Design and Praxis in the Post-Digital Era

A. P. C. Buccellato
School of Architecture, University of Notre Dame, Notre Dame, Indiana, USA

C. F. Vardeman, II & H. T. Ferguson
The Center for Research Computing, University of Notre Dame, Notre Dame, Indiana, USA

ABSTRACT: Like similar revolutions transpiring in science, engineering, and industry, the future of building design will be driven by the power of structured data and the development of new computational tools, frameworks, and methodologies capable of bringing heterogeneous design data and information together in order to make it more accessible, reliable, and effective in modern practice. And like these associated disciplines, leveraging the full body of data, knowledge, and expertise in our discipline – including that which is already embedded in the built environment – will require new structures to propagate critical cross-disciplinary virtual communities and collaborative networks, new tools to facilitate the use, reuse, and sharing of knowledge, and new platforms to grow and sustain new networks of information. Authors will describe current research efforts to advance building design in the Post-Digital Era, specifically: multi-criteria decision support tools and web-based strategies to streamline the creation and discovery of building information and design data.

1 INTRODUCTION

1.1 *Motivations and Background*

The advent of ubiquitous information catalyzed by the development of the World Wide Web (WWW) and by extension the Semantic Web (Hendler & Berners-Lee 2010) has transformed the way humans work, socialize, and interpret the world. As a consequence, humans are becoming increasingly dependent on data in nearly every aspect of life. Practically speaking, this immense influx of data and information, coupled with the rapid rate of change in all disciplines, means that it is getting more and more difficult to synthesize and apply the state-of-the-art knowledge in one's own field, in closely allied fields, and even in one's own sub-disciplines. Indeed, Design Knowledge and Practice are more distributed now than ever, due in large part to vast amounts of heterogeneous data related to the design, construction, and ecological impact of the built environment that span the currently incompatible levels of open access, proprietary, and other types of special interest databases.

Meanwhile, new knowledge in allied fields and disciplines is influencing architecture in major and conceivably unprecedented ways, i.e., in digital technology and material science; in energy production and consumption by buildings; and systems innovation.

Consistent with past revolutions in discovery (Hey Tansley, & Tolle 2009), technology is once again transforming the way disciplines observe and describe the world, generate new knowledge, communicate with one another, and ultimately make impactful decisions. The instruments that are used to support design-learning and advance both the scholarship of design and the physical execution of design have evolved far beyond traditional modes. Society, at large, produces and digests information by vastly different mechanisms than just 50 years ago.

1.2 *What Does This Mean For Architecture?*

Whereas the greatest shifts in general discovery were previously achieved by innovation in empirical and theoretical models and later, by harnessing the power of advanced computational modelling (Hey Tansley, & Tolle 2009), the new paradigm for discovery in this era - that is the post-digital era and the era of “Big Data” - will be marked by newfound abilities to access and consider large amounts of dis-associated data in order to make discoveries that would not be possible in a single view of the data or from a single data set. The future of discovery in our discipline, as in others, is in the development of new tools, frameworks, and methodologies capable of bringing design data and information currently compartmentalized across areas of expertise, databases, and sources back together, in order to make the data and information more accessible, reliable, and useful in building design, construction, and operation.

Gaining access to building design data and related information presents new challenges as well as a myriad of timely opportunities that will influence the future conception and execution of the built environment, particularly as concern over the impact of the built environment on the natural environment continues to grow. When up to 90% of a building's lifetime energy use potential can be determined in the earliest stages of the design process (Keolien, Menery & Curran 1993), expanded access to reliable and useful data is essential.

The first step is to add structure to the data with formal, machine readable, conceptual models (Janowicz, Harmelen, Hendler, & Hitzler 2014), capable of facilitating more complex decision-support (Ishizaka & Nemery 2013) for architects. In order to guide the design of impact-aware buildings, methods and tools must, at minimum, be capable of assessing up-front energy costs alongside operating energy costs, repair and replacement life-cycle costs, and potential repair or replacement costs related to hazard mitigation or loss. However, current tools are not yet capable of generating this full picture – or rather, the data that is needed to achieve this type of robust, holistic analysis is not in a format that is well structured, reliable, or readily accessible.

These circumstances are changing, in light of the sheer quantity of building-related information that exists in digitally fabricated design documents, material and manufacturer databases, and even the "experiential data" in the minds of the design professionals. These developments mean that distinctly new tools and methods are needed to facilitate the use, reuse, and sharing of design data and knowledge (Gruber 1995) over distributed localities and processing components. However, in order for data to effectively influence – or inform – the design process, mechanisms that are capable of harnessing data effectively must be developed and then overlaid by frameworks capable of integrating data into the design process (Fig. 1).

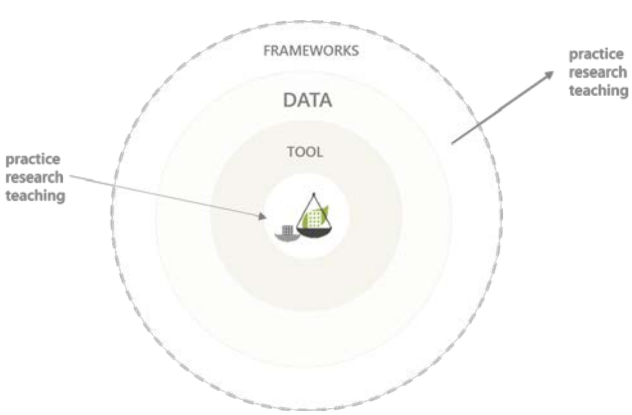

Figure 1: Integrating Data into the Design Process

2 METHODOLOGY

2.1 *Background*

Recognizing the need for enhanced tools and methods capable of weighing the broader ecological impacts of building design decisions from the outset of design is essential if the goal is to achieve the greatest possible influence over design outcome. Enhanced decision-support by these computational tools will be enabled by harnessing building design data and information in new and powerful ways. The GreenScale Research Project (The Green Scale 2014), an interdisciplinary team with expertise in architecture, computer science, and knowledge engineering, has been working along three related workflows: the development of a new multi-criteria design tool called the GreenScale Tool, theory plus use-case based approaches for Linked Open Data (LOD) methodologies that are beneficial for generating the data structures and translations required, and the creation of modern computational architecture infrastructures capable of intelligent processing in our world of distributed data processing for Decision Support System consumption.

Simultaneously, our work considers other non-trivial challenges surrounding ubiquitous data in the age of data-driven design. For example, we strive to answer questions including:

- Beyond harnessing the data itself in a structured manner, what are the threats to authorship and intellectual property?
- What are the barriers to reliable, validated data, both computational and psychological?
- What are the opportunities within these same challenges?
- How might the built world adapt to reflect our data-driven society?
- And how might unprecedented – or unfettered – access to building information and design data change the process of design, the practice of architecture, & pedagogy?

Our research commenced with a study and rigorous testing of existing analysis methods and tools in 2009. This study revealed the short-comings of many current design analysis and support tools in the face of a Big Data driven world, and meanwhile inspired the development of a novel tool and analytical data models that we discovered to be essential for advancing the testing and validation of existing tools (discussed in detail in the following section). Development of the novel tool also enabled the research of methods for expanded data discovery by tools and the development of novel strategies for achieving enhanced decision-support beyond what is currently possible with the tools tested and most-widely used in AEC (Ferguson, Buccellato, Vardeman II 2015). The tool we created, the GreenScale Tool (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016), underwent thorough testing and validation, which led to the realization and subsequent hypothesis: prevailing tools and methods are only one part of a much larger set of challenges for our built environment. We believe that data is, in fact, the more critical – and non-trivial – barrier to achieving higher resolution and reliability in building performance and impact analysis simulations.

In response, in 2013, the research team launched the Sustainability Data Initiative which brought together experts across the academic, industrial, governmental, and technology research domains to discuss the challenge of data and information related to the construction of buildings and cities, and to find opportunities to leverage existing knowledge, tools, technologies, and data to overcome the status quo. These conversations led to a series of workshop events (Geospatial Semantics Workshop and GeoVoCampDC, 2013 and GeoVoCamp, 2014) sponsored by the Spatial Ontology Community of Practice – or SOCoP (SOCoP 2013) – a community of spatial knowledge experts who develop spatial ontologies (Ontology Design Patterns.org 2010) or data pattern languages for use on the semantic web (Berners-Lee, Hendler, & Lassila 2001). The initial workshop (GeoVoCamp, 2013) was held at the National Science Foundation and lead to sustained collaborations with experts in semantic web frameworks and decision theory focused on finding practical ways to structure building design information and data and advance frameworks capable of bringing the data back to the user in new and more meaningful ways.

2.2 *A Study of Tools and Methods*

Prevailing computational tools available to and used by architecture and engineering professionals purport to gather and present thorough and accurate perspectives of the environmental impacts associated with their contributions to the built environment. Research of building modeling and analysis software used by the Architecture, Engineering, Construction, and Operations (AECO) industry reveals that many of the most heavily relied-upon industry tools are isolated in functionality, utilize incomplete models and data, and are, in fact, disruptive to normative design and building optimization workflows (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016).

At the outset of this research and as mentioned above, a study of current architectural design tools revealed large gaps in the capabilities to advance the design and execution of both data-aware and impact-intelligent buildings. A review of primary functions and limitations of prevailing tools (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016) led to a series of case studies to further evaluate the current state-of-the-art in tools and building energy analysis, followed by the development of more advanced models (Ferguson, Vardeman II, Buccellato 2015) to assess, among various factors, lifetime building energy consumption alongside operating energy use. The novel GreenScale Tool was developed in response to the functional comparison of tools and the need for this specific functionality. The fundamental goals of this particular effort, as well as some of the goals of our research, in general, are to a) better understand the efficacy of tools used by architects and engineers to influence the design process and b) to increase the use and efficacy of building impact studies conducted during the building design process.

An analysis of the current state of AECO modeling and analysis platforms (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016) revealed a number of tools that endeavor to process one type of energy analysis or another, whereas they either have limited – or no – capacity to evaluate material variations against overall building energy impact, including pre-construction material energy costs. Additionally, many of the tools are not able to operate beyond a narrow set of functions within each application, and are almost always limited to single-metric calculations, meaning that no comparative analysis can be completed without extensive manual intervention. For example, there are individual tools that are able to produce an embodied energy simulation that categorizes total energy summations into specific categories, such as transportation impacts, construction processes, and manufacturing impacts; however, if only one type of analysis is possible, separate analysis is required to anticipate lifetime building energy impact, including operating energy.

Other, distinct tools used frequently in industry can consider isolated metrics like daylighting potential, carbon emission totals, water usage, and ventilation capabilities, conduct performed carbon emission analysis and thermal calculations that include daylight and shadow. However, many of these tools function in a non-iterative manner, making it difficult to meaningfully analyze – or respond to – how changes made during the design process might ultimately influence a building's overall performance. There are also tools, such as SuAT ®, that will separate analyses into distinct categories, such as thermal flux, zoning, and climatology, and while these calculations are important, all are based on thermal calculations, which is what we refer to as single-metric. The multi-metric comparative analysis that we are referring to exists where broader and more disparate types of simulation comparisons can be made and enhanced by decision support. For example, the multi-metric analysis required by the modern industry would be able to take these thermal results and dynamically compare them to embodied energy data types, material shipping information, market costs for parts, and other data sets that are currently not compatible – but need to be.

In addition to these function-based comparisons, several experiments were conducted to gain an understanding of prevailing industry tools' relative strengths and weaknesses, focused in two key areas: capacity for estimating key metrics of building energy use in operation and initial (embodied) energy consumption. To accomplish this, simulations were conducted with a consistent set of architectural models using a select set of prevailing, commercially-available analysis tools (Fig. 2) and a control model, the GreenScale Tool (GST) (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016).

Existing Sustainability Tool Comparisons	GreenScale Tool ®	Tally ®	SUaT ®	Ecotect ®	Athena ®	GB Studio ®	DOE/BEES ®	GaBi ®
Automated and Iterative Model/Tool Interaction	x							
Zoning	x		x					
Thermal Heatflux	x		x	x			x	
Embodied Energy	x	x			x			
Embodied Water	x					x		
LCA	x	x			x			x
Global Warming Potential		x				x		x
Environmental Potentials		x				x		
Carbon Emmissions				x		x		x
Other Energy Analysis						x		
System Types			x					
Daylighting						x		
Heating/Ventilation			x					
Shadow Calculations	x		x	x		x		
Statistical Data Analysis			x					
Climatology	x						x	
Energy Star and LEED Support						x		

Figure 2: Standard AEC Tools and Respective Functions

The results of the experiments revealed large differences in calculated results for the same architectural building models. This was due, in part, to differing levels of precision and accuracy in the settings and calculation methods between the tools. Further investigation, including verification by manual calculation methods, revealed that the majority of the numerical differences (Fig. 3) resulted from discrepancies originating in the raw data values and the associated material properties that the commercial tools were using to tabulate the results. For example, prevailing analysis tools may use significantly different density values for the same material, yielding significantly different results when aggregated over an entire building. Furthermore, observations of the outputs of these tools (i.e. aggregated energy use totals, for example) demonstrate that building simulation programs used by architects today lack uniform data and processes to facilitate the consistent and reliable evaluation of both customary and novel building practices, meaning that there is a high degree of uncertainty behind the critical decisions that architects are making using these tools. At minimum, this suggests that the decision-support tools that are currently available in the marketplace may vary widely in accuracy.

Preliminary Industry Tool Comparisons	Embodied Energy	Operating Energy/Year	Total (100 Years)
Revit Energy Analysis	N/A	1.478e9 BTU	N/A
OpenStudio/EnergyPlus	N/A	2.743e9 BTU	N/A
Athena Impact Estimator	2.501e9 BTU	N/A	N/A
GreenScale Tool	2.369e9 BTU	6.044e7 BTU	8.413e9 BTU

Figure 3: Industry Standard Tool Experiment Results

Specifically, there are differences in numerical values due to rounding and approximation techniques as well as issues with the semantics of the naming system for materials, among other properties, leading to comparisons being a best guess in a lot of cases. For example, certain databases will record material data (assume for pine), but the values can also differ for the exact same pine because of unit translation precision or because they are referring to two different types of the same pine but from different geographical regions that yield slightly different wood due to different soil, rainfall, or mineral availability. As expected, the potential propagation of these differing values can make building simulations vary drastically to say the least.

The study of the existing tools, the results of the experiments conducted with them, and industry case studies allied with these concerns (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016) reveal the following:

- Single-metric analysis tools are not capable of evaluating the full energy impact of a building over its lifespan; a single view of data, or a series of independent views, is less efficient – and potentially less effective – than multi-metric analysis tied to consistent, uniform data sources.
- Whether single or parametric utilities, data and processes that are utilized by prevailing analysis tools are currently non-standard and embody too much uncertainty (in the data sourcing and thus in the outputs). Uncertainty is also not adequately considered in the decision-making process.
- Regardless of the capability of a tool (GST or otherwise), data reliability and consistency is critical to achieve a complete picture of environmental impacts, and ultimately, data is the central factor for advancing the use of simulation in architecture.

3 DATA

3.1 *Confronting the data challenge in architecture*

Gaps in data – whether in provenance, reliability, accuracy, etc. – increase uncertainty in the results handed to the user by the tool and meanwhile contribute to widening gap between intended, simulated, and achieved building performance. In the data-rich disciplines, like the life sciences, or even in more closely-allied disciplines, like engineering, gaps in data are mitigated as these disciplines turn to current computational and Big Data analytics technologies to enable broader access to and reliability of their data. When compared to other disciplines, like automotive and aerospace design, or even closely allied industries, like civil engineering, architecture is currently underutilizing these potentially very powerful technologies; our data remains largely siloed, fragmented across areas of expertise, allied in sub-disciplines, or tied up in proprietary databases.

Our discipline's relatively restricted use of advanced computational models in conjunction with limitations of the data it is currently using, means that decisions are being made by humans considering a smaller, and often skewed, scope of parameters and criteria than could be possible by utilizing Big Data practices and modern computing power applied to that data. If such methodologies could be harnessed by the building professions -- by the architects and engineers who will determine the future of the built environment -- we could achieve a much greater understanding and practical ability to affect the way that the built environment is conceived, how resources are consumed or conserved, and the effect of design choices today on the future.

3.2 Here's How: LOD Principles for Architecture

The WWW with which we are all familiar – and dependent upon – has evolved from a network of documents into a Semantic Web of intelligent data that enables access to a volume of information previously unimaginable. However, in order for humans to make sense and proper use of the vast amount of data available on the web, strategies that have been developed to support data interoperability and contextualization need to be fully implemented. Machine-readable data formats based on the web standard Resource Description Framework (RDF) (Resource Description Framework 2014) computationally appropriate data structures (Ontology Design Patterns 2010) can all be enlisted for use in modern Decision Support Systems. However, for this to be successful, we must be able to correctly describe, classify, and relate cross-domain data. The aforementioned formalizations of data enable computations to "understand" human conceptualizations and link – or connect – the data that matters to that conceptualization or specific human effort (like design).

Because ontologies have the capability to capture complex logical relationships, frameworks for promoting data sharing, discovery, and interoperability via the semantic web are being widely adopted, from the scientific research community (Kandil, Hastak, Bridges 2014) to the US government (DATA.GOV 2016) to the private sector (Samwald, Coulet, Huerga & others 2012).

There is significant ongoing research aimed at developing tools and techniques to lower the barrier to adopting semantic technologies, including several on-going efforts in architecture and allied professions to create extensive controlled vocabularies, like the Industry Foundation Classes (IFC) (Building Smart 2014), a commonly-used data model designed to facilitate interoperability within the AEC communities.

In fact, the origin of pattern languages is not remote from architecture. It was Christopher Alexander who first introduced the concept of pattern languages to describe problems and specify potential solutions in his A Pattern Language (Alexander, Ishikawa, & Silverstein), which describes design patterns for architectural design, building, and planning. The goal of Alexander's patterns for architecture is to enable people to describe their own conceptualization of design at different scales: for houses, streets, and cities. Inspired by this work, the software engineering community began to adopt software design patterns that created reusable, successful solutions to commonly-occurring programming problems. (Gamma, Helm, Johnson, Vlissides 1995). Ontology Design Patterns (ODPs) share a similar goal: to enable communities to describe their domain and domain-specific data using (their) natural vocabularies and meanwhile provide additional guidance to the data and formalization of the process.

To this end, our research group recognizes the data challenge in architecture and also the capability of ontology design patterns to tie together these data schema with varying degrees of formal specification in conjunction with existing controlled vocabularies used by the AECO community. Our research team, in collaboration with experts in ontology, developed a formal Material Transformation Pattern (MTP) (Vardeman, Krisnadhi, Cheatham & others 2014) to contextualize necessary data and provide automated reasoning support for Life Cycle Inventory analysis related to construction materials (Janowicz, Krisnadhi, HU & others 2015).

Briefly, the MTP (Fig. 4) "tells" a computer there exists a class of things in the world called a MaterialTransformation that has inputs (hasInput) and has outputs (hasOutput) that these both refer to MaterialObjects. It also tells the computer that a MaterialTransformation has some spatial extent, a Neighborhood, in which the transformation must occur and some specific time interval associated with the transformation. Lastly, we specify a class of MaterialObjects that are necessary for the transformation (manufacturing equipment, tools, etc) and that are necessary for the process to proceed, but are not part of the transformation. These we call Catalyst. Behind these logic classes are a set of machine readable axioms in formal logic, the most interesting of which expresses that a MaterialTransformation has at least one input that is not part of the output and at least one output that is not part of the input, which is the crux of the notion of transformation.

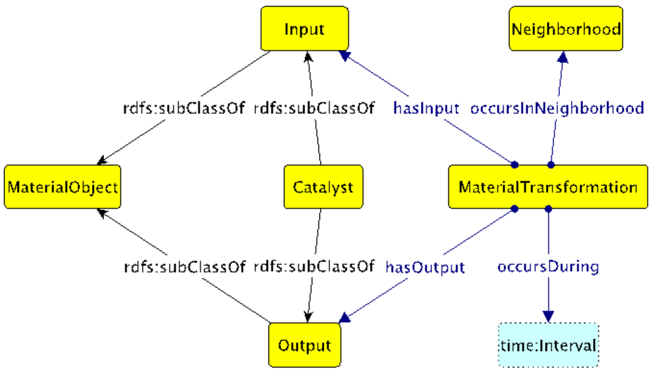

Figure 4: Material Transformation Pattern Data Model

The MTP is designed to work in conjunction with other ontology design patterns, such as the semantic trajectory pattern (Hu, Janowicz, Carral & others 2013) that defines "a path through space on which a moving object travels over time". An example of the most basic use case would be to GPS tag an object as it traverses some path through space and time resulting in a GPS data output that is a Semantic Trajectory. Such a pattern might also be applied to material supply chains to describe the path that objects take during their life cycle. MTP lets us then describe the situation in the data where one object may stop, be transformed, and continue its journey with a new identity. When a computer detects this situation in the data, it may "infer" that a transformation has occurred.

Other patterns such as "a Minimal Ontology Pattern for Life Cycle Assessment Data" provide a conceptual model for Life Cycle Assessment and Life Cycle Inventory at a higher domain specific level. This pattern can be formally "aligned" against lower level patterns, such as the trajectory and transformation patterns, to give domain practitioners a natural conceptual "view" to consume the data and ask questions of the data.

Assuming that material and material supply chain data were published as Linked Open Data using these design patterns, it is conceivable that machines could automatically discover relationships in data atoms using probabilistic machine learning algorithms, such as the multiway neural network approach that Google (Dong, Gabrilovich, Heitz & others 2014) has developed as part of its Knowledge Vault effort. Methods such as these add new information to a knowledge graph along with the probability that the relationship represented in the graph is a true relationship. This form of algorithm may hold advantages over traditional reasoning algorithms which are strictly coupled to the definitions expressed in mathematical logic.

4 FRAMEWORKS: ALIGNING DATA & TOOLS

4.1 *Rules Engine Layer*

As indicated by our initial study, many simulation tools in use today perform data analysis on a single design metric, but in the era of Big Data, tools will need the capacity to run efficiently through several metrics simultaneously and over a large quantity of design choices. To reach an optimal level of usefulness and efficacy, modern multi-criteria decision tools must be capable of performing routine calculations efficiently, considering a much broader range of data and information, including, as we propose, data harvested from heterogeneous sources linked to all aspects architectural material manufacturing, construction, and usage.

In response, Green Scale Research aims to discover more efficient methods to both aggregate and process data. This work is further motivated by our understanding that within certain existing simulation tool databases, there is syntactical and semantic ambiguity that is not adequately addressed by the application of regular expressions or even in generating the correct nominal values for material data. By extension, adjustments in energy calculations are often overlooked due to these misinterpretations because of the structural nature of Open Green Building XML-based (gbXML) (Green Building XML 2016) schemas, which are commonly implemented in architectural design tools. Additionally, the capacity of current tools to interpret XML tag-structured data is limited. However, by using the PyKE (Python Knowledge Engine 2016) rules engine, rules can be adaptively constructed to fix these discrepancies and can be extrapolated for use across other rules engines and applications.

PyKE is a knowledge-based inference engine inspired by Prolog11 (Logic Programming 2016) but implemented completely in a programming language, called Python. It is a type of logic programming allowing customized, domain-specific system rules. Incorporating rule sets with existing simulation processes means that there is an extra set of knowledge: in addition to a set of instructions, the model now has an additional set of information describing what those pieces of information mean as well as how to proceed in special cases. It also means that originally missing schema data, imperfections in schema layout, and capturing human-implicit "rules of thumb" can all be resolved computationally instead of manually.

In our own Tool study, we specifically focus on the incorporation of construction material data into a multi-metric application and develop ways to access additional data without harming the performance of a multi-metric application. Next, studies were completed to explore how these methods could communicate beneficially with the larger Semantic Web.

One of the avenues of research included developing a PyKE Rules Engine (Ferguson, Vardeman, Buccellato 2015). A rules engine is a part of a computing infrastructure that captures categories of scenarios and the associated set of actions or computations that need to happen when a certain schema instance is encountered.

For example, a rule might be to use the T-by-2T "rule of thumb" for describing spread footings. This piece of knowledge about a building structure is translated into computer code and stored electronically in what is called a Knowledge Base, so that the machine knows how to process something of this category each time certain types of building footings are encountered in computational representations. Layers of computations, such as these, lead to more accurate results than would be generated without knowing the rule(s). The ultimate goal is to seek machine-assisted solutions that will positively impact the way that the built environment is conceived and executed.

For this portion of the study, we implemented the previously developed GS Tool (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016) with sets of rules as they were needed to fill in information that is not explicit in the schema file that is handed to the tool itself. To explore experimental methods of rules communication used for handling the types and quantities of data that we anticipate means that we will need to collectively work toward bridging the gap between our tools, the larger semantic web, and decision support, all using linked open data principles (Janowicz, Hitzler, Adams & others 2012). The Green Scale (GS) Tool (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016) itself was built to seamlessly integrate with the normative workflow of the architect and to provide a broader perspective of the potential ecological impacts of design decision-making through a multi-criteria, comparative analysis including BEAM thermal heat flux model and Embodied Energy life cycle inventory model (Ferguson, Buccellato, Paolucci, Yu, & Vardeman II 2016).

4.2 What Does the Horizon Look Like

This above instance of a PyKE implementation builds on the GS Tool by adding a flexible framework for mapping regular expressions to normative conditions and by linking regular expressions to an associated action for several types of architectural design decisions (material choice, geometry discrepancies, and architectural best-practice rules).

However effective, this is only one set of solutions to a much larger problem. These rule sets and tools need to become more accessible and eventually need to be part of a larger cyberinfrastructure that can handle distributed processing. Currently the efforts of the research still function within a closed computational environment often a single machine. Some organizations are beginning to move data stores to cloud-based locations and are providing their tools as a service via the internet. Similarly, the computational components we are researching and developing need to have access to many other elements, such as reasoners, triple stores, and other computational models (The Digital Manufacturing and Design Innovation Institute 2016). Upon further development of a Linked Data Platform (W3C: Linked Data Platform 2015) used for distributed computing, the rules layers we develop, the tools we create, and the resulting knowledge bases generated can all be moved to the Cloud and subsequently enhanced and improved upon by Machine Learning [10]. Additionally, light weight (W3C: SPIN 2015) rules can be integrated with query mechanisms like the web standard (W3C: SPARQL 2015) for generating inference based connections with relevant methods in the expanding knowledge base. Additionally, SPIN could provide mechanisms for data consistency and integrity checks within the knowledge base.

Using a framework such as this, varying architectural perspectives and experience can be leveraged to generate a broader set of accessible and reliable data. Using specific types of interfacing and web services such as REST APIs (Fielding 2000), the accessibility of our data can be expanded and the framework can be made effective for cloud-computing environments. This means that advanced implementations of our research could enable wireless and web-based applications to communicate with other applications and more usefully connect a series of elements, tools, sensors, etc., within the same system or for the same simulation purposes.

5 CONCLUSION

By virtue of the mechanisms that we use to observe and describe the world, generate new knowledge, and communicate with one another, we arguably "know" more than has ever been known to civilization, but consequently – due to this proliferation of information – there is that much more to know. Yet, in many disciplines, like architecture, this abundance of information has made the design and execution of the built world more complicated – not less – and the utilization of design related data and information more challenging, due to the sheer quantity of data, the myriad places where the data "lives", and the limitations of the tools that we depend on to use the data effectively. And this is just the data that already exists. What about the future, and the data that is generated daily, in BIM, in simulations, and by our buildings?

Given the predominant use of digital design and analysis tools in building design today, a building's lifespan now begins in computer simulation. Therefore, the consequences of consistency and reliability – of both the tools and the data – impact the environment from the commencement of the design process throughout each iterative step thereafter, including the selection of materials, the methods of their assembly, and the myriad long-term implications of a building's design on the environment. And yet, without a more efficient way to discover data, we will never become more effective in harnessing it, understanding it, or, perhaps most importantly, in applying the data.

While concerns about the proprietary nature of building design – and by extension, the data "created" in the process of design – will persist, by recognizing the opportunity of utilizing existing Big Data practices along with modern computing power applied to that data, it would be possible, in this post-digital era, to advance beyond the status quo, in tools and in data, to achieve greater understanding about the impact of our design decisions and enhanced practical ability to affect the way that the built environment is conceived and how resources are consumed or conserved.

6 REFERENCES

Alexander, C., Ishikawa, S., & Silverstein, M. 1977. A Pattern Language: Towns, Buildings, Construction. NY: Oxford University Press.

Berners-Lee, T., Hendler, J., & Lassila, O. 2001. 'The Semantic Web', Scientific American.

Keoleian, GA., Menery, D., Curran, MA (1993) *Life Cycle Design Guidance Manual*.

Building Smart: International Home of Building BIM. 2014. [Online] Available: http://www.buildingsmart.org/ 2014.

Data.gov US Government. 2016. [Online] Available http://www.data.gov.

Dong, X., Gabrilovich, E., Heitz, G., Horn, W., Lao, Ni., Murphy, K., Strohmann, T., Sun, S., & Zhang, W. 2014. Knowledge Vault: A Web-Scale Approach to Probabilistic Knowledge Fusion. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 601–10. ACM.

Fielding, Roy Thomas (2000). "Architectural Styles and the Design of Network-based Software Architectures". Dissertation. University of California, Irvine.

Ferguson, H., Vardeman, C., Buccellato, A. 2015. Capturing an Architectural Knowledge Base Utilizing Rules Engine Integration for Energy and Environmental Simulations. Proceedings of the Symposium on Simulation for Architecture and Urban Design. SimAUD 2015.

Ferguson, H., Buccellato, A., Paolucci, S., Yu, N., Vardeman II, C. 2016. Green Scale Research Tool for Multi-Criteria and Multi-Metric Energy Analysis Performed During the Architectural Design Process. [Online] Available: arXiv.org.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. 1995. Design Patterns: Elements of Reusable Object-oriented Software. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Geospatial Semantics Workshop and GeoVoCamp. 2014. UW-Madison WI. [Online] Available: http://www.ssec.wisc.edu/meetings/geosp_sem/.

GeoVoCampDC. 2013. CA. [Online] Available: http://vocamp.org/wiki/GeoVoCampDC2013.

Green Building XML (gbXML). 2016. [Online] Available: http://gbxml.org/.

Gruber, T.R. 1995. Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies, 43. London; San Diego: Academic Press, c1994, pages 907–928.

Hendler, J., & Berners-Lee, T. 2010. From the Semantic Web to Social Machines: A Research Challenge for AI on the World Wide Web. Artificial Intelligence 174, no. 2: 156–61. doi:10.1016/j.artint.2009.11.010.

Hey, A., Tansley, S., & Tolle, K. 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research Redmond, WA.

Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M., & Kolas, D. 2013. A geo-ontology design pattern for semantic trajectories. In Spatial Information Theory, pages 438–456. Springer.

Ishizaka, A. & Nemery, P. 2013. Multi-criteria Decision Analysis: Methods and Software. John Wiley & Sons, Ltd.

Janowicz, K., Harmelen, F., Hendler, J., & Hitzler, P. 2014. Why the Data Train Needs Semantic Rails. AI Magazine.

Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman, C. 2014. Five Stars of Linked Data vocabulary use. Semantic Web Journal V. 5:3, pages 173-176.

Janowicz, K., Krisnadhi, A., Hu, Y., Suh, S., Pedersen, B., Weidema, Rivela, B., Tivander, J., et al. 2015. A Minimal Ontology Pattern for Life Cycle Assessment Data. Proceedings of the 6th Workshop on Ontology and Semantic Web Patterns (WOP2015) Co-Located with the 14th International Semantic Web Conference {(ISWC} 2015), Bethlehem, PA, USA, October 11, 2015.

Kandil, A., Hastak, M., Bridges, P. S. D. 2014. An ontological approach to building information model exchanges in the precast/pre-stressed concrete industry. Volume 10 pp. 9780784412329–112.

Logic Programming. 2016. [Online] Available: http://en.wikipedia.org/wiki/Logic_programming.

Matthias, S., Coulet, A., Huerga, I., Powers, R., Luciano, J., & others. 2012. Semantically enabling pharmacogenomic data for the realization of personalized medicine. Pharmacogenomics, 13, 201–212 <doi:10.2217/pgs.11.179>.

Ontology Design Patterns .org (ODP). 2010. [Online] Available: http://ontologydesignpatterns.org/wiki/Ontology_Design_Patterns_._org_%28ODP%29 2010.

Python Knowledge Engine (PyKE) 2016. [Online] Available: http://pyke.sourceforge.net/.

Resource Description Framework (RDF). 2014. [Online] Available: https://www.w3.org/RDF/.

Spatial Ontology Community of Practice (SOCoP). 2013. [Online] Available: http://socop.org/ 2013.

SPIN: SPARQL Inferencing Notation. 2015. [Online] Available: http://spinrdf.org/.

The Digital Manufacturing and Design Innovation Institute (DMDII). 2016. [Online] Available: http://dmdii.uilabs.org/.

The Green Scale. 2014. Green Scale Research Project, University of Notre Dame. [Online] Available: www.greenscale.org.

Vardeman, C., Krisnadhi, A. A., Cheatham, M., Janowicz, K., Ferguson, H., Hitzler, P., Buccellato, A., Thirunarayan, K., Berg-Cross, G., & Hahmann, T. 2014. An Ontology Design Pattern for Material Transformation. WOP2014.

W3C: Linked Data Platform 1.0. 2015. [Online] Available: https://www.w3.org/TR/ldp/.

W3C: SPARQL 1.1 Overview. 2015. [Online] Available: https://www.w3.org/TR/sparql11-overview/.