The Institution of Engineering and Technology

Data integration approaches for smarter operation of large commercial buildings

Background

Buildings are expensive not only in terms of real estate, financing and operational costs, but also in terms of environmental impacts such as climate change. Building owners and occupants are therefore increasing their focus on carbon reduction and energy management, as a means of transforming their buildings into more efficient and sustainable infrastructure. The trend for systems integration and data collection in buildings is on the rise, as end users seek to realise the benefits of cyber-physical integration through improved control and management of resources.

Most large buildings will often feature a range of technology systems (such as access control, CCTV, elevators, HVAC, etc) as well as multiple sensors, all integrated back into a building management system (BMS), SCADA, or Element Management System. For instance, in a typical 200 square metre office building, one may encounter approximately 15 systems and over 250 sensors to support BMS operation.

Recent developments in digital technologies and the Internet of Things (IoT) applications have created opportunities to collect increasing volumes of new data from individual infrastructure systems, leading to increased reliance on automated machine-to-machine (M2M) interactions, particularly in buildings.

Motivation

In order to manage a building holistically and model the implications of connections across systems, it is becoming essential to derive knowledge from across different data sources and technology domains. However, there is often little interaction between the traditional islands of building data (e.g. BMS, SCADA) and emerging systems, making it difficult to interconnect and interpret cross-domain information.

Industry experience of smart building projects has seen existing structures and technology silos being taken as a starting point, overlooking the wider needs of users and machines for better discovery, processing, and use of new, relevant data sources. Even where interconnection can be achieved, the potential range and volume of data can quickly become overwhelming without a clear and focused methodology for data management. Interoperability of emerging technologies and data management tools is therefore a key concern.

Built Environment

Information & Communications

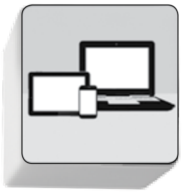
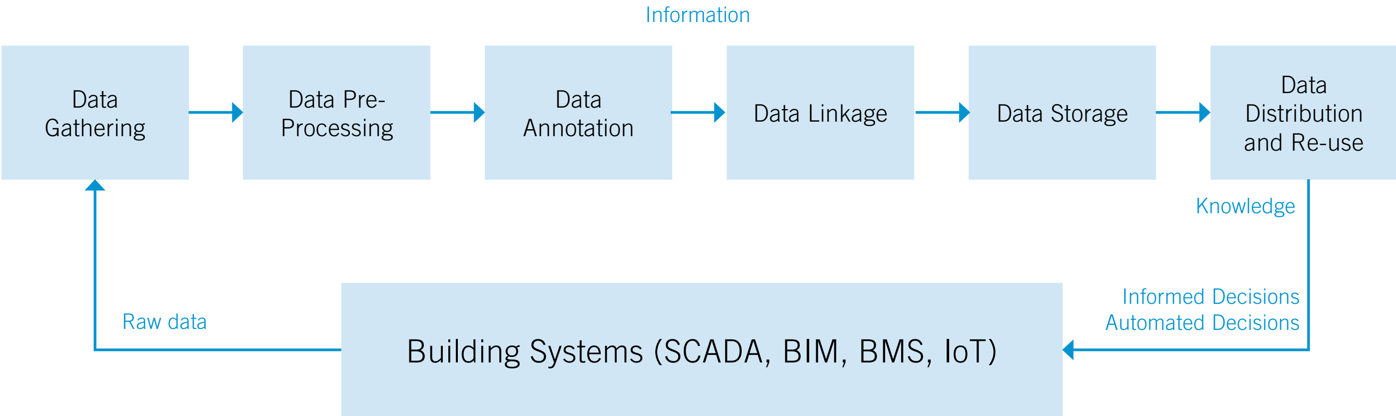


Figure 1 – Data integration in buildings



Development

Integrating data in buildings to derive the benefits of smarter operation faces numerous challenges that are often neglected, including:

- **Data find-ability**: Users can become frustrated at the existence (and prevalence) of non-reusable data sources and the lack of discoverability of appropriate data in a desired format.
- **Data understand-ability**: Lack of accuracy and timeliness of data and difficulties in defining data policies and licences. The publication of obsolete and non-valid data can result in data publication that serves no value-adding function.
- **Data usability**: Release of data that requires substantial human workload to clean it up and make it comprehensible prior to machine processing can present a resource (and motivational) barrier.
- **Data compatibility**: Lack of widely-accepted standards for expressing the syntax and semantics of data can inhibit the potential for data integration across systems/services.
- **Data quality and provenance**: Metadata structures are often insufficient or poorly documented, which can leave users unclear as to the provenance, quality and purpose of any data and the way that it was gathered and measured.

Figure 1 outlines the process for integration of data in buildings. In this figure, data gathering starts with sensors and actuators set up to capture data from activities in buildings and to enable interactivity between systems. A key point of connectivity for data collection is also required, likely to be based on standard protocols and interfaces.

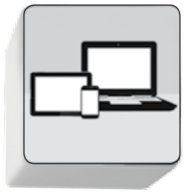
Once collected, pre-processing allows for verification of the completeness, provenance and quality of the data. Provenance can be used to give a historical account of when and how data has been produced, its ownership, and all of the necessary information to reproduce the results.

One of the most difficult tasks involves how to automatically generate the right metadata, to describe the data and how it is recorded and measured. The data must be semantically annotated and modelled in a standard and established format to provide uniform access to data. At this point, the raw data is transformed into information i.e. the context of the data is known. Sensitive information regarding users and systems needs to be protected at this stage to assure data privacy.

Once the data is modelled it can be linked to other internal or external data repositories in order to enable data discovery and reuse. The ability to reuse data, not having to collect the same data again alleviates unnecessary duplication of data and associated costs. At this point, the information is transformed into knowledge i.e. the relevance of the data is known.

Moving beyond the integration of data described, a further challenge exists in confirming how best to manage the overall process in such way that the data is easily stored, processed and retrieved across different systems and value chains, and infused back into smart buildings systems to generate informed and automated decisions.

IET The Institution of Engineering and Technology



Results

Lessons learned from earlier smart building initiatives indicate a tendency to enforce proprietary standards, or to partially adopt open standards for data communication (for instance, through deployment of cloud-based technologies to overcome interoperability issues). However, fundamental barriers to semantic interoperability remain in these cases, due to mismatch in data abstraction levels and data quality.

Recent literature on emerging technologies, shows that the cyber-physical integration at the system-of-systems level has the potential to create drivers for innovation in infrastructure.

A key effort in this domain is the Linked Data model (**http://data.gov.uk/linked-data**) as a means to overcome the barriers of data interoperability and the interconnection of disparate and fragmented silos of information. Linked data is expressed in open and non-proprietary formats, with modular approach to data instances so that datasets can be combined (mashed-up) with any other pieces of linked data.

Figure 2 illustrates a high level view of a Linked-data platform architecture. As Linked Data works at a higher information level (data / application layer) and not at the physical infrastructure-level (discrete hardware standards / software protocols), systems architectures can be designed independently and later linked at the edges, building in interoperability incrementally where it is most needed and most cost-effective.

This approach allows for scalability even when the terms and definitions of the datasets used change over time, while also enabling data exchange and reuse across different systems and by a wide range of stakeholders.

Next steps

Linked data provides a mechanism through which information silos can exist in a homogeneous and interconnected format. Representing building data, such as a Building Information Model (BIM) and the IoT, using the linked data model would allow for systems to co-exist and evolve using relevant data from a range of previously siloed domains.

By implementing Linked Data as part of a solution, organisations could generate and extract additional value from existing stand-alone repositories, across value chains and for a wider range of stakeholders. The resulting integrated, cross-domain linked building data would provide a holistic view of a building's operations and enhanced control/use of resources.

Acknowledgement

Larissa Suzuki - Researcher at Arup. Arup is an independent firm of designers, planners, engineers, consultants and technical specialists offering a broad range of professional services.

Figure 2 – Linked data platform high level view