

# Knowledge Graphs for Analyzing Thermal Comfort: A Proposal

Jorge-Luis Jácome-Domínguez<sup>1</sup>, Edgard Benítez-Guerrero<sup>1</sup>,  
Guillermo Molero-Castillo<sup>2</sup>

<sup>1</sup>Universidad Veracruzana, Xalapa, Veracruz, Mexico  
jorge.jacomdiguez@gmail.com, edbenitez@uv.mx

<sup>2</sup>Universidad Nacional Autónoma de México, Mexico  
gmolero@fi-b.unam.mx

**Abstract.** Knowledge Graphs (KGs) can be used to provide a unified, homogeneous view of heterogeneous data, which then can be queried and analyzed. In this paper, we explore the use of KGs to analyze the Thermal Comfort (TC) of users in specific environments (e.g. classrooms, hotel rooms). This implies the integration of several data sources that provide environmental variables (i.e. temperature, humidity), but also users' physiological variables (i.e. temperature). Thus, this paper proposes a method to integrate and analyse environmental and physiological data through KGs in the context of TC.

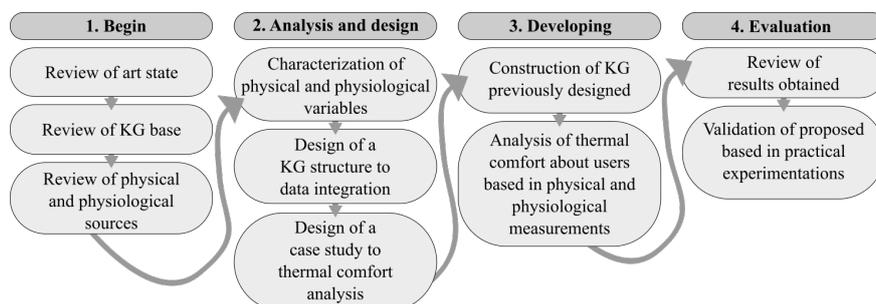
**Keywords:** Knowledge Graphs, Thermal Comfort, Data Integration.

## 1 Introduction

A Knowledge Graph (KG) is a structure designed to represent and organize knowledge [1–3], such as the structure of an information system [1] or specific topics for intelligent services, among others [4, 5]. In recent years, KGs have attracted the interest of diverse research groups and companies (Google, Microsoft, University of Leipzig, University of Mannheim, to name a few) [2, 6], which have implemented KG-based solutions in diverse fields, such as the Internet of Things, big data, semantic web, information management, among others [4, 5, 7, 8].

KGs can be used to integrate several, potentially heterogeneous data sources [5], with the objective of organizing the content of the sources and present them as a single source of information [3, 7]. This is particularly needed in Context-Aware systems, where the system uses context information to provide data and/or services to its users. That information usually needs to be acquired from heterogeneous data sources and integrated to enable its use in decision-making processes. An application domain of this kind of systems, that is of special interest to our research project, is the analysis of Thermal Comfort (TC). It refers to the level of satisfaction of a person with respect to the thermal conditions of the space (building, room) where he/she is [9, 10]. It has been recognized that several physical (temperature, humidity, air quality) and physiological factors (heart rate, stress) influence the thermal comfort of an individual. The physical factors

refer to environmental features, while the physiological are people characteristics [9, 11, 12]. The factors described above can be found in the ANSI/ASHRAE 55-2017 standard [9, 13] and some recommendations of the World Health Organization (WHO) [10], and have been traditionally used for the design of spaces in static contexts. However, environmental and individual conditions change over time, and thus it is necessary to consider this dynamism in a computer system in order to provide a more exact and personalized solution to guarantee the comfort of users. Our work is aimed to propose such a system.



**Fig. 1.** Research Method.

Let us note that to achieve the aims of our work, we are following a research method consisting of four stages as shown in Figure 1. The first stage is aimed to analyze the literature related with the topics of Knowledge Graph, integration of heterogeneous data sources, and Thermal Comfort. The second stage corresponds to the characterization of physical and physiological variables needed to determine Thermal Comfort and, based on that characterization, design a KG-based solution to connect sources exposing that data (e.g. sensors). The third stage refers to the development and testing of the KG-based solution. For the development it is foreseen the use of tools for creating and querying KG. Finally, the fourth stage encompasses the evaluation of the results obtained. This will be done from the point of view of the capacity of the solution to connect heterogeneous data sources and to determine the Thermal Comfort of the users.

This paper particularly proposes a framework to integrate and analyse environmental and physiological data through KGs in the context of the study of TC. This implies a) characterization of physical and physiological variables, b) Knowledge Graph design, c) construction, and d) evaluation of solution proposal.

## 2 Background

### 2.1 Knowledge Graphs

A Knowledge Graph (KG) is a structure designed to store, organize and represent knowledge [1–3]. A KG is formed by an indefinite number of entities that are

linked to other entities through relationships; the entities and relations of a KG represent facts [14]. A KG is usually represented as a directed acyclic graph, where each node contains an entity and an arc between two nodes represent a relationship [2, 14]. Associated to a KG it is possible to find tools to query, interpret and present its content [1, 14].

Nowadays, there are no formal classification of KGs. However, it has been possible to observe three different types [1]:

1. **Prior knowledge.** A KG of this kind is constructed to work with a set of existing data and its objective is to solve some problems and inconsistencies in a well-defined context [1, 2].
2. **Existing model.** KGs of this kind model information through a highly expressive and complex semantic scheme [2, 3], that is usually an ontology [1]
3. **Defined structure.** A KG of this kind is based on a structure of data or objects, as the Resource Description Frameworks (RDF) [15], and its objective is the study of disordered data and the discovering of implicit hierarchies between its entities [1, 5].

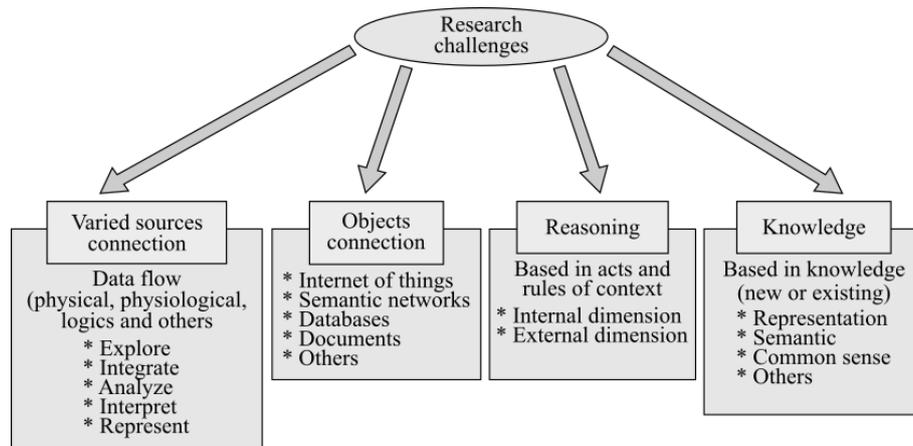


Fig. 2. KGs research challenges.

The popularity of KGs has grown in recent years, as shown by the increasing number of research groups and companies that work on the topic [2, 6]. KGs are usually used in works that focus on the representation and use of knowledge in the Semantic Web and Natural Language Processing [5, 6, 16]. However, there are still a number of research challenges related to KGs, as shown in Figure 2.

The role of KGs regarding the representation, querying and analysis of heterogeneous data, especially in dynamic contexts [5, 16], needs to be more explored.

Here, a number of problems that arises while acquiring data from highly sensitive environments need to be addressed [5, 8, 11]: errors during the measurement of context parameters, data loss, speed, size and meaning [5, 11, 13].

Other challenges associated with KGs are related to the connection of objects (e.g. IoT-devices), and the representation of knowledge (e.g. common sense [8]) and associated reasoning mechanisms. This project, as previously mentioned, focuses on the integration of heterogeneous data sources, in specific, physical and physiological data sources.

## **2.2 Heterogeneous Data Sources**

Traditionally, data that is needed to make an informed decision does not come from a single source, but from several potentially heterogeneous data sources. This heterogeneity relates to differences among the sources with respect to formats, amount of data issued, acquisition speed, meaning, among others [5, 9, 11, 13]. In the case of the analysis of thermal comfort, it is possible to identify the need for physical and physiological data that comes naturally from heterogeneous sources [11, 17, 18].

**Physical Data and its Sources.** These are the result of the measurement of environmental aspects, for example temperature, humidity, air quality, airspeed, intensity of light, noise level, among others [9–11, 13]. These can be obtained through sensors such as: a) DHT11, for measuring temperature and humidity, b) MQ-135, for measuring air quality, c) KY-038, for measuring noise, and d) LDR, for measuring slight intensity.

**Physiological Data and its Sources.** These are the result of measuring aspects such as heart rate, or brain activity, among other parameters [9–11, 13]. These data is usually obtained using devices with integrating sensors, such as smartwatches, smartphones, headbands, among others.

## **2.3 Thermal Comfort**

Thermal Comfort is the level of acceptance of thermal conditions (e.g. temperature or air quality [9]) of a specific space by its occupants. TC is studied because it has been shown that it favors not only the comfort, but also the wellness, the health, the productivity and the performance of the occupants of a space [13, 19]. So, the research in this domain is aimed to determine parameters and variables to help upgrading the conditions of habitability and quality of specific spaces [19].

In this context, it is important to know about the aspects that might change the conditions of a space [20]: the space itself, external environment, and its occupants. Concerning the space itself and its external environment, variables such as radial temperature and relative humidity are of importance, while regarding the occupants, variables of interest are for instance the body mass index and the

level of isolation of thermal clothing [11, 17, 20]. Based on this discussion, the approaches for the study of Thermal Comfort are two [10, 13, 17, 18]:

- **Traditional:** It is based on the measurement of different internal and external environment aspects, and statistical models are used to analyze the collected data.
- **Varied:** It involves the analysis of unusual variables, for example, characteristics of the users. Also, the revision of adaptive models is considered, such as the adaptive model of ASHRAE 55 standard.

Thermal Comfort involves then the analysis of a highly complex context, that can be easily influenced by a large number of aspects [10, 17, 19]. However, in spite of being a highly studied topic, it is commonly tackled with a traditional perspective. The study of TC considering other physiological or even psychological variables is then of relevance [17, 18].

### 3 Proposal

The disconnection of different heterogeneous data sources is a problem that persists today [6, 12]. So, the tasks of acquiring, integrating, and analyzing diverse data flows, are challenging [3, 21]. To solve this problem, KGs are an alternative with useful advantages, since its construction implies the use of structures with the ability to represent, explore and analyze different types of data [5, 6].

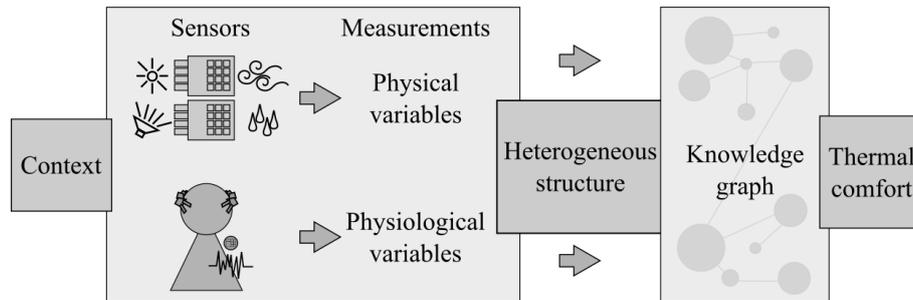
This project is aimed to develop a KG-based solution to the problem of integrating physical and physiological data sources in a dynamic context. Thus, the Thermal Comfort provide an ideal case of study in a highly dynamic environment and with distinct types of data. Moreover, in this proposal we consider to set our study in the context of a classroom. Thus, we are considering the physical and physiological variables shown in Table 1, which are considered by the ASHRAE 55 standard and the literature [17, 22–25].

**Table 1.** Physical and physiological variables.

Variables	Type	Units
Air temperature	Physical	°Celsius
Relative humidity	Physical	Water vapor % in 1 m <sup>3</sup>
Airspeed	Physical	Distance/Time
Amount of energy expended by user	Physiological	MET
User heart rate	Physiological	BPM

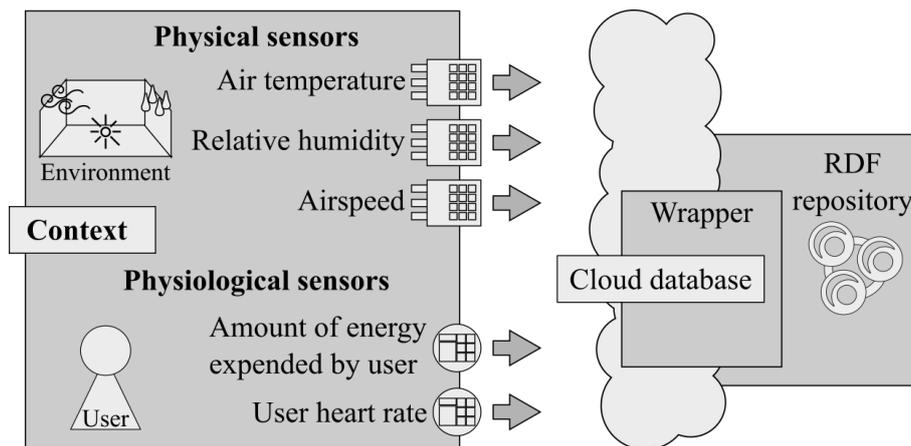
In order to tackle the problem, we propose to follow a process divided in four stages, as seen in Figure 3: a) physical and physiological data acquisition, b) integration of physical and physiological data, c) representation and analysis

of context based on the obtained physical and physiological data, and finally d) determine the Thermal Comfort of users. Each stage is further explained next.



**Fig. 3.** Process for the acquisition, integration, and analysis of data to determine thermal comfort.

Stage one is aimed to the acquisition of physical and physiological data. In this stage, we will implement three tools, one for each of the steps shown in Figure 4: a) construction of a sensor network to measuring the variables of Table 1, b) creation of a database in the cloud, and connecting the sensor network to it, c) implementation of a wrapper to export the database as a RDF repository. For this last step, it is foreseen to use the Semantic Sensor Network (SSN) ontology, that was developed by W3C based on SSO (Stimulus-Sensor-Observation) patterns, which was especially create to represent sensor information [5, 26].



**Fig. 4.** Physical and physiological data acquisition.

Stage two is aimed to integrate and analyze the acquired physical and physiological data. In this stage, we will use of the RDF repository which it made in stage one, as seen in Figure 5, and involves ones steps: a) query the RDF repository by SPARQL [27], to get RDF Triples which are an integrated representation of the physical and physiological data, b) analyze the RDF Triples acquired through the ASHRAE 55 standard methods and statistical methods according to the literature, c) use the results obtained to expand the knowledge and get ontologies.

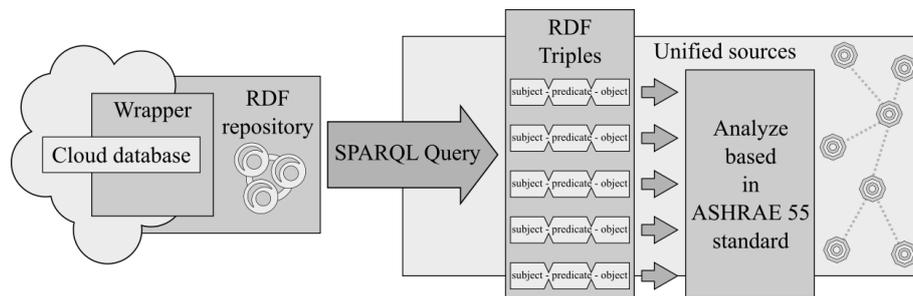


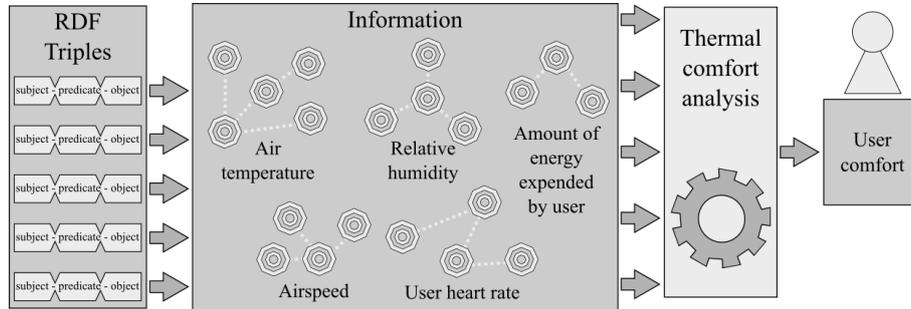
Fig. 5. Integration and analysis of physical and physiological data.

Stage three is aimed to represent and analyze the context through the RDF objects. In this stage, we will use of the RDF Triples analyzed in the stage two, to get a KG, as seen in Figure 6. To this, we will use the RDF Triples to get an ontology for each variable referred in Table 1, which we will combine with the user marks "provide a good comfort" and "provide a bad comfort" in order to distinguish the state of user comfort. Thus, each ontology will have connect with between both to make up a KG, that will have the ability to be querying by SPARQL. Also, KG obtained will have an inference engine to expand its knowledge and have more precision to know and customize the users comfort.

Finally, in stage four, based on the case study, we propose to analyze and understand the user Thermal Comfort through the application of ASHRAE 55 standard, and compare the results obtained against the knowledge that could be acquired by the KG.

## 4 Related Work

KGs enable the possibility of having a unified view for a set of heterogeneous data sources, which can be further queried and analyzed it. This has been investigated in several works. Particularly, in [5], the authors observed a big potential in the connection of an enormous amount of data that are generated by the billions of devices of the Internet of Things (IoT), and in [16], the authors similarly identified the need of unify data coming from energy systems.



**Fig. 6.** Analysis and representation of context with base in the obtained physical and physiological data.

In the case of studies on Thermal Comfort, in general they follow a traditional approach, do not exploit the use of tools for handling heterogeneous data, and they are limited to following standards such as ASHRAE 55 [17]. In addition, these studies do not usually check the influence of physiological variables on the comfort of a given environment [17, 22]. For example, in [28], the authors explained that in the south of Europe exist uninhabited and obsolete homes, which need to be rehabilitated in favor of human occupation. Given this, the authors reviewed the Thermal Comfort conditions in those homes through on a model developed for the prediction of Thermal Comfort based on Artificial Intelligence. So, they provided a mechanism capable of supporting the rehabilitation tasks in these places. However, the proposed model focused on the revision of environmental variables, and was unable to analyze the conditions of the environment in real-time.

In [29], the authors mentioned that in Cuba, the approach to determine the punctual thermal load of a building used by the companies that install HVAC systems was wrong, causing high maintenance costs and high energy consumption by this type of machines. Thus, the authors studied Thermal Comfort conditions in spaces with the double-folded objective of exposing the bad implementation of HVAC systems and of determining how to improve the Thermal Comfort level in those spaces. However, parameters related to users were not considered, only taking into the occupational affluence in those spaces.

In [30], the authors presented RoomFort, an intelligent system for the management, adjustment and improvement of the comfort conditions in hotel rooms. They used ontologies and a sensor network in hotel rooms to acquire, process and reason about different aspects related to the comfort status and how to improve it. However, although they considered parameters related to the hotel guest, this was not significant on the Thermal Comfort of the hotel room, since the study was mainly aimed to study Visual Comfort.

In [31], the authors proposed E-Cabin, a framework based on IoT which take advantage of the potential of ontologies to represent knowledge and maintain a

consistent relationship between ontology knowledge and user profiles. The authors used this framework to design and construct an intelligent cabin on a cruise, equipping the cabin with the ability to provide its inhabitants with personalized interior comfort and, at the same time, better management of the energy consumed. However, in the case of Thermal Comfort, the variables related to the user were not significant in the environment thermal settings.

In [32], the author explored the capabilities of the Wireless Sensor Network (WSN) for the study of the Human Comfort problem and the creation of Habitable Environments. Thus, he revised the potential of a WSN to analyze the environment and determine the Visual, Acoustic, Indoor Air Quality and Thermal Comfort state. However, in the case of Thermal Comfort, his analysis was based on physical parameters related to the environment.

## 5 Conclusions

The popularity that KGs have gained in the last years, has allowed to improve the form of construction of this kind of tools, highlighting the ability to manage large volumes of heterogeneous data.

Thus, KGs have served as a solution to satisfy problems related to the use of data sources of different nature, providing a base to develop systems for presenting an integral view of those sources.

Nonetheless, KGs have still some challenges related to the management, connection and integration of data. Particularly, the challenges related with acquiring and analyzing data flows coming from sources with high levels of variability and interoperability.

In this sense, Thermal Comfort analysis through KGs can provide a dynamic panorama with heterogeneous variables of multiples sources, helping and expanding the study of KG abilities to integrate, connect, storage and represent heterogeneous data sources. Also, contribute to Thermal Comfort studies, and the use of ontologies on topics related to intelligent environments and the development of spaces for human occupation.

Thus, as part of our research, this paper presents a proposal to manage and analyze heterogeneous data based on their integration through KG by knowing the user Thermal Comfort which are the study case. This proposal consists of a process of four stages: a) physical and physiological data acquisition, b) integration and analysis of physical and physiological data in a structure, c) representation and analysis of context with base on the obtained physical and physiological data, and finally d) determine the Thermal Comfort of users. Then, the knowledge obtained by this proposal could be used to develop intelligent, healthy and productive environments, as well as a help to manage heterogeneous sources.

As future work of this project, it is planned to design a network of sensors to obtain physical and physiologic sources, obtained from a specific working environment and users, respectively, to integrate them using a KG and, then analyze the users' Thermal Comfort in closed environments.

**Acknowledgments.** This work was partly sponsored by the Universidad Veracruzana (UV) and the National Council of Science and Technology (CONACYT) of México as part of the project "Infraestructura para agilizar el desarrollo de sistemas centrados en el usuario" (ref. 3053). Also, the first author was supported by CONACYT through the MSc. scholarship, Ref. 930925.

## References

1. Ehrlinger, L., Wöß, W.: Towards a definition of knowledge graphs. In: Posters and Demos Track of 12th International Conference on Semantic Systems. Volume 1695., Leipzig, Germany (2016)
2. Paulheim, H.: Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web* 8 (2016) 489–508
3. Fathalla, S., Vahdati, S., Auer, S., Lange, C.: Towards a knowledge graph representing research findings by semantifying survey articles. *Lecture Notes in Computer Science* 10450 (2017)
4. Yu, T., Li, J., Yu, Q., Tian, Y., Shun, X., Xu, L., Zhu, L., Gao, H.: Knowledge graph for tcm health preservation: Design, construction, and applications. *Artificial Intelligence in Medicine* 77 (2017) 48–52
5. Le-Phuoc, D., Mau-Quoc, H., Ngo-Quoc, H., Tran-Nhat, T., Hauswirth, M.: The graph of things: A step towards the live knowledge graph of connected things. *Journal of Web Semantics* 37-38 (2016) 25–35
6. Wang, C., Gao, M., He, X., Zhang, R.: Challenges in chinese knowledge graph construction. In: 2015 31st IEEE International Conference on Data Engineering Workshops. (2015) 59–61
7. Ernst, P., Siu, A., Weikum, G.: Knowlife: a versatile approach for constructing a large knowledge graph for biomedical sciences. *BMC Bioinformatics* 16 (2015) 157
8. Freitas, A., Da-Silva, J., Curry, E., Buitelaar, P.: Approximate and selective reasoning on knowledge graphs: A distributional semantics approach. *Data and Knowledge Engineering* 100 (2015) 211–225 Special Issue of 19th International Conference on Applications of Natural Language Processing to Information Systems.
9. De-Vecchi, R., Sorgato, M., Pacheco, M., Cândido, C., Lamberts, R.: Ashrae 55 adaptive model application in hot and humid climates: the brazilian case. *Architectural Science Review* 58 (2015) 93–101
10. Tartarini, F., Cooper, P., Fleming, R.: Thermal perceptions, preferences and adaptive behaviours of occupants of nursing homes. *Building and Environment* 132 (2018) 57–69
11. Sharif, M., Alesheikh, A.: Context-aware movement analytics: Implications, taxonomy, and design framework. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8 (2018) 1–19
12. Aggarwal, N., Shekarpour, S., Bhatia, S., Sheth, A.: Knowledge graphs: In theory and practice. In: *Conference on Information and Knowledge Management' 17*, Singapore (2017)
13. Kim, J., De-Dear, R.: Thermal comfort expectations and adaptive behavioural characteristics of primary and secondary school students. *Building and Environment* 127 (2018) 13–22
14. Ebisu, T., Ichise, R.: Toruse: Knowledge graph embedding on a lie group. In: *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*. (2018)
15. W3C: Resource description framework (rdf). <https://www.w3.org/RDF/> (2014)

16. Chun, S., Jin, X., Seo, S., Lee, K., Shin, Y., Lee, I.: Knowledge graph modeling for semantic integration of energy services. In: 2018 IEEE International Conference on Big Data and Smart Computing (BigComp). (2018) 732–735
17. Forgiarini-Rupp, R., Kim, J., De-Dear, R., Ghisi, E.: Associations of occupant demographics, thermal history and obesity variables with their thermal comfort in air-conditioned and mixed-mode ventilation office buildings. *Building and Environment* 135 (2018) 1–9
18. Lu, C.Y., Tsai, M.C., Muo, C.H., Kuo, Y.H., Sung, F.C., Wu, C.: Personal, psychosocial and environmental factors related to sick building syndrome in official employees of taiwan. In: International journal of environmental research and public health. Volume 15. (2017) 7
19. Ghaffarianhoseini, A., AlWaer, H., Omrany, H., Ghaffarianhoseini, A., Alalouch, C., Clements-Croome, D., Tookey, J.: Sick building syndrome: are we doing enough? *Architectural Science Review* 61 (2018) 99–121
20. Corona-Patricio, C., Dueñas-Sosa, C., Castro-Enríquez, J., Avendaño-Juárez, J., Olivieri, F.: Diseño e implementación de un sistema inalámbrico para el monitoreo de las condiciones ambientales externas y de confort térmico en casetas prototipo. *Pistas Educativas* 36 (2015) 207–227
21. Molero-Castillo, G., Benítez-Guerrero, E., Mezura-Godoy, C.: Interacción humano computadora y minería de datos para la generación y representación de conocimiento útil. *Ciencias de la Información* 48 (2017) 3–5
22. Kim, J., Schiavon, S., Brager, G.: Personal comfort models – a new paradigm in thermal comfort for occupant-centric environmental control. *Building and Environment* 132 (2018) 114–124
23. ASHRAE: Thermal environmental conditions for human occupancy (2017)
24. Raish, J.: (Thermal comfort: Designing for people)
25. Pancardo-García, P.: Personalización basada en sensores para la estimación del estrés por calor en ambientes laborales. PhD doctor en ciencias de la computación, Univesidad Autónoma de Aguascalientes, Universidad Juárez Autónoma de Tabasco, Universidad Veracruzana (2016)
26. Gim, J., Lee, S., Joo, W.: A study of prescriptive analysis framework for human care services based on ckan cloud. *Journal of Sensors* (2018) 10 Art. ID 6167385.
27. Doan, A., Halevy, A., Ives, Z.: Principles of Data Integration. 1st edn. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2012)
28. Escandón, R., Ascione, F., Bianco, N., Mauro, G.M., Suárez, R., Sendra, J.J.: Thermal comfort prediction in a building category: Artificial neural network generation from calibrated models for a social housing stock in southern europe. *Applied Thermal Engineering* 150 (2019) 492–505
29. Madrigal, J., Cabello, J., Sagastume, A., Balbis, M.: Evaluación de la climatización en locales comerciales, integrando técnicas de termografía, simulación y modelado por elementos finitos. *Información tecnológica* 29 (2018) 179–188
30. Spoladore, D., Arlati, S., Carciotti, S., Nolich, M., Sacco, M.: Roomfort: An ontology-based comfort management application for hotels. *Electronics* 7 (2018)
31. Nolich, M., Spoladore, D., Carciotti, S., Buqi, R., Sacco, M.: Cabin as a home: A novel comfort optimization framework for iot equipped smart environments and applications on cruise ships. *Sensors* 19 (2019)
32. Mohamed-Rawi, M.: Sensor Network Embedded Intelligence: Human Comfort Ambient Intelligence. PhD doctor of philosophy, School of Engineering Auckland, New Zealand (2013)