International Journal on Advances in Intelligent Systems, vol 7 no 3 & 4, year 2014, http://www.iariajournals.org/intelligent_systems/

Modelling Spatial Understanding: Using Knowledge Representation to Enable Spatial Awareness and Symbol Grounding in a Robotics Platform

Martin Lochner, Charlotte Sennersten, Ahsan Morshed, and Craig Lindley
CSIRO Computational Informatics (CCI) Autonomous Systems (AS)
Commonwealth Scientific and Industrial Research Organization (CSIRO)
Hobart, Tasmania, Australia
Contact: martin.lochner@csiro.au, charlotte.sennersten@csiro.au,
ahsan.morshed@csiro.au, craig.lindley@csiro.au

***Abstract*—Robotics in the 21st century will progress from scripted interactions with the physical world, where human programming input is the bottleneck in the robot's ability to sense, think and act, to a point where the robotic system is able to autonomously generate adaptive representations of its surroundings, and further, to implement decisions regarding this environment. A key factor in this development will be the ability of the robotic platform to understand its physical space. In this paper, we describe a rationale and framework for developing spatial understanding in a robotics platform, using knowledge representation in the form of a hybrid spatial-ontological model of the physical world. Further, we describe the proposed CogOnto (cognitive ontology) model, which enables symbol grounding for a cognitive computing system, using sensor data gathered from diverse and heterogeneous sources, associated with humanly crafted symbolic descriptors. While such a system may be implemented with classical ontologies, we discuss the advantages of non-hierarchical modes of knowledge representation, including a conceptual link between information processing ontologies and contemporary cognitive models.**

Keywords-Human Robot Interaction; Artificial Intelligence; Autonomous Navigation; Knowledge Representation; Symbol Grounding; Spatial Ontology.

I. INTRODUCTION

The process of transitioning away from hard-coded robotics applications, which carry out highly pre-determined actions such as the traditional manufacturing robot, is already well underway. This paper follows our previous work [1] in which we describe a methodology for using ontological data representation to encode 3D spatial information in robotics applications. With notions such as cloud robotics [2] entering the *zeitgeist*, and highly publicized events such as the Defense Advanced Research Projects Agency (DARPA) Robotics Challenge (Dec. 19-21, 2013, Miami FL) bringing public attention to these advances, it is foreseeable that robots will be entering the mainstream realm of human activity – more than in fringe applications (robotic vacuum cleaner; children's toys), but in key areas such as caring for the aged [3], operating vehicles [4], disaster management [5], and undertaking autonomous scientific investigation [6].

The hurdles that must be overcome in reaching these goals, however, are neither few nor small. This can be plainly seen, for example in the aforementioned 2013 Robotics Challenge, in which simple spatial tasks that are routine for a human being (open a door, climb a ladder) are still critically difficult for even the most advanced and highly funded robotics projects. While the state-of-the-art is impressive, it is evident that physical robotics hardware is far in advance of the control systems that are in place to guide the robot. The challenge is, thus, to develop systems whereby a robot can perceive a physical space and understand its position in that space, the components that exist within the space, and how it can or *should* interact with these components in order to achieve implicit or explicit goals. This is furthermore impacted by the requirement that robotic systems be able to operate in outdoor environments where distributed connections may not be available; however, describing the development of long-range data networks for robotic communication is beyond the scope of this paper.

While there are a number of ways that the problem of providing a robot with a spatial understanding can be approached (e.g., neuro-fuzzy reasoning [7], dynamic spatial relations via natural language [8]) it is our proposition that leveraging the current advancements in knowledge representation via ontologies [9][10], in combination with an understanding of human spatial-cognitive processing [11][12], and enabled by real-time scene modeling [13] will provide a powerful and accessible methodology for enabling spatial understanding and interaction in a mobile robotics platform. As argued by Sennersten et al. [14], the advantage of using cloud-based repositories of perceptual data annotated with ontology and metadata information is to take advantage of humanly-tagged examples of sense data (e.g., images) to overcome the symbol grounding problem. Symbol grounding refers to the need for symbolic structures to have valid associations with the things in the world that they refer to. Achieving symbol grounding is an ongoing challenge for robotics and other intelligent systems [15]. Using cloud-based annotations attached to sensory exemplars takes advantage of the human ability to ground symbols, obviating the need for robots to achieve this independently of human symbolic expressions.

This paper provides a conceptual overview of how spatial understanding can be developed in a robotics platform. We discuss traditional knowledge representation (classical information processing ontologies), describe the development and use of "cognitive" ontologies, and how this may be transitioned into the development of a physical-spatial ontology, including a possible system of comprehension for spatial position. Finally, we discuss the notion that truly non-hierarchical systems such as complex chemical structure, and such as the human cortex, may require the development of systems of knowledge representation that transcend the structural limits of today's systems.

II. STATE OF THE ART: KNOWLEDGE REPRESENTATION

The development of specific nomological hierarchies for concept representation is currently taking place across many fields of academic endeavor (e.g., genetics, medicine, neuroscience, biology, chemistry, physics). Under the guise of the philosophical concept of an *Ontology*, such applications seek to outline the knowledge, which exists within a domain at three levels of representation: Classes, Properties, and Relationships. These nomological hierarchies provide a way of describing the precise relationship that terms in a given domain have to one another. As an information processing construct, the definition of an ontology is refined as an "explicit formal specification of the terms in the domain and relations among them", or more concisely, "a specification of a conceptualization" [16].

A system that operates with such knowledge representation within its core functionality may be considered to be 'knowledge-based'. A knowledge-based system is a computer program that stores knowledge about a given domain (also known as an "expert system", when the knowledge is considered to be from a highly specialized domain). However, an ontology does not intrinsically represent the kinds of truth-functional mappings or procedures captured by rules in more complete knowledge bases. Hence, an ontology provides classifications and the ability to infer associations via subclass/superclass relationships. More complex forms of reasoning required for most forms of useful cognitive task performance require task-oriented rules. As such, the domain knowledge in a knowledge base includes ontology representations, while most task-oriented reasoning is achieved by the use of rules that refer to ontological constructs in the form of domains within rule tuples.

The system attempts to mimic the reasoning of a human specialist by conducting reasoning across rules and in reference to a database of atomic facts. Matching sense data against metadata/ontology-annotated sense data on the web can provide a method of automatically mapping a current sensed situation to the annotations of past situations stored in the cloud. This allows the system to retrieve representations of the situation in an atomic form, as statements formulated using the symbolic forms of annotations, which are retrieved by matching against associated sense data. Ontologies hold the potential, therefore, to provide the constructs for symbolic atomic fact expressions that rule-sets can then process for automated cognitive task performance.

A. *Cognitive Ontologies*

An increasing number of ontologies are available on-line that can potentially support this symbolic structure generation process. Knowledge representation via ontological structure has been applied to the field of cognitive science, both in relation to terminology used within the domain (e.g., DOLCE - Descriptive Ontology for Linguistic and Cognitive Engineering [17][18]) and for concepts relevant to empirical testing paradigms (e.g., CogPo [19]). Indeed, several cognitive ontologies have been developed in the recent years, including DOLCE, WordNet [20], CYC [21], and CogPo.

WordNet is an online lexical knowledgebase system, whose design is inspired by current psycholinguistic theories of human lexical memory, where each cognitive artifact can be semantically classified into English nouns, verbs, and adjectives, with different meanings and relationships in real-world scenarios. DOLCE is developed by Nicola Guarino and his associates at the Laboratory for Applied Ontology (LOA) [22]. It captures the ontological categories underlying natural language and human common sense. DOLCE, however, does not commit to a particularly abstract level of concepts that relate to the world (like imaginary thoughts); rather, the categories it introduces are thought of as cognitive artifacts, which are ultimately dependent on human perception, cultural imprints and social conventions.

The Cyc project goal is to build a larger common-sense background knowledgebase, which is intended to support unforseen future knowledge representation and reasoning tasks. The Cyc knowledgebase contains 2.2 million assertions (fact and rules) describing more than 250,000 terms, including nearly 15,000 predicates.

Finally, the Cognitive Paradigm Ontology (CogPo) is developed based on two well-known databases, namely, the Functional Imaging Biomedical Informatics Research Network (FBIRN) Human Imaging Data base [23] and the BrainMap database [24]. The CogPo Ontology has categorized each paradigm in terms of (1) the stimulus presented to the subjects, (2) the requested instructions, and

(3) the returned response. All paradigms are essentially comprised of these three orthogonal components, and formalizing an ontology around them is a clear and direct approach to describing paradigms. This well-formed standard ontology guides cognitive experiments in formalizing the cognitive knowledge.

While these ontologies are of great value to the community of researchers, and while the knowledge-based mapping of concepts within particular domains may enable robotic systems to rapidly access the linguistic identity of physical objects and their relations within the domain, they do not provide a means whereby the robot may become spatially aware. To achieve this goal, we will need to provide the robot with the ability to identify the spatial characteristics particular to an identified object, and the physical relations between these objects and the surrounding environment. A robot requires an internal representation of 3D space. It could access two dimensional images on the web, by content-matching those images with contents of its own visual system. This would aid the robot by enabling real-time identification of unfamiliar objects, including spatial parameters that may not be immediately visible to on-board sensors. The matching process, and especially the ongoing 3D interpretation of the images, could be greatly aided if the ontology/metadata associated with images includes representation of the 3D context of image capture. The "ontological" schema of knowledge representation for images may provide this means if it is extended to include 3D spatial annotations.

III. REPRESENTING RELATIONSHIPS IN THREE DIMENSIONS: SPATIAL ONTOLOGIES

We propose here that this same methodology for specifying semantic relationships between concepts (the ontological structure of knowledge representation, i.e., Classes, Properties, and Relationships) may also be useful in specifying spatial relationships between physical objects. While a traditional ontology will hierarchically represent a concept and its relation to other concepts in a domain, a spatial ontology (e.g., Fig. 1) will represent an object, (class), its spatial properties including a detailed 3D representation in a language such as the X3D XML-based file format, and its positional relation (x,y,z) to other objects existing within the scene by using the datatype properties.

An entity (the "*individual*") in a prototypical ontology is comparable to an entity in a spatial ontology, being an object in the physical world. *Class* indicates the category, into which the individual falls, for example "person", or "boat". *Attributes* traditionally describe the *individual* – features, properties, or characteristics of the object: a person has arms; a boat has a hull. In a spatial ontology this information will be appended with configural information regarding the object, for example the parent-child node relationship of a human body, including torso, appendages, etc. The *relation* between individuals is where the power of the traditional ontology arises, by specifying the precise ways, in which different individuals relate to one another (e.g., "a catamaran is a subclass of boat"). Once again, in a spatial ontology the *relation* will be a precise indicator (a reference, or an 'object index') of the relative positionality of items in the physical space, as described in the following section. By thus, leveraging the existing functionality of ontological representation, augmented with relevant and necessary spatial referencing information, we may develop a knowledge-based system that enables a level of spatial awareness in a robotic platform.

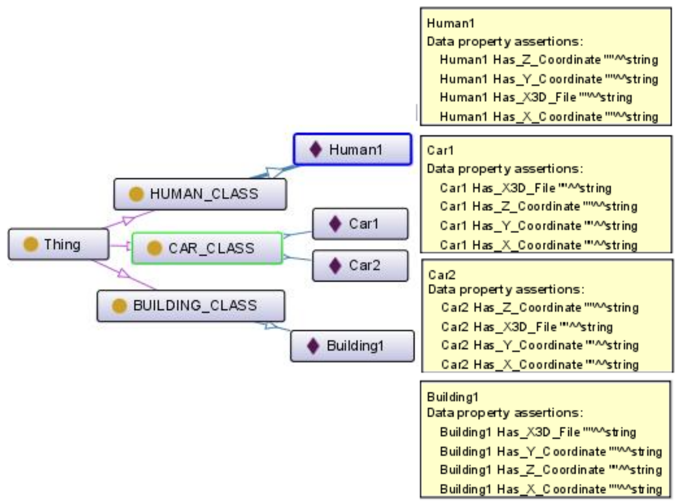

Figure 1. Example of a simple spatial ontology
(Note that the relations between objects are represented via "Data Properties" here.)

A. *A system of comprehension for spatial position*

Following the above discussion about relationships in 3D space, we look into how coordinate systems can be synchronized for objects whose positions and local configurations are non-static. The physical scale requirement that a robot needs to have can be measured by the accuracy the robot needs to operate in via its navigation system. An autonomous robot must be able to determine its position in order to be able to navigate and interact with its environment correctly (e.g., Dixon and Henlich, 1997 [25]). When the *Class* of "robot" navigates from A to B it is a basic motion, which is similar to the movement of an in-game character via a default keyboard set-up where the key "W" moves the character forward, turning left using key "A", turning right using key "D" and go backwards using key "Z". The 3D digital world uses the X, Y, Z coordinate system called the Cartesian Coordinate Method (CCM) and is expressed in meters (m). To measure distance between two spherical points; X^1, Y^1, Z^1 and X^2, Y^2, Z^2 we take the

Euclidean distance using a Cartesian version of Pythagoras' Theorem (1). The distance is the sum of their individual point differences in square.

$$\sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2)} \qquad (1)$$

To determine a position in the physical world and navigate the robot in map-referenced terms to a desired destination point from A to B, Dixon and Henlich use what they call 1) *Global Navigation*. The positioning accuracy with a standard consumer Geographical Positioning System (GPS) is accurate within a range of 8 feet, which is approximately 284 centimeters. This does not give high fidelity position accuracy. As such, when the robot has to operate in a typical indoor manufacturing environment, it needs detailed position support in order to create 3D reference points within the space. What Dixon and Henlich call 2) *Local Navigation* is to determine one's own position relative to the objects (stationary or moving) in the environment, and to interact with them correctly. If we think of Human Robot Interaction (HRI) and the robot arm and its gripper(s) (hand/s), the gripper(s) must via eye(s) be able to recognize the object it will manipulate and how it shall be manipulated. The spatial centre points for individual objects are of importance, as well as group of objects and the robot's own centre point in relation to actual manipulation centre point for gripper. From a spatial ontology point of view, the centre points have to be able to change dynamically depending on interaction purpose.

For example, the Puma robot arm series has three different arms with slightly different sophistication and these are Puma 200, Puma 500, and the Puma 700 Series. These robot arms execute 3) *Personal Navigation* [D&H], which makes the arm aware of the positioning of the various parts, its own positioning, and also in relation to each other and in handling objects. The Puma 200 Series has been used for absolute positioning accuracy for CT guided stereotactic brain surgery [26]. The Puma 200 robot has a relative accuracy of 0.05 mm. There are already 3D Spatial Vision Systems for robots out on the market, which are driven via several cameras. This creates a local world solution for 3D vision robot guidance, where the software first makes the user calibrate the cameras and the robot, and then loads standard Computer Aided Design (CAD) files of parts, which the system shall track.

IV. THE 3D WORLD

The ability to scan a real-world environment makes it possible to extract digital information about the physical world, and the way in which it functions. Three dimensional perception is a key technology for robotics applications where obstacle detection, mapping and localization are core capabilities for operating in unstructured environments. Laser scanning creates a surface point cloud of a 3D physical environment [34] making it possible to map any environment in a rather short time (the Leaning Tower of Pisa was scanned in 20 minutes). This technology can be used in a robotic intelligence system for Simultaneous Localization Mapping (SLAM) and higher level reasoning regarding location and position. However, object recognition and manipulation requires deriving 3D object information from the overall point cloud and building cognitive models with task reasoning for using object and scene data in real time.

Object extraction [35][36][37] makes it possible to know what a robot is looking at, supporting manipulation or collection actions. This can be achieved by an Environmental Scanning-Object Extraction (ES-OE) engine. For human-robot collaboration, a robot can be enabled to use deictic visual references from human gaze by integrating an eye tracker with the ES-OE engine.

A. Background

In a previous work [38], a 3D simulation engine was integrated with an eye tracker. The integrated system allows the human point of gaze on 3D objects within a 3D digital world projected onto a computer screen to be tracked automatically. This development made it possible to log gaze in various task-related environments in a simulated world. From a Human Factor's perspective, the simulation and human observation can be investigated, including collaborative actions performed by groups with various workloads, stressors and decisions. There have been several studies made using the technological framework with different stimuli [39][40][41], but no substantial theoretical framework has been developed in relation to this object-based approach *per se*. A bottleneck in relation to this visual approach has been that 2D image, film and visual stimuli have not met the requirements for incorporating a knowledge-based approach for dynamic 3D worlds, whether the real physical world or a digitized 3D world. The object approach needs to address how both modeled and real world objects can be perceived and manipulated [42] by a robot, allowing the system to sense, think and act in real time: the computer needs to understand how to define an object and how to ontologically and semantically make sense out of such an object in a dynamic spatial world.

1) 3D objects in a 3D world

In [38], a simulation engine integrated with an eye tracker took a gaze fixation (x and y screen coordinates) and ray casted/traced from that position onto the underlying 3D virtual object's collision box, a volume corresponding with the shape of a virtual object as recognized and processed by a physics engine that is also used to designate objects by interface devices, like a mouse. This made it possible to track gazed objects in real time every 17 ms (using a 60Hz eye tracker). The same principle can be used in a physical world context where an ES-OE engine could be integrated with eye tracking glasses to allow a computational system to know what object a person wearing the glasses is looking at.

2) Structuring a noisy world

The 3D world scenario, simulated or physically real, constitutes an event or scene. A scenario includes objects that are instances of their classes. A class could be something like a *CarClass*, *HumanClass*, *FlowerClass*, etc.

In a constrained world, we can name all objects beforehand so when they are logged we know what they are and what position (x, y, z, θ_1, θ_2, θ_3) they are in. In an unconstrained environment that is scanned and has extracted objects, we must also have a capability to know what the objects are and to be able to classify them. A cloud-based approach of the kind proposed in this paper presents a middle ground, being more open than a highly constrained environment, but still being limited to objects of types that are represented and labeled within the cloud.

V. INTELLIGENT ACTION IN A STRUCTURED WORLD

Knowledge by definition is "1. Facts, information, and skills acquired through experience or education; the theoretical or practical understanding of a subject and 2. Awareness or familiarity gained by experience of a fact or situation" [43]. To gain an understanding of how robots might learn and operate on knowledge, we have looked at several established models that can fit within an initial architecture that enhances these established models by the ingestion of information from the web. Our overall aim is to build a computational comprehension system for 3D object information, assisted by a hybrid computational ontology (i.e., combining several existing and new ontologies).

A. Existing Models

Extensive effort has been put into the task of understanding and attempting to re-create/simulate the processes, by which a human being thinks. Using the underlying assumption that intelligence is wholly "the simple accrual and tuning of many small units of knowledge" [44], production-based models of cognition have had success in displaying human-like performance on a number of tasks (e.g., visual search [45] and natural language processing [46]). While there are debates regarding the similarity of what humans actually do to what we have achieved using the above assumption [47], there is little doubt that such systems can produce intelligent-seeming behavior, which can facilitate the development of vitally useful control structures in the field of robotics and computational intelligence [46].

One of the most influential models of human cognition is the ACT-R, or "Adaptive Character of Thought – Rational" model [44], developed over many years by John Anderson, who was a student of the seminal Cognitive Scientist Alan Newell (1927-1992). Anderson's model is a hybrid symbolic/sub-symbolic system that incorporates various "modules" that are deemed necessary for rational behavior, and are thought to have biological correlates. These include the modules *Declarative* (manages creation, storage and activation of memory "chunks"), *Procedural* (stores and executes productions based on expected utility), *Intentional/Imaginal* (goal formulation for directed behavior), and *Visual (2D)/Audio* (theoretically plausible implementation of visual and auditory perception), see Fig. 2. An internal pattern-matching function searches for a production that matches the current state of the buffers.

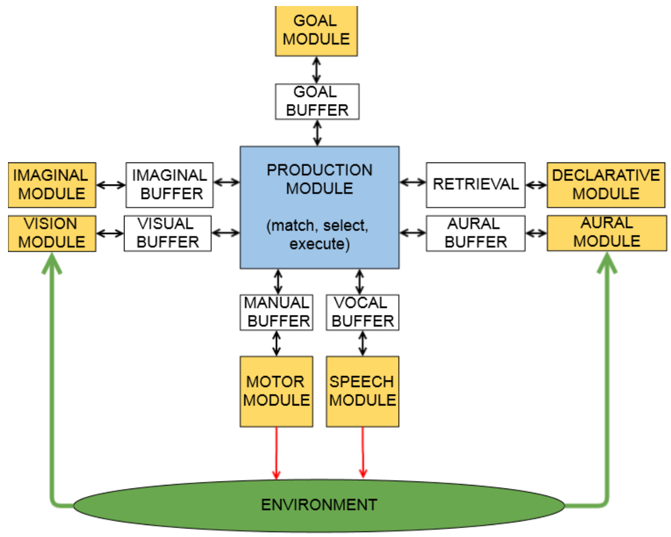

Figure 2. A schematic representation of the canonical ACT-R cognitive model.

ACT-R is formed as a knowledge model where the "chunks" are the elements of declarative knowledge in the ACT-R theory and are used to communicate information between modules through the buffers. A chunk is defined by its chunk type, that is described by its slots (here compared with properties), see Table I. Chunk types can be organized as a hierarchy of parent (SuperType)-child (SubType) relationships. The subtype will inherit all of the slots (properties) of the parent node(s).

Other models that take a similar symbolic approach to model human cognition include Soar [48], EPIC (Executive-Process/Interactive Control) [49], CLARION (Connectionist Learning with Adaptive Rule Induction On-line) [50], and others (for a detailed review see [51]). While these have been successful to varying degrees at modeling specific human cognitive task(s) performance, it is becoming evident that such models are intrinsically limited by their disconnection from the real world, in which humans (or robots) operate. A production based system is only as adaptive as its rule set allows given the inputs provided to it, which have generally been limited to "screen as eye" and "keyboard/mouse as hands" mappings. A new wave of thought surrounding the development of cognitive models is embracing the need for "embodied" cognition, improving the ability of the system to sense and act. One example of this is the ACT-R/E ("E" for "Embodied") framework, used as an operating system for mobile robotics developed by the American Naval Research Lab [52], depicted in Fig. 3.

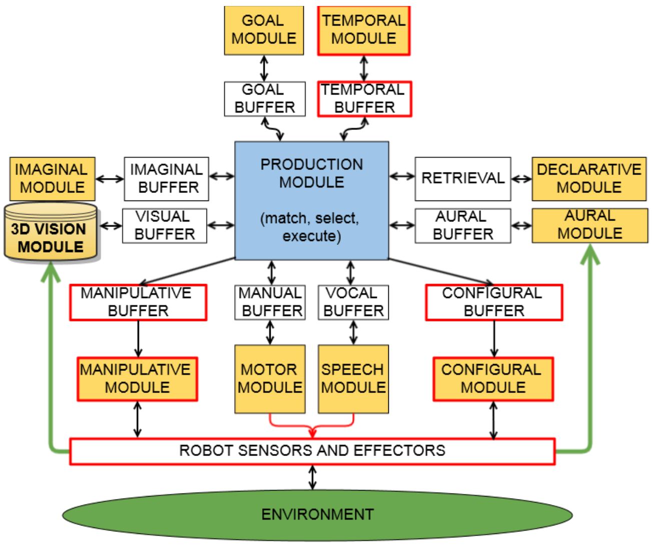

Figure 3. The "embodied" (Visual 3D) modifications introduced by Trafton et al. 2012. Additions in the ACT-R/E are highlighted in red.

The Object-Attribute-Relation (OAR) model of Wang, 2007 [53], specifies the elements of a cognitive model in the fashion of an ontology, the logical model of memory. In an attempt to formally describe the mechanism of human Long Term Memory (LTM), which he states is the "foundation of all forms of natural intelligence" (p. 66), Wang decomposes the construct into three elemental components – Objects, Attributes and Relations (OAR). This OAR model allows the computational specification of the human LTM formation and storage process, and is put forth as having sufficient explanatory power as to describe the "mental process and cognitive mechanisms of learning and knowledge representation" (p.72). This model has a strong parallel with the specification of knowledge in information processing ontologies. This parallel is direct, as described by the relations given in Table I.

TABLE I. COMPARISON OF MODEL TYPE CONSTRUCTS

OAR Model	Ontology Components	ACT-R ACT-R/E
Object(s) Attribute(s) Relation(s)	Class(es) Propert(ies) Relationship(s)	Chunk Type(s) Chunk Slot(s) Function(s)

A critical issue for any of these kinds of models is the relationship of their constructs to the environments, in which they are expected to provide foundations for action. The core notion of *embodiment* is to provide the heretofore functionally "disembodied" computational model with sensors and effectors that allow its direct interaction with the physical world. In such a way, the inherent limitation of human-defined input may be overcome. In addition to physical sensory perception and manipulative ability, a human may have access to a detailed semantic understanding of the surrounding world. In the quest to produce a non-human intelligent actor within a physical space, we must provide the actor with an understanding of underlying structures, i.e., specific denotations in the physical world.

VI. PROPOSED MODEL

In the CogOnto model, we propose a further augmentation of the cognitive models discussed above, providing the robot with detailed 3D schematic representations of objects that it encounters in real time, supported via task models, knowledge models and ontologies.

The CogOnto model is composed of five parts $\triangleq$ $<S_i,C_i,A_i,O_i,R_i>$, where $i = 1..\,N$, and where S_i is a finite set of situations, C_i is a finite set of classes, A_i is a finite set of attributes for characterizing a class, O_i is a finite set of objects in a class, and R_i is a finite set of relationships among the objects. In the CogOnto model (Fig. 4), we consider the following features [54][56]:

- Situation: represents an interactive (i.e., dynamic) real world scenario.
- ConceptNet: is a network of class-to-class relationships applicable in a given situation.
- ObjectNet: an object is an instance of a class. ObjectNet is a network of object-to-object relationships.
- AttributeNet: is a network between properties of classes and objects.
- Relation: is a function associating concepts, classes, objects and attributes; e.g., a *robot* is *part-of* an *Intelligent Agent (IA)*, were the "part-of" relation connects two concepts. The relations (associations) may be modeled or created by an autonomous learning process.

These constructs are not defined in detail here, but unlike the other models are not limited to textual/linguistic meanings. The CogOnto model illustrated in Fig. 4 has four major functional elements that share information: 1) the ES-OE engine, 2) the eye tracking system interconnected with the ES-OE engine, 3) the OAR model functioning as the basis of the Cognitive System, and 4) the knowledge cloud, including external resources such as WordNet or Cyc. The latter is also called the Linked Open Data and may be used to illustrate the intelligent process for sharing and exposing information in machine readable form by using uniform resource identifiers based on Berners-Lee's [55][56] principles. These principles enable data communication guiding perception from procedural memory.

The knowledge system of the CogOnto model can be perceived as a storage system that accesses real world object information and external semantic resource information via the existing knowledge cloud [57].

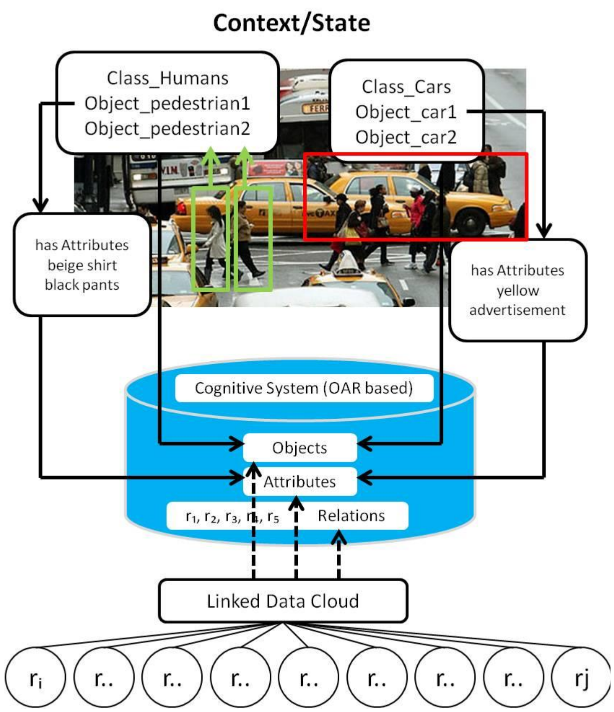

Figure 4. The CogOnto model and its operative states. The relations are build up for the current scene via object gaze tracking, and past stored scenes using a match function.

The knowledge system represents the integration of formal symbolic and free text descriptors of an object.

VII. INTEGRATING SEMANTIC WEB CONCEPTS, TECHNOLOGIES AND RESOURCES

CogOnto integrates its own knowledge resources with external resources accessible via the web. For example, WordNet is a lexical database where nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets). To recall an object, the 'synsets (WordNet 2.1)' [58] and the W3C [59] standard can be used at a text level, to describe what an object is when it is text-labeled. Ontologies can be expressed by using Semantic Web tools, e.g., Web Ontology Language (OWL) [60] and the Resource description framework Schema (RDFS) [61].

The OAR model, with its Object, Attribute and Relation parts, and the ontological framework, containing Class/Instance, Relationship and Properties, can be inter-mapped so the object world can be comprehended using existing resources and using the 3D information represented internally within an object model. The 3D object's internal structure and shape can either be structured as Free Form Geometry (FFG) with surfaces and curves, or as Polygonal Geometry (PG) with points, lines and faces. The objects can be extracted and exported into different file formats, such as, e.g., .obj files, .stl files. The .stl file format is a triangular representation of a 3D object, where each triangle is uniquely defined by its normal and three points representing its vertices. The format is native to the stereolithography Computer Aided Design (CAD) software created by 3D Systems (in this kind of format it is also possible to print the object out from a 3D printing machine).

The 3D object file contains different layers cognitively (form, volume, size, other descriptive attributes, etc.), supporting our senses and perception operating in parallel when performing allocated manipulation tasks. A human looking at an object can relate to the object both on a denotative- and on a connotative level. The denotative level is understood as a pure noun level without any cultural associations, nor any emotional or associative signifiers to the object, it is purely instrumental. The connotative layer is, on the other hand, the level of cultural and personal associations attached to an object with experience over time. Geometrical information within the 3D object can be represented using the X3D XML-based file format, an ISO standard for representing 3D computer graphics.

VIII. BEYOND ONTOLOGIES – COMPLEX RELATIONSHIPS, AND ALTERNATIVES TO HEIRARCHICAL DATA REPRESENTATION

As we move from relatively canonical data sets, for which the information processing ontology was designed (i.e., semantic relations within a particular knowledge base) to more complex relationships (such as ad-hoc physical relations), in which the hierarchical order is not nearly so explicit, or potentially non-existent, will the classical ontology suffice? Or alternately, will something more adaptive need to take its place? Because relationships in the physical world are multifaceted and multidirectional, it is useful to have a schema that can represent this interconnectedness. The key strength of an ontology is that it provides a concrete nomological environment, from which to operate within the chosen domain. Table II summarizes the traditional information processing ontology.

TABLE II. TRADITIONAL ONTOLOGY CHARACTERISTICS

- allows a common understanding of the structure of information - enables reuse of domain knowledge - makes domain assumptions explicit - separates domain knowledge from operational knowledge - defines a common vocabulary for researchers - provides machine readable definitions of basic concepts and the relationships among them

However, there are instances (albeit few as of this writing), in which it is being recognized that the intrinsic limitations of the "ontology" such it is commonly understood in 2014, (e.g., OWL-based) are sufficient as to demand a modification whereby the innate complexities of real-world phenomenon may be modeled. That is: complex, potentially non-hierarchical relationships.

For example, it has been noted in the field of chemical molecular informatics that while ontologies are able to represent tree-like structures, they are unable to represent cyclical or polycyclical structures [27]. Similarly, the difficulty in building classifications of nano-particles has led some researchers to begin to look into taxonomies based on "physical / chemical / clinical / toxic / spatial" characteristics of an object, supplemented by structural information, in order to account for shapes, forms and volumes [28]. Other examples of representing complex structural relations that stretch the boundaries of ontological representation include using Description Graph Logic Programs (DGLP) to represent objects with arbitrarily connected parts [29], and a hybrid formalism whereby the authors propose a "combination of monadic second order logic and ordinary OWL", where the two representations are bridged using a "heterogeneous logical connection framework" [30].

It is evident that the potential applications of a formalism such as the ontological method of information representation far outreach the initial conceptualizations of the language. While it may be possible to model 3D spatial information within the constraints of a hierarchical ontology, it is also to be considered that this notion, as well as applications such as those described above, may require the development of progressive, flexible alternatives, which capture the strengths of the ontology (i.e., the points from Table II), while managing to represent arbitrary or non-hierarchical relationships.

A. *Cognitive Models and Ontologies*

One information system where a non-hierarchical organization may be necessary, when attempting to map the internal structural relations, is the human brain. For more than half a century, researchers across many fields (e.g., Cognitive Psychology, Neuroscience, Cognitive Science) have been using models to posit and test hypothetical interpretations of how the human brain is structured. These range from the very simple (e.g., Baddely's working memory model, [31]) to complex neurological models (e.g., [32]), though no current model has even begun to approach the actual complexity of the human brain. On a neuronal level, and certainly even on a functional level such as between brain regions, this is a non-hierarchical system. It is once again remarkable that, at a superficial level, the development of ontologies draws a strong parallel with theoretical interpretations of how the human cognitive system might be structured (refer back to Table I). This relation is further discussed in Sennersten et al. [13].

In OAR (Object, Attribute, Relation), Wong [10] develops a model that most certainly shares conceptual roots with ontological knowledge representation. Likewise, parallels may be drawn with Anderson's ACT-R model [11] and Trafton's "embodied" version [32] ACT-R/E. In each model, *Objects* in the real world possess characteristics (i.e., *attributes*, or *properties*) and also *relations* with one another. If we can augment these heretofore largely semantic components with a functional representation of 3D space (e.g., at the 3 levels *Global*, *Local*, and *Personal*), we may have the fundaments of a system of Spatial Understanding for a robotic platform.

IX. CONCLUSION AND FUTURE WORK

The CogOnto model with support from the technological implementation of the eye tracker system with the ES-OE engine can represent cognitive relations that can be processed by a robot operating in a spatial world [62].

Formal knowledge structures within CogOnto face similar challenges to other knowledge representation formalisms, and this paper has shown isomorphism with a number of examples. However, the primary advance proposed is to use cloud-based resources that are not limited to formal representations to enhance the robustness of knowledge processing by the integration of similarity-based search. Those cloud-based resources may use text and images. But more interesting extensions for future work include new forms of cloud content, such as multi-spectral images, point clouds and behavior tracks. The main ongoing research challenge is to provide suitable similarity metrics for these data forms, integrating search results with formal structures, and developing methods for integrating them in unified search, or meta-search, results.

One of the few certainties regarding the immediate future is that robotic control technology will advance from systems that are coded for specific applications, to systems that are designed with an innate adaptability to unexpected environmental situations. This will require new methods of providing on-the-fly relational information to the robot, in order for it to gain an understanding of both its spatial position, and the position of other objects in the vicinity, their characteristics, and the ways that it can relate to them. A reworking of the traditional OWL-based ontology, with an eye for 3-dimensional spatial relations on 1) Global, 2) Local, and 3) Personal levels of specificity may be sufficient to this end.

It is also noted that as data sets become more complex, and especially as we begin to consider that most complex of biological control systems, the human cognitive system, it may very well become necessary to develop hybrid ontological-type systems of knowledge representation, which 1) encompass the full realm of advantages provided by the use of specific nomologial hierarchies, and 2) enable the encoding of arbitrary or non-hierarchical relationships.

The development knowledge-based systems that can account for abstract, non-hierarchical relations could potentially facilitate the next generation of spatially aware robotics applications.

REFERENCES

[1] M. Lochner, C. Sennersten, A. Morshed, and C. Lindley, "Modelling Spatial Understanding: Using knowledge representation to enable spatial awareness in a robotics platform," The 6th International Conference on Advanced Cognitive Technologies and Applications, 25-29th of May 2014, Venice, Italy.

[2] J. Kuffner, "Cloud Enabled Robots," Presentation, IEEE Humanoids conference, Nashville, Tenn. 2010. http://www.scribd.com/doc/47486324/Cloud-Enabled-Robots [retrieved: 2014.11.26]

[3] R. Khosla, M. Chu, R. Kachouie, K. Yamada, and T. Yamaguchi. "Embodying Care in Matilda: An Affective Communication Robot for the Elderly in Australia," In Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, pp. 295–304, 2012. http://dl.acm.org/citation.cfm?id=2110398. [retrieved: 2014.11.26]

[4] J. M. Lutin, A. L Kornhauser, and E. Lerner-Lam. "The Revolutionary Development of Self-Driving Vehicles and Implications for the Transportation Engineering Profession," Institute of Transportation Engineers, ITE Journal, vol. 83(7), July 2013, pp. 28.

[5] L. M. Hiatt, S. S. Khemlani, and J. G. Trafton, "An Explanatory Reasoning Framework for Embodied Agents," Biologically Inspired Cognitive Architectures, vol. 1, July 2012, pp. 23–31, doi:10.1016/j.bica.2012.03.001.

[6] A. Elfes, J. L. Hall, E. A. Kulczycki, D. S. Clouse, A. C. Morfopoulos, J. F. Montgomery, J. M. Cameron, A. Ansar, and R. J. Machuzak, "An Autonomy Architecture for Aerobot Exploration of the Saturnian Moon Titan," IEEE Aerospace and Electronic Systems Magazine, vol. 23(7), July 2008, pp. 1-9.

[7] K. K. Tahboub and S. N. Al-Din Munaf, "A Neuro-Fuzzy Reasoning System for Mobile Robot Navigation," JJMIE vol. 3(1), March 2009, pp. 77-88. http://pdf.aminer.org/000/361/105/a_neuro_fuzzy_approach_to_autonomous_navigation_for_mobile_robots.pdf. [retrieved: 2014.11.26]

[8] J. Fasola and M. Matarić, "Using Spatial Language to Guide and Instruct Robots in Household Environments," Refereed Workshop, AAAI Fall Symposium: Robots Learning Interactively from Human Teachers, Arlington, VA, Nov 2012. http://www.aaai.org/ocs/index.php/FSS/FSS12/paper/viewFile/5582/5880. [retrieved: 2014.11.26]

[9] C. Hudelot, J. Atif, and I. Bloch, "Fuzzy Spatial Relation Ontology for Image Interpretation," Fuzzy Sets and Systems, vol. 159(15), August 2008, pp. 1929–1951. doi:10.1016/j.fss.2008.02.011.

[10] G. Fu, C. B. Jones, and A. I. Abdelmoty, "Ontology-Based Spatial Query Expansion in Information Retrieval," On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE, Springer, 2005, pp. 1466–1482. http://link.springer.com/chapter/10.1007/11575801_33. [retrieved: 2014.11.26]

[11] Y. Wang, "The OAR model for knowledge representation," Proc. The 2006 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE'06), Ottawa, Canada, May 2006, pp. 1692-1699.

[12] J. R. Anderson, "ACT," American Psychological Association, vol. 51(4), 1995, pp. 355-365.

[13] M. Bosse, R. Zlot, and P. Flick, "Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping," IEEE Transactions on Robotics, vol. 28(5), October 2012, pp. 1104–1119. doi:10.1109/TRO.2012.2200990.

[14] C. Sennersten, A. Morshed, M. Lochner, and C. Lindley, "Towards a cloud-based architecture for 3D object comprehension in cognitive robotics," The 6th International Conference on Advanced Cognitive Technologies and Applications, 25-29th of May 2014, Venice, Italy.

[15] C. A. Lindley. "Synthetic Intelligence: Beyond A.I. and Robotics," in Integral Biomathics: Tracing the Road to Reality, Simeonov, Plamen L.; Smith, Leslie S.; Ehresmann, Andrée C. (Eds.), Springer, 2012.

[16] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications," Knowledge Acquisition, vol. 5(2), 1993, pp. 199–220.

[17] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, and Schneider, L, "Dolce: a descriptive ontology for linguistic and cognitive engineering," WonderWeb Project, Deliverable D17, vol. 2(1), 2003.

[18] C. Masolo, et al. "The WonderWeb Library of Foundational Ontologies," IST Project 2001 - 33052 WonderWeb: Intermediate Report, May 2003.

[19] J. A. Turner and A. R. Laird, "The cognitive paradigm ontology: design and application," Neuroinformatics, vol. 10(1), 2012, pp. 57-66.

[20] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to wordnet: An on-line lexical database," International journal of lexicography, vol. 3(4), 1990, pp. 235-244.

[21] C. Matuszek, J. Cabral, M. J. Witbrock, and J. DeOliveira, "An Introduction to the Syntax and Content of Cyc," AAAI Spring Symposium: Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering, March 2006, pp. 44-49.

[22] Laboratory for Applied Ontology (LOA) (ISTC-CNR) http://www.loa.istc.cnr.it/ [retrieved: 2014.11.26].

[23] D. B. Keator, et al. "A national human neuroimaging collaboratory enabled by the Biomedical Informatics Research Network (BIRN)," Information Technology in Biomedicine, IEEE Transactions on, vol. 12(2), 2008, pp. 162-172.

[24] A. R. Laird, J. J. Lancaster, and P. T. Fox, “Brainmap,” Neuroinformatics, vol. 3(1), 2005, pp. 65-77.

[25] J. Dixon and O. Henlich, “Mobile Robot Navigation,” Final Report, Information Systems Engineering, Imperial College, UK, 1997.

[26] Y. Kwoh, J. Huo, E. Jonckheere, and S. Hayati, “A Robot with Improved Absolute Positioning Accuracy for CT Guided Stereotactic Brain Surgery,” IEEE Transactions on Biomedical Engineering, Feb 1988, vol. 35(2), pp. 153-160.

[27] J. Hastings, C. Batchelor, and M. Okada, “Shape Perception in Chemistry,” Proceedings of the Second Interdisciplinary Workshop The Shape of Things (SHAPES 2013), Rio de Janeiro, Brazil, April 3-4, 2013, pp. 83-94. http://ceur-ws.org/Vol-1007/paper6.pdf. [retrieved: 2014.11.26].

[28] V. Maojo, et al. “Nanoinformatics: Developing New Computing Applications for Nanomedicine,” Computing, vol. 94(6), March 7, 2012, pp. 521–539, doi:10.1007/s00607-012-0191-2.

[29] D. Magka, B. Motik, and I. Horrocks, “Modelling structured domains using description graphs and logic programming,” Lecture Notes in Computer Science, vol. 7295, 2012, pp. 330-344, Department of Computer Science, University of Oxford, 2011.

[30] O. Kutz, J. Hastings, and T. Mossakowski. “Modelling Highly Symmetrical Molecules: Linking Ontologies and Graphs Artificial Intelligence: Methodology, Systems, and Applications,” Lecture Notes in Computer Science, vol. 7557, chap. 11, pp. 103–111, Springer Berlin / Heidelberg, Berlin, Heidelberg, 2012.

[31] A. D. Baddeley and G. Hitch, “Working memory,” In The psychology of learning and motivation: Advances in research and theory, vol. 8, G.H. Bower, Ed. New York: Academic Press, 1974, pp. 47–89.

[32] M. Riesenhuber and T. Poggio, “Hierarchical models of object recognition in cortex,” Nature Neuroscience, vol. 2(11), November 1999, pp. 1019–1025.

[33] G. Trafton, L. Hiatt, A. Harrison, F. Tanborello, S. Khemlani, and A. Schultz, “ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction,” Journal of Human-Robot Interaction, vol. 2(1), March 2013, pp. 30–55. doi:10.5898/JHRI.2.1.Trafton.

[34] M. Bosse, R. Zlot, and P. Flick, “Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping,” IEEE Transactions on Robotics, vol. 28, no. 5, 2012.

[35] R. Zeibak and S. Filin, “Object extraction from Terrestrial Laser Scanning Data,” TS 8E –Terrestrial Laser Scanning, Visualization and LIDAR, FIG Working Week 2009, Surveyors Key Role in Accelerated Development, Eilat, Israel.

[36] S. Westerberg and A. Shiriaev, “Virtual Environment-Based Teleoperation of Forestry Machines: Designing Future Interaction Methods,” Journal of Human-Robot Interaction, vol. 2, no. 3, 2013.

[37] A. El Daher and S. Park, “Object Recognition and Classification from 3D Point Cloud,” student project, 2006, http://cs229.stanford.edu/proj2006/ElDaherPark-ObjectRecognitionAndClassificationFrom3DPointClouds.pdf [retrieved: March 2013], Stanford Education at Stanford University, USA.

[38] C. Sennersten and C. Lindley, “Evaluation of Real-time Eye Gaze Logging by a 3D Game Engine,” 12th IMEKO TC1 & TC7 Joint Symposium on Man Science and Measurement, Annecy, France, 2008.

[39] C. Sennersten, M. Castor, R. Gustavsson, and C.A. Lindley, “Decision Processes in Simulation-Based Training for ISAF Vehicle Patrols,” NATO-OTAN, MP-HFM-202-17, 2010.

[40] P. Jerčić, et al., “A Serious Game Using Physiological Interfaces for Emotion Regulation Training In The Context of Financial Decision Making,” ECIS 2012 Proceedings. AIS Electronic Library (AISeL), 2012.

[41] H. Cederholm, O. Hillborn, C. Lindley, C. Sennersten, and J. Eriksson, “The Aiming Game: Using a Game with Biofeedback for Training in Emotion Regulation,” 5th Digital Games Research Association (DIGRA) Conference THINK DESIGN PLAY 2011, Utrecht, Netherlands.

[42] J. R. Flanagan, G. Rotman, A. F. Reichelt, and R. S. Johansson, “The role of observers’ gaze behavior when watching object manipulation tasks: predicting and evaluating the consequences of action,” Philosophical Transactions of the Royal Society –B: Biological Sciences, 2013, UK.

[43] Oxford English Dictionary entry: http://oxforddictionaries.com/definition/english/knowledge, [retrieved: 2014.11.26].

[44] J. R. Anderson, “ACT,” American Psychological Association, Vol. 51, No.4, 1995, p. 355-365.

[45] D. Kieras and D. Meyer, “An overview of the EPIC Architecture for Cognition and Performance with Application to Human-Computer Interaction,” University of Michigan, EPIC report No. 5 (TR-95/ONR-EPIC-5), 1995, DTIC Document 1995, 43 pages.

[46] D. P. Benjamin, D. Lonsdale, D. Lyons, and S. Patel, “Using Cognitive Semantics to Integrate Perception and Motion in a Behavior-Based Robot,” In Learning and Adaptive Behaviors for Robotic Systems, 2008, LAB-RS’08. ECSIS Symposium on, pp. 77–82. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4599431 [retrieved: 2014.11.26].

[47] E. Hutchins, Cognition in the Wild. Chapter 9 – Cultural Cognition, 1996, pp. 353-374.

[48] J. Laird, A. Newell, and P. Rosenbloom, “SOAR: An Architecture for General Intelligence,” Technical report AIP-9, University of Michigan, Carnegie-Mellon University, Stanford University, Artificial Intelligence 33, 1987, pp. 1–63, DTIC Document 1988, 63 pages.

[49] D. Kieras and D. Meyer, “The EPIC architecture for modeling human information-processing and performance: A brief introduction,” EPIC Report No.1 (TR-94/ONR-EPIC-1), University of Michigan, 1994, DTIC Document 1994, 43 pages.

[50] R. Sun, “The CLARION Cognitive Architecture: Extending Cognitive Modeling to Social Simulation,” Cognition and

Multi-Agent Interaction, Oct. 2004, Cambridge University Press, New York, 2006.

[51] D. Vernon, G. Metta, and G. Sandini, “A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents,” IEEE Transactions on Evolutionary Computation 11 (2), April 2007, pp. 151–180, doi:10.1109/TEVC.2006.890274.

[52] G. Trafton, et al., “ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction,” Journal of Human-Robot Interaction, vol.2, No.1, 2013, pp. 30-55.

[53] Y. Wang, “The OAR model for knowledge representation,” Proc. The 2006 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE’06), Ottawa, Canada, pp.1692-1699.

[54] A. Oltramari and C. Lebiere, “Extending Cognitive Architechtures with Semantic Resources,” Department of Psychology, Artificial General Intelligence Lecture Notes in Computer Science, Vol. 6830, 2011, pp. 222-231.

[55] T. Berners-Lee, “Linked data-the story so far,” International Journal on Semantic Web and Information Systems, 5(3), 2009, pp. 1-22.

[56] Linked_Data_Design_Issue. http://www.w3.org/DesignIssues/LinkedData.html [retrieved: 2014.11.26].

[57] C. D’este et al., “Sustainability, Scalability, and Sensor Activity with Cloud Robotics,” Proceedings of Australasian Conference on Robotics and Automation, 2-3 Dec 2013, University of New South Wales, Sydney, Australia.

[58] WordNet: http://wordnetweb.princeton.edu/perl/webwn [retrieved: 2014.11.26].

[59] World Wide Web Consortium is the main international standards organisation for the World Wide Web.

[60] OWL W3C: http://www.w3.org/TR/owl-features/ [retrieved: 2014.11.26].

[61] RDFS W3C: http://www.w3.org/TR/rdf-schema/ [retrieved: 2014.11.26].