Webs of Belief and Chains of Trust

Semantics and Agency in a World of Connected Things

Pete Rai – Cisco-SPVSS

There is a common conviction that, in order to facilitate the future world of connected things, we simply need to come to a collective agreement about vocabularies and predicates that are required to exchange data and meaning. The thinking is that (somehow) this common understanding will inevitably lead to applications that can freely exchange meaning in order to create utility in the real world. Perhaps the only other addition might be a set of covering ontologies that help to manage and formalise the predicate space.

In Cisco-SPVSS we have been examining this area via a number of internal initiatives. We have come to the conclusion that this requirement is certainly not sufficient and perhaps not even necessary.

In this short paper we will outline the main themes of our research. We will draw out and frame what we believe to be some major issues going forward. The goal of this paper is to shift the debate away from simple vocabulary standardisation and more towards underlying models of belief in the data and trust in the actions.

Uncontrolled Predicate Growth

Can we assume that people sitting in meetings will agree standards around vocabularies that will allow for the free flow of meaning? In some key areas this will certainly be true. Areas that have existing industry bodies and existing standards can rise to this challenge: Governmental systems, military systems, heavy industry, agriculture, energy suppliers, etc. These are areas where there is generally a large overlap in the motivations of the actors in the system.

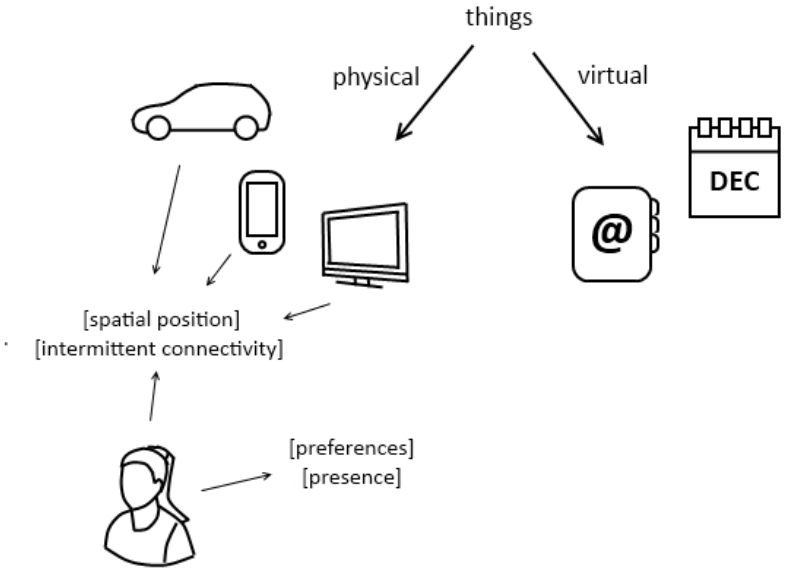
However, we believe that systems that live or die via their adoption and reception by ordinary people will undergo continuous and uncontrolled predicate growth. The data from sensors, the meaning of this data, the distillation of it into knowledge and its application to provide end-user utility, will be based on vocabularies that rise and survive by fitness and not by prescription. Like any evolved system, the space will be messy and confusing; full of overlaps and contradictions. Some predicates will have ephemeral social value based on the current buzz or zeitgeist. Whatever emerges will, by definition, be best suited to the widest definition of the problem to domain.

The challenge here is not to stop this evolutionary process, and not even to try to control or impede it. Rather the challenge is to embrace it and even encourage it, as the fastest route to the fittest semantic structures. We must accept that clean ontologies will never be possible; their imposition will in fact hinder growth. As we will outline further in this paper, we believe what is needed is to underpin this organic activity in models that allow for firm and clear concepts of belief and trust.

Capabilities not Things

Today we often talk about “things”, but in the future we might be better to consider “capabilities”. Our world is full of things; some of these things are physical, some are virtual. The physical things

have a special attribute of a spatial position. Some physical things can move about, so may experience intermittent connectivity. A special class of physical thing is a Person. At this level, they are only special because they have some unique attributes like preferences and presence.



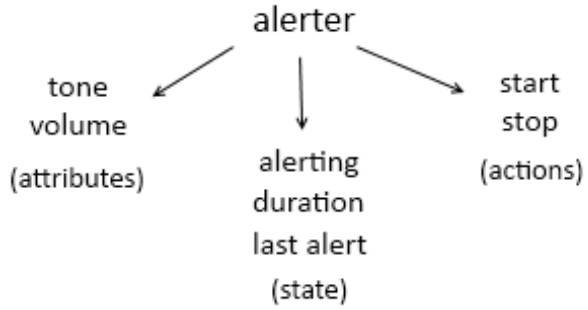
Going forwards, things can be better thought of as a collection of capabilities. These capabilities may be collected together in a physical thing for practical, historical, economic or even arbitrary purposes. We can decompose a thing into its constituent capabilities:

$$\text{thing} = \sum \text{capabilities}$$

$$\text{phone} = \text{address book} + \text{alerter} + \text{microphone} + \text{speaker} + \text{etc...}$$

$$\text{smoke alarm} = \text{smoke detector} + \text{alerter}$$

We find that, things from different real-world domains often share the same logical capabilities – in this example an “alerter”. These capabilities have their own attributes, actions and state:



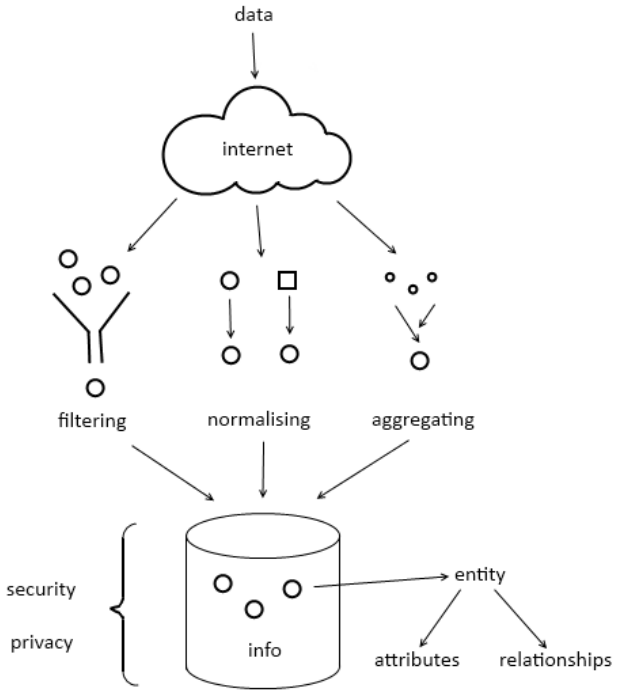
We believe that capabilities are a better model on which to hang concepts like trust, belief, privacy, security, etc than the cruder concept of things. For example, a user may be happy to allow access to her phone alarm, but reluctant to allow access to her phone call logs.

The Imposition of World Views

This world of things with capabilities produces lots and lots of data. This data may be attributes, relationships, state, preferences, etc. All this data is subject to issues of security and privacy. Some of it has a timeliness element, in that it may expire or more generally have a limited temporal value. Some of it will have an imperative element, in that it may have a great deal of importance in its own right.

In order to produce real world utility, it's likely that the data from things will at some point be placed within a structured store. It is important to understand that this act of storage imposes a world-view on the data which limits the way it can be queried, accessed and manipulated by later components. Indeed, the same underlying data may be housed in many different stores each with different world views. This, of course, is not wrong. However, all storage structures encapsulate a subjective model which inevitably means there exists an axis of trust.

For example, storing the vast and divergent data that will emerge from things will inevitably involve some steps like filtering, normalising and aggregating. Further, each world-view will likely require its own classification of things into known enumerations of entity types.



Different implementations of models and drivers will have different techniques for achieving this shoe-horning exercise. It will become important to know which implementation has been used in any given solution, as results will differ between them. Users will need to know which implementations they believe and which they do not.

Positive vs Normative Predicates

Going forward it will be important to classify predicates as either Positive or Normative. Statements made of positive predicates are objective facts, whereas statements made of normative predicates

are subjective. In a TV based system, examples of positive predicates are Title, Year of Production, Running Time, etc. Examples of normative predicates are Synopsis, Rating, Review, Genre, etc.

Generally normative data involves a creative act on behalf of a person (perhaps via algorithms that they authored). Hence this type of data has an axis of ownership. There is an intrinsic element of trust in normative data, as some sources will be more acceptable than others. You may, for example, prefer IMDB ratings to Rotten Tomato ratings. Valuable normative statements bubble-up to the top via consensus amongst users of the data.

Positive predicates are facts and there is ultimately one objectively true value. However there may well be multiple conflicting views of this fact. There is an intrinsic element of belief in positive data, as some sources will be more accurate than others. For example running times for movies may be more factual on IMDB that Wikipedia. Since positive statements are fact based (and there is a true answer) they are not owned by anyone. Valuable positive statements bubble-up to the top by being quoted by multiple independent sources.

Open World Assumption and the Domain of Consideration

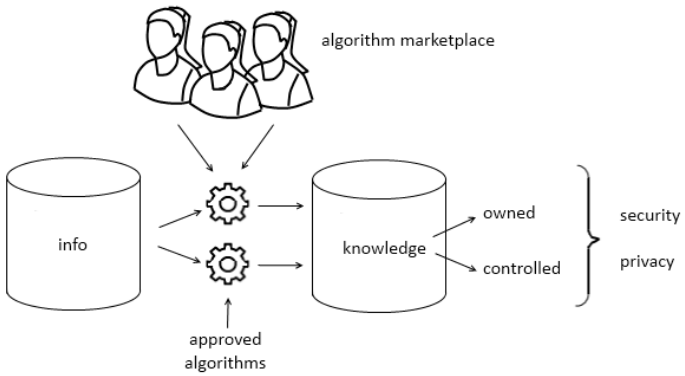
Semantic models already allow for an Open World Assumption, in that they don't assume that they have all the data in one store. Critically they don't make assumptions from the absence of data. For example, not knowing that Tom Hanks is male does not make him female.

We would like to propose that an overlapping concept is the Domain of Consideration. This is the subset of predicates that are required by an application or algorithm in order to make determinations and useful effects within their problem domain. It also encompasses the trusted sources and world views of those predicates.

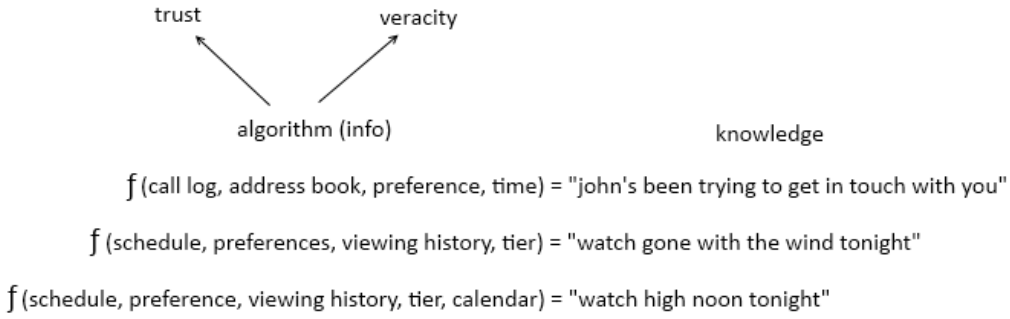
For example, one TV system may have a Domain of Consideration that does not consider at all the edit of a movie. Whereas an airline entertainment system, may lay great importance on a knowing which edit is suitable for general family viewing. Further, different airline entertainment systems may put their trust into different vendors of this information. These systems have distinct domains of consideration. Each domain is a quantifiable function of a set of predicates together with belief and trust in the sources of the predicate values.

From Information to Knowledge

Information is transformed into knowledge by the application of discrete algorithms. These algorithms create new normative information within the system. This information may require the creation of new predicates, but generally not new entity types. Anyone is free to create new knowledge through the application of their algorithms. The knowledge they create is rights owned and controlled by them. Eventually, we believe there will be an open market place for knowledge generation.



Knowledge is generally more qualitative and qualified than the information that led to its creation. The normative knowledge created in the system will be of varying value and can be duplicated, conflicting or even contradictory.



Knowledge gains value by use and affirmation via the enabling factors such as availability, reputation, trust and veracity.

Agency Control over Actuators

If applications are given direct and free access to actuators, we predict there may be a cacophony of confusing outputs that could annoy end users. There will be times when real world actions affect multiple individuals at the same time, generally because the actions effect shared resources. For example, the decision to tune a family's shared TV set. This could lead to a "fight for control" with actuators ping-ponging back and forth. Further there will be times when actions are conflicting with the behaviour of people in the real world. For example, a viewer maybe tuning the TV set manually at the same time as applications are trying to tune it.

We believe there is a role for agent technology which arbitrates and moderates the control of over actuators in order to prevent chaotic outcomes. High quality agents will apply DRY principles (Don't Repeat Yourself) by providing a hysteresis on the outputs of applications. They may also monitor how well the outputs are received by end users and qualify, prioritize or supress them as necessary.

Like an old world butler, agents will protect end users from the raw outputs that could otherwise swamp and confuse them. Different implements of agents will results in different outcomes; hence they too have an intrinsic axis of trust.