Fleiner, R., **Szász, B.**, Micsik, A.: OLOUD: an ontology for Linked Open University Data.

Acta Polytech. Hung. 14 (4), 63-82, 2017. ISSN: 1785-8860.

DOI: <http://dx.doi.org/10.12700/APH.14.4.2017.4.4>

IF: 0.909

Foreign language conference proceedings (1)

9. **Szász, B.**, Saraniva, A., Bognár, K., Unzeitig, M., Karjalainen, M.: Cultural heritage on the semantic web: the Museum24 project.

In: Digital Semantic Content across Cultures. Eds.: Eero Hyvönen, Pekka Silvennoinen, Francis Jutand, Semantic Computing Research Group (SeCo), Paris, [1-10], 2006.





List of other publications

Foreign language international books (1)

10. Guta, G., **Szász, B.**, Schreiner, W.: A lightweight model driven development process based on XML technology. Research Institute for Symbolic Computation, Linz, 29 p., 2008. (RISC Report Series; No. 3.)

Foreign language Hungarian book chapters (1)

11. Demeter, J., **Szász, B.**: Data mining based on medical diagnosis.
In: Proceedings of the 7th International Conference on Applied Informatics, Eger, Hungary, January 28-31, 2007.. Szerk.: Emőd Kovács, Péter Olajos, Tibor Tómacs, BVB Nyomda és Kiadó, Eger, 261-267, 2007. ISBN: 9789630693417

Total IF of journals (all publications): 1,818

Total IF of journals (publications related to the dissertation): 1,818

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

18 April, 2019





Nyilvántartási szám: DEENK/160/2019.PL
Tárgy: PhD Publikációs Lista

Jelölt: Szász Barnabás
Neptun kód: T76821
Doktori Iskola: Informatikai Tudományok Doktori Iskola
MTMT azonosító: 10060289

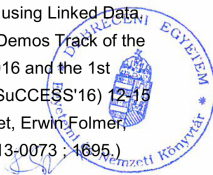
A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű, hazai könyvrészletek (1)

1. **Szász, B.**, Fleiner, R., Micsik, A.: Practical uses of location and event data as Linked Open University Data.
In: Proceedings of the 1st International Conference and Exhibition on Future RFID Technologies. Eds.: Emőd Kovács, Gábor Kusper, Tibor Juhász, Tibor Tómacs, [EKTF], [Eger], 151-159, 2014.

Idegen nyelvű, külföldi könyvrészletek (5)

2. **Szász, B.**, Fleiner, R., Micsik, A.: A case study on Linked Data for university courses.
In: On the Move to Meaningful Internet Systems : OTM 2016 Workshops : Confederated International Workshops: EI2N, FBM, ICSP, Meta4eS, and OTMA 2016 Rhodes, Greece, October 24-28, 2016 Revised Selected Papers. Eds.: Ioana Ciuciu, Christophe Debruyne, Hervé Panetto, Georg Weichhart, Peter Bollen, Anna Fensel, Maria-Esther Vidal, Springer, Cham, 265-276, 2018, (Lecture Notes in Computer Science, ISSN 0302-9743 ; 10034.)
ISBN: 9783319559605
3. Fleiner, R., Simon-Nagy, G., **Szász, B.**: Accessible indoor navigation based on Linked Data in hospitals.
In: 2016 IEEE International Conference on Systems, Man, and Cybernetics : Proceedings, IEEE, Piscataway, 2425-2430, 2016. ISBN: 9781509018970
4. **Szász, B.**, Fleiner, R., Micsik, A.: iLOC: building indoor navigation services using Linked Data.
In: SEMANTICS Leipzig 2016 : Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTICS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) 12-15 September 2016 Leipzig, Germany. Eds.: Michael Martin, Marti Cuquet, Erwin Folmer, CEUR, Leipzig, [1-4], 2016, (CEUR Workshop Proceedings, ISSN 1613-0073 ; 1695.)
5. Fleiner, R., **Szász, B.**, Piros, P.: Indoor navigation Linked Data at Óbuda University.
In: SACI 2016 11th IEEE International Symposium on Applied Computational Intelligence and Informatics : Proceedings : May 12-14, 2016 Timișoara, Romania. Ed.: Anikó Szakál, IEEE Computer Society, Piscataway, 25-30, 2016. ISBN: 9781509023790





6. **Szász, B.**, Fleiner, R., Micsik, A.: Linked Data enrichment with self-unfolding URIs.
In: SAMI 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics : Proceedings : January 21-23, 2016, Herl'any, Slovakia, IEEE Computer Society, Piscataway, 305-309, 2017. ISBN: 9781467387408

Idegen nyelvű tudományos közlemények hazai folyóiratban (2)

7. Fleiner, R., **Szász, B.**, Simon-Nagy, G., Micsik, A.: Indoor navigation for motion disabled persons in medical facilities.
Acta Polytech. Hung. 14 (1), 111-128, 2017. ISSN: 1785-8860.
DOI: <http://dx.doi.org/10.12700/APH.14.1.2017.1.8>
IF: 0.909
8. Fleiner, R., **Szász, B.**, Micsik, A.: OLOUD: an ontology for Linked Open University Data.
Acta Polytech. Hung. 14 (4), 63-82, 2017. ISSN: 1785-8860.
DOI: <http://dx.doi.org/10.12700/APH.14.4.2017.4.4>
IF: 0.909

Idegen nyelvű konferencia közlemények (1)

9. **Szász, B.**, Saraniva, A., Bognár, K., Unzeitig, M., Karjalainen, M.: Cultural heritage on the semantic web: the Museum24 project.
In: Digital Semantic Content across Cultures. Eds.: Eero Hyvönen, Pekka Silvennoinen, Francis Jutand, Semantic Computing Research Group (SeCo), Paris, [1-10], 2006.





További közlemények

Idegen nyelvű, külföldi könyvek (1)

10. Guta, G., **Szász, B.**, Schreiner, W.: A lightweight model driven development process based on XML technology. Research Institute for Symbolic Computation, Linz, 29 p., 2008. (RISC Report Series; No. 3.)

Idegen nyelvű, hazai könyvrészletek (1)

11. Demeter, J., **Szász, B.**: Data mining based on medical diagnosis.
In: Proceedings of the 7th International Conference on Applied Informatics, Eger, Hungary, January 28-31, 2007.. Szerk.: Emőd Kovács, Péter Olajos, Tibor Tómacs, BVB Nyomda és Kiadó, Eger, 261-267, 2007. ISBN: 9789630693417

A közlő folyóiratok összesített impakt faktora: 1,818

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):
1,818**

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2019.04.18.

