

A Linked Sensor Data Cube for a 100 year homogenised daily temperature dataset

Laurent Lefort¹, Josh Bobruk¹, Armin Haller¹, Kerry Taylor¹ and Andrew Woolf²

¹CSIRO ICT Centre, GPO Box 664, Canberra, Australia, {firstname.lastname}@csiro.au

²Australian Bureau of Meteorology, Canberra, Australia, A.Woolf@bom.gov.au

Abstract: The Australian Bureau of Meteorology (BOM) has recently published a homogenised daily temperature dataset, ACORN-SAT, for the monitoring of climate variability and change in Australia. The dataset employs the latest analysis techniques and takes advantage of newly digitised observational data to provide a daily temperature record over the last 100 years. In this paper, we present a case-study to publish the ACORN-SAT as Linked Data. We use the Semantic Sensor Network ontology to deliver the publicly available metadata about the BOM weather stations and their deployment history as linked data. Additionally, for concepts that are not covered by existing vocabularies, we have developed domain ontologies to define the adjusted aggregate variables and associated parameters for the ACORN-SAT homogenised observation data, the BOM weather stations and the BOM Rainfall districts. We use the RDF Data Cube Vocabulary to publish the originally released tabular time series data and structure it into slices to support multiple views and query endpoints. We further describe how these linked open vocabularies have been used and combined in the context of this project to make this dataset linkable to existing or future linked open data resources. We also discuss the versatility of the new service for the consumers of the ACORN-SAT dataset and uncover some issues which are specific to such long term climate data time series. The resulting Linked Sensor Data Cube is now accessible online via a pilot government linked data service built on the Linked Data API at lab.environment.data.gov.au.

Keywords. Ontology, Semantic Sensor Network, Data Cube, Time series.

1 Introduction

The Australian Climate Observations Reference Network - Surface Air Temperature (ACORN-SAT) dataset [1-3], [10], [16-17], a flagship data product of the Australian Bureau of Meteorology (BoM), has been developed for monitoring climate variability and change in Australia. To produce this dataset, climate data experts have used all the available information about weather station relocations, changes in technology and changes in observational procedures to detect breakpoints in time series and to compute adjustments for each station. The dataset provides a

daily temperature record over the last 100 years. Its primary objective is to underpin better understanding of long-term climate change.

The pilot government linked data service presented in this paper provides access to the observation network metadata and to the data via an API which allows users to retrieve subsets of the published data. We have combined and extended a number of available ontologies to develop this capability, in particular the W3C Semantic Sensor Network ontology [4] and the W3C RDF Data Cube vocabulary [6]. The originally released tabular data has been transformed into a Linked Sensor Data Cube and is now accessible online as a linked data service built on top of a Linked Data API.¹

The rest of this paper is structured as follows. In Section 2, we describe the ACORN-SAT dataset and the observation network metadata. In Section 3, we describe the role of the Semantic Sensor Network (SSN) and RDF Data Cube ontologies in the ACORN-SAT Linked Sensor Data Cube structure and in Section 4, we describe how we have built the ACORN-SAT API. In Section 5, we discuss the opportunities for climate data producers to further improve their production and publication to better match the increased demand for a transparent and reproducible data production process.

2 Overview of the ACORN-SAT data and metadata

The ACORN-SAT dataset originally released by the Bureau of Meteorology is available² as a set of tab-delimited data files which contain the homogenised minimum and maximum temperature and the raw rainfall data recorded daily at each selected site.

The temperature time series for the 112 ACORN-SAT locations are sourced from a set of single or composite stations selected according to the availability and quality of the data [17]. 10 locations have been added and one removed since the previous selection was made for the High Quality Temperature dataset [8], now superseded by ACORN-SAT. This new dataset utilises improved analysis techniques ([1], [16], [17]) which exploit pairwise comparison between one station and up to 10 comparable stations used as references. The algorithm applies differential adjustments to different parts of the daily temperature frequency distribution to better estimate the deltas with reference stations before and after an inhomogeneity. The ACORN-SAT method uses all the available background information or metadata about station moves, changes in technology and in observational procedures to identify and locate these breakpoints and to evaluate and validate the adjustments. The general knowledge used for this homogenisation process is described in several peer reviewed technical reports ([1], [16], [17]).

The site-specific knowledge is compiled in a separate station catalogue document [3] which contains the description of the 112 ACORN-SAT weather station sites. Each site is described in one page with a map and a photo of the site, the name,

¹ <http://code.google.com/p/linked-data-api/>

² <http://www.bom.gov.au/climate/change/acorn-sat/>

number, geographical coordinates and the locality of the station currently used plus the list of the nearest ACORN-SAT sites and some text about the site and its history. The information included in the site history is also available through a second document, which explains the numbering system used by the Bureau of Meteorology and the methods used to manage the changes of stations at each sites ([17], section 2.4 and 3.4). During each transition period, one of the sites, generally the old one, is kept as a comparison site [14] for a minimum period of five years of parallel observations. These modifications of the network structure are related to factors such as the urbanisation of the original site, in particular, the construction of new buildings affecting the quality of the observations, and the systematic transfer of bureau-staffed sites from city centres to airports. For example, the Darwin observations have been recorded at the Darwin Post Office (PO) from 1910 to 1942, and at two different sites at the Darwin Airport (AP), from 1941 to 2007 and from 2001 to now, with an overlap period of one year for the first transition and of just under six years for the second transition.

The BoM system allows for the same station code (014015) to be used for two different sites at different periods: from 1941 to 2001, it is used for a first location at the airport, which is later turned into a “comparison” station (014040) and from 2011 to now, it is used for a second location.

The six-letter station codes are “logical” codes used by the Bureau of Meteorology for the publication of observation data and station metadata for single or composite stations. In the other BoM systems, the raw data, from which the ACORN-SAT data is derived, appears as three separate time series with three different codes: 014016 for the Darwin Post Office station, 014015 for the Darwin Airport Station and 014040 for the Darwin Airport Comparison station. The current or last known location of a station with a given BoM code can be retrieved from the BoM Weather Station Directory.³ For Darwin, we have three different physical sites: the Post Office site and two airport sites (AP1 and AP2) which are one kilometre away from each other.

Table 1 shows the sub-phases and the key relationships between them from three distinct perspectives: the physical weather stations deployed for each phase, and the BoM and respectively ACORN-SAT time series for these stations. The information included in this table is important for users wishing to compare the homogenised ACORN-SAT data with raw or adjusted data sourced from these stations.

There is not enough useable information prior to 1997 [16] about the additional “minor” moves and the changes in the instrumentation which have occurred at each site to reconstruct the full sequence of changes.

³ <http://www.bom.gov.au/climate/cdo/about/sitedata.shtml>

Period	PHYSICAL STATIONS			COMPOSITE OR SINGLE STATIONS (BOM)			HOMOGENEISED STATIONS (ACORN-SAT)		
	Darwin PO	Darwin AP1	Darwin AP2	014016 (PO)	014015 (AP)	014040 (AP comp)	Darwin PO	Darwin AP1	Darwin AP2
1910-1941	014016			PO			014016		
1941-1942	014016	014015		PO	AP1		014016 + 014015 overlap		
1942-2001		014015			AP1			014015	
2001-2007		014040	014015		AP2	AP1	014040 + 014015 overlap		
2007-now			014015		AP2				014015

Table 1. History for the Darwin ACORN-SAT site

3 The ontologies used in the Linked Sensor Data Cube

3.1 Representation of the station history with the SSN ontology

In this paper we present OWL ontologies in the visual form produced by the Cmap Ontology Editor tool [9].

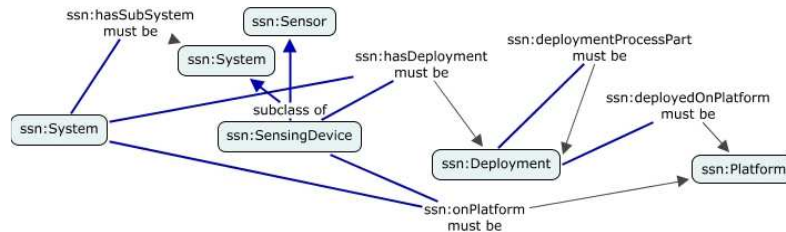


Fig. 1. Extract of the SSN Ontology

The Semantic Sensor Network Ontology [4] defines four classes and several relationships (Fig. 1) to model a system, its sensors, the deployment phases and deployment site (or platform).

To capture the historical sequence of changes at each ACORN-SAT site and manage the three perspectives discussed above, we define the ACORN-SAT network (Fig. 3) as a system of composite weather stations (`bom-station:System`) deployed on multiple sites (`bom-station:Station`).

The deployment sequence for each site is divided into three sub-phases with a pre-deployment at the beginning and a post-deployment at the end, to specify when two stations are used in parallel and a standalone phase for the middle period (Fig. 2).

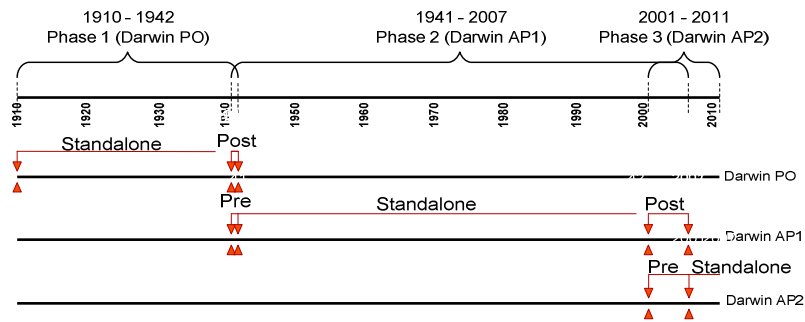


Fig. 2. Deployment phases and sub-phases for Darwin

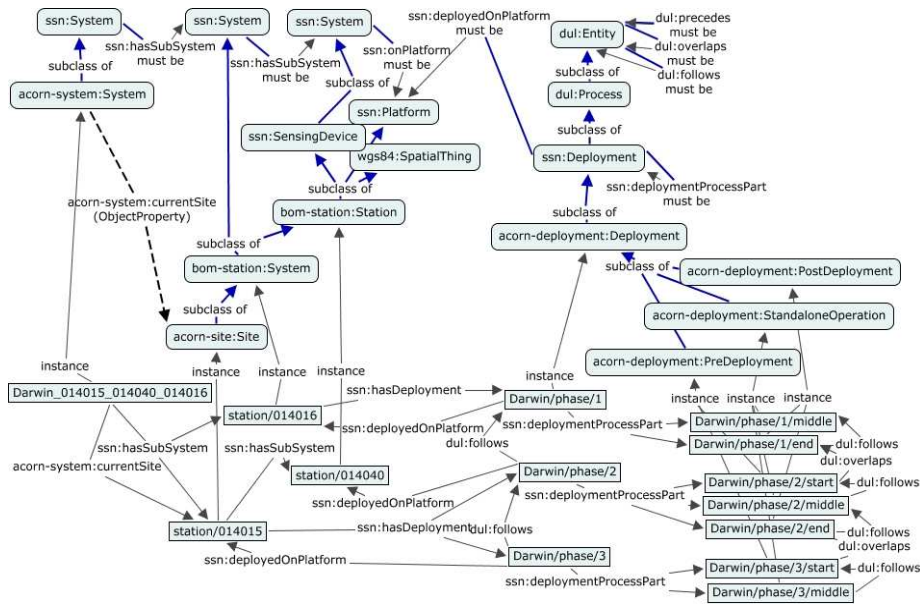


Fig. 3. ACORN-SAT system, sub-systems, deployment phases and sub-phases for Darwin

Using the SSN and the DOLCE Ultra Lite⁴ (DUL) ontologies, we can define the ACORN-SAT deployment class and its three sub-classes for the pre-deployment,

⁴ <http://www.loa-cnr.it/ontologies/DUL.owl>

standalone operation and post-deployment sub-phases. We reuse the `dul:follows` and `dul:overlaps` properties to specify the temporal relationships between phases and sub-phases (Fig. 3). The nature of the sub-phase is also documented through the URI scheme used for the instances of the deployment and sub-deployment classes shown in this figure.

3.2 Structure of the ACORN-SAT data cube

The RDF Data Cube vocabulary [6] is a vocabulary for the publication of statistical data in RDF [5], published by the W3C Government Linked Data working group.⁵ Its design has been influenced by earlier efforts like SCOVO⁶ and by the Statistical Data and Metadata Exchange (SDMX⁷) data model. The result is a versatile specification which can be applied to a large range of application domains.⁸

The design of the ACORN-SAT Linked Sensor Data Cube is based on the Bathing Water Linked Data pilot⁹ developed by the UK Environment Agency to meet its obligations under the EU Bathing Water Directive to report weekly on the water quality measured at more than 500 sampling sites. The URI¹⁰ and API schemes¹¹ developed for this project apply the URI Sets design principles defined for the UK government [7].

The ACORN-SAT data cube has four *dimensions*, one for the ACORN-SAT site and three for the date of the observation. Each *observation* contains three daily *measures*: the minimum and maximum temperature, the rainfall amount and two additional boolean *attributes* to indicate if there are missing values. Each ACORN-SAT observation refers to a date, or more precisely, to a 24 hour interval during which the maximum and minimum temperature and the amount of rainfall have been measured. For this purpose, we use the `interval:CalendarInterval` class from the UK interval ontology¹² and the URI pattern¹³ for a 24 hour period starting and ending at 09:00AM.¹⁴

The data cube itself is divided into *slices* using the site id first, and then the year and month of the observation. All the slices are compound observations enriched with some extra statistical attributes. For the temperature measures, we pre-compute the minimum, maximum, mean, and standard deviation indicators and the count of available measures for the time series period. For the rainfall measure, we have the maximum, the sum and the count of available measures. The slice size is provided to support the estimation of data quality factors e.g. the percentage of missing data

⁵ <http://www.w3.org/2011/gld/>

⁶ <http://vocab.deri.ie/scovo/>

⁷ <http://www.sdmx.org/>

⁸ <http://wiki.planet-data.eu/web/Datasets>

⁹ <http://environment.data.gov.uk/lab/>

¹⁰ www.epimorphics.com/web/wiki/bathing-water-quality-structure-published-linked-data

¹¹ <http://environment.data.gov.uk/lab/doc/api-bwq-reference-v0.2.html>

¹² <http://reference.data.gov.uk/def/intervals>

¹³ <http://www.epimorphics.com/web/wiki/using-interval-set-uris-statistical-data>

¹⁴ Example: <http://reference.data.gov.uk/id/gregorian-interval/2005-06-05T09:00:00/PT1D>

points for each measure. The start and end dates of the period are encoded as `interval:CalendarInstant` instances. Fig. 4 shows the relationships between the dataset, the slices and the observations and illustrates how the URI scheme mirrors the data cube structure. Inserting keywords like `station`, `year` and `month` into the URI makes it easier to interpret and reduces the risk of errors by end users. In the figure we use the notation “{...}” within a URI to generically represent many URIs of the indicated pattern.

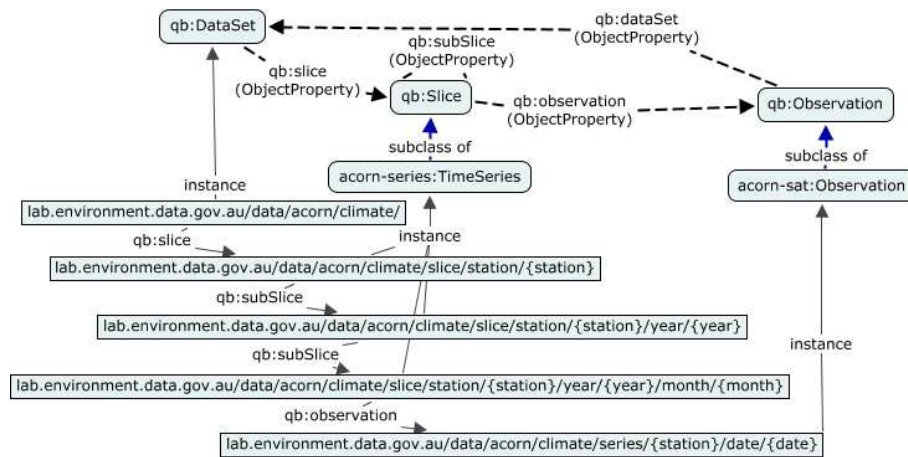


Fig. 4. Linked Sensor Data Cube structure

With such a scheme, it is possible to configure the Linked Data API¹⁵ with two types of endpoints: *item endpoints* to access the data attached to an instance and *list endpoints* to access the sub-objects (Table 2).

REST API	LDA ENDPOINT
http://lab.environment.data.gov.au/data/acorn/climate/slice/station/014015	Item data for station slice
http://lab.environment.data.gov.au/data/acorn/climate/slice/station/014015/year	List of all sub-slices

Table 2. API scheme for ACORN-SAT

3.3 Joint use of the RDF Data Cube and the SSN ontology

The SSN ontology and the RDF Data Cube vocabulary have been developed independently of each other. Here we report how they have can be integrated together.

¹⁵ <http://elda.googlecode.com/hg/deliver-elda/src/main/docs/index.html>

In the Linked Sensor Data Cube, the coupling of the two ontologies is done at two levels (Fig. 5): `ssn:Observation` and `qb:Observation` are tied together via `acorn-sat:Observation`; `acorn-sat:TimeSeries` is also defined as a `qb:Slice` and a `ssn:Observation`.

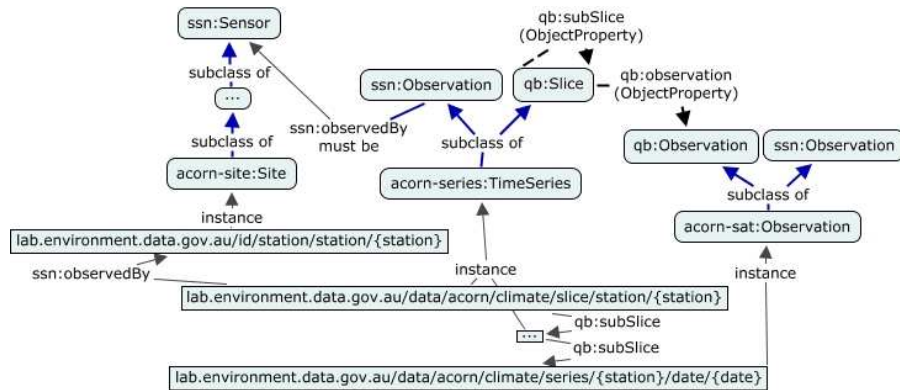


Fig. 5. Coupling of the RDF Data Cube and SSN ontologies

The links between observations and sensors are defined for time series, but not at the level of individual data points. This approach is preferred because of the size of the ACORN-SAT dataset (~ 61 million triples). For the same reason, we have not generated the links between `qb:Observation` and `qb:DataSet` instances.

The datatype properties attached to the `acorn-sat:Observation` and `acorn-series:TimeSeries` classes are imported from two purpose-made ontologies: `sat.owl` (Fig. 6) and `time-series.owl`.

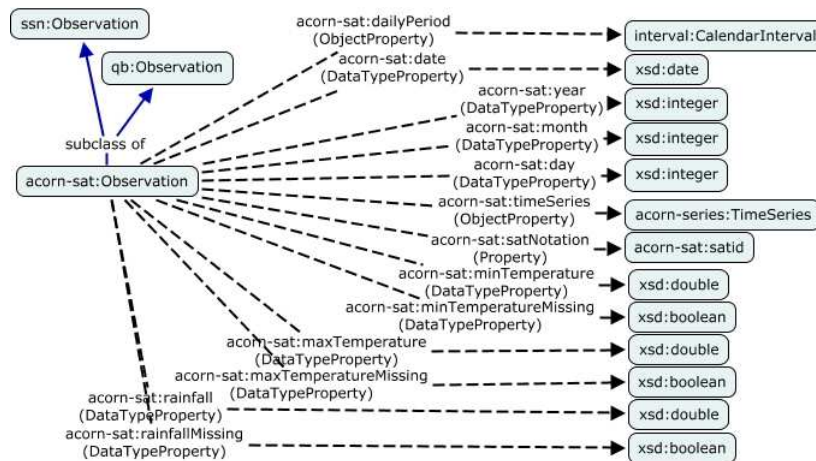


Fig. 6. The Surface-Air-Temperature observation ontology (sat.owl)

Pre-existing definitions for minimum temperature and maximum temperature are not applicable for the ACORN-SAT dataset because they do not specify the boundaries of the time interval during which the daily observations are made and that they are the output of a homogenisation algorithm used to adjust the original measurements. The dotted arrows in Fig. 6 represent domain and range axioms.

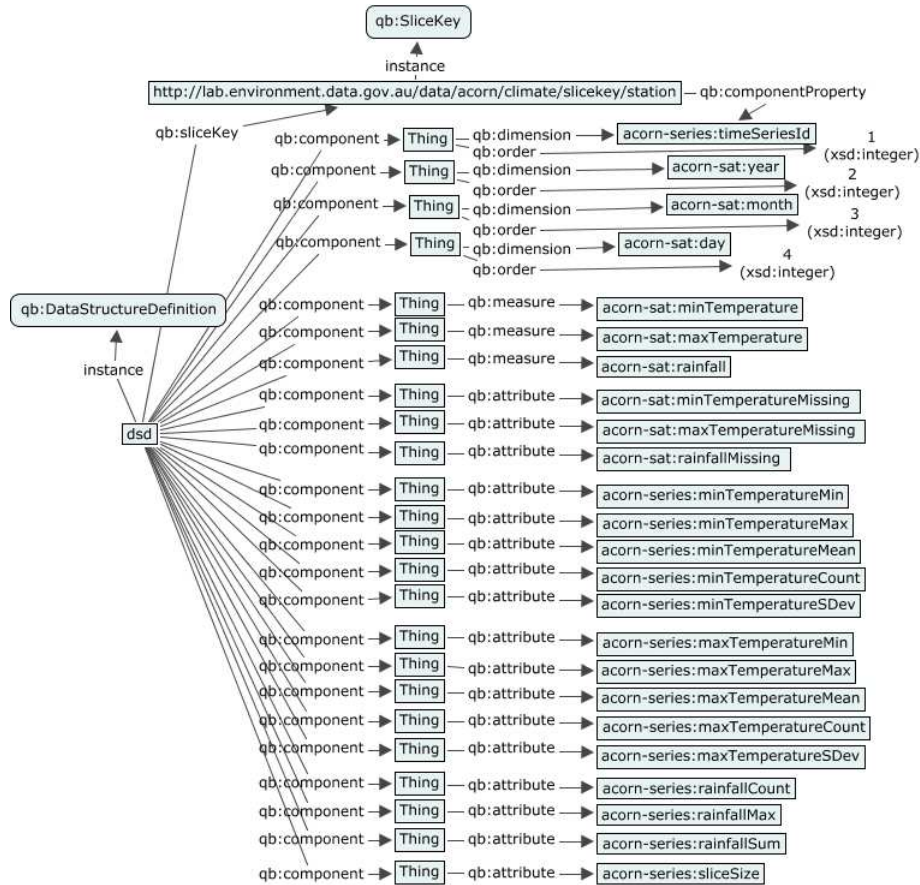


Fig. 7. The Linked Sensor Data Cube DSD (Data Structure Definition) - extract

The Expert Team on Climate Change Detection and Indices [17] has specified the statistical attributes which can be attached to monthly or annual slices. We have transformed these definitions into an ontology¹⁶ and used it to characterise the acorn-series properties when possible.

These properties are (respectively) used in the observations and time series instances and are also registered in the Data Structure Definition (DSD) instance

¹⁶ <http://purl.oclc.org/NET/ssnx/etcddi>

which documents the dimensions, measures, and attributes of the Linked Sensor Data Cube (Fig. 7).

We have found that the declarations of the observed properties in the SSN ontologies as classes are not directly compatible with their declarations in the RDF Data Cube as properties. Further investigation is needed to define a common ontology design pattern to bridge the two ontologies. The registration of the properties in the DSD is also an aspect of the RDF Data Cube vocabulary which the W3C Government Linked Data working group has identified as an area for further work.

4 The ACORN-SAT Linked Sensor Data service

The ACORN-SAT Linked Sensor Data service¹⁷ uses open source software such as the SESAME triple store¹⁸ and the ELDA¹⁹ implementation of the Linked Data API. The architecture of our solution is shown in Fig. 8.

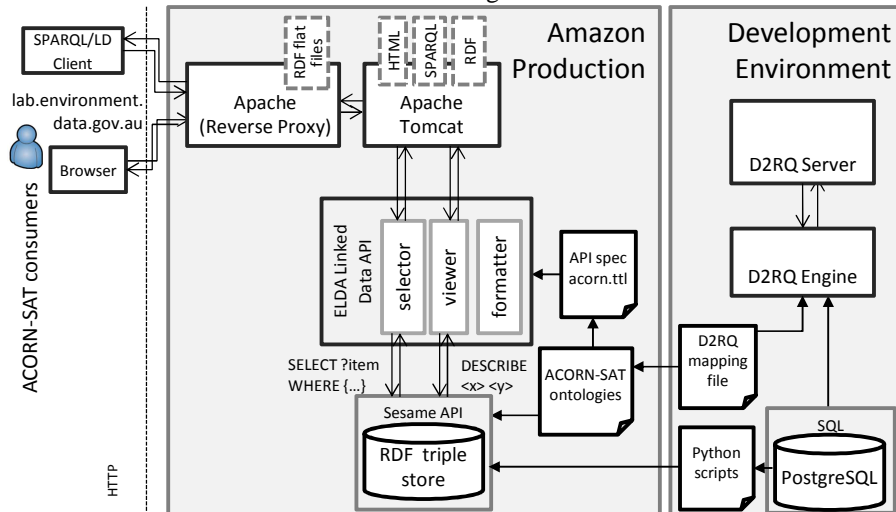


Fig. 8. Architecture of lab.environment.data.gov.au

We mapped the tabular time series data of the original ACORN-SAT to RDF using D2RQ and custom-built Python scripts. The mapping file for D2RQ is configured to produce RDF data according to the ACORN-SAT ontologies (see above). We used Jena EyeBall²⁰ to validate the linked data because of the dependencies on several RDF(S) vocabularies. Since ACORN-SAT is a relatively static data set (we expect no more than one update per year), the generation and validation tools are deployed outside of the production environment.

¹⁷ Accessible via <http://lab.environment.data.gov.au/>

¹⁸ <http://www.openrdf.org/>

¹⁹ <http://elda.googlecode.com/>

²⁰ <http://jena.apache.org/documentation/tools/eyeball-getting-started.html>

For the configuration of the Linked Data API we exploited the URI and API scheme introduced in section 3.1 to define all the item and list endpoints which give access to all the individual observations (Fig. 9, Fig. 10) and slices (Fig. 11, Fig. 12) but also to the site metadata for each time series (Fig. 13). The site data is also enriched with additional information from the Weather Station Directory²¹ in particular the rainfall district²² and state.

Due to the size of ACORN-SAT (~61million triples) we have put particular focus on the usability (performance) of the Linked Data API. We defined custom viewers for the different list endpoints in order to avoid expensive DESCRIBE SPARQL queries. Our production environment serving lab.environment.data.gov.au runs on an Amazon cloud with ELDA scaling horizontally at peak demand. Since our data is static, we only replicate ELDA and its cache, but access only a single Sesame triplestore instance.

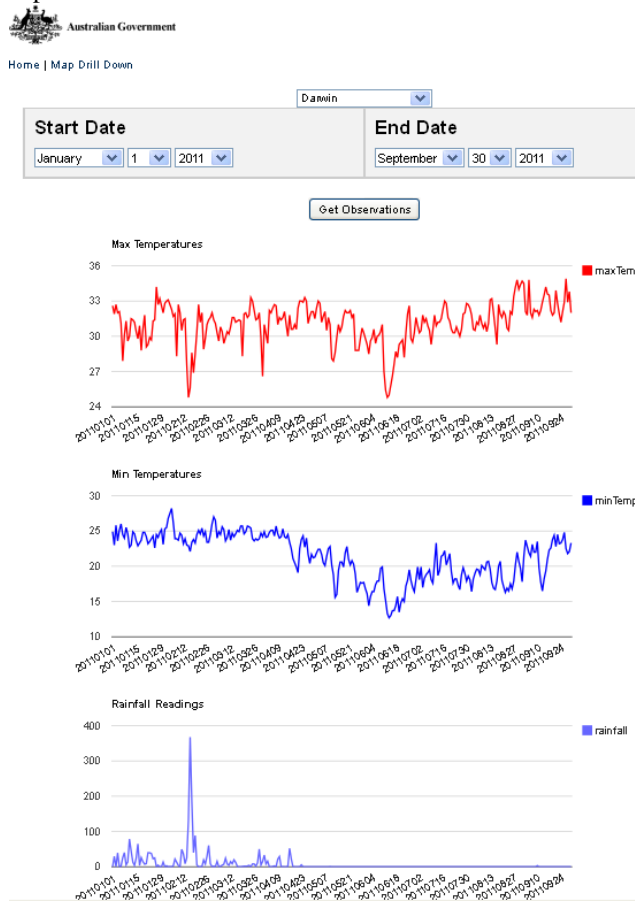


Fig. 9. Access to observation data (Darwin, January-September 2011)

²¹ <ftp://ftp.bom.gov.au/anon2/home/ncc/metadata/sitelists/stations.zip>

²² <http://www.bom.gov.au/climate/cdo/about/rain-districts.shtml>

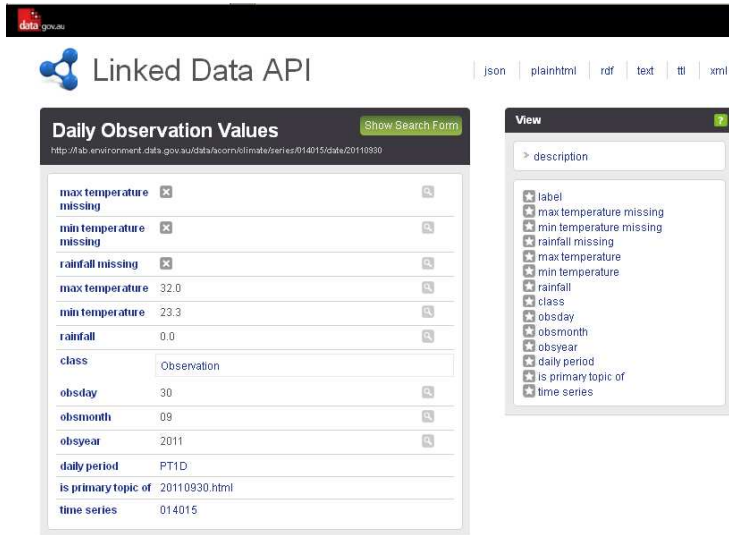


Fig. 10. Access to observation record via the API (Darwin, 30 September 2011)

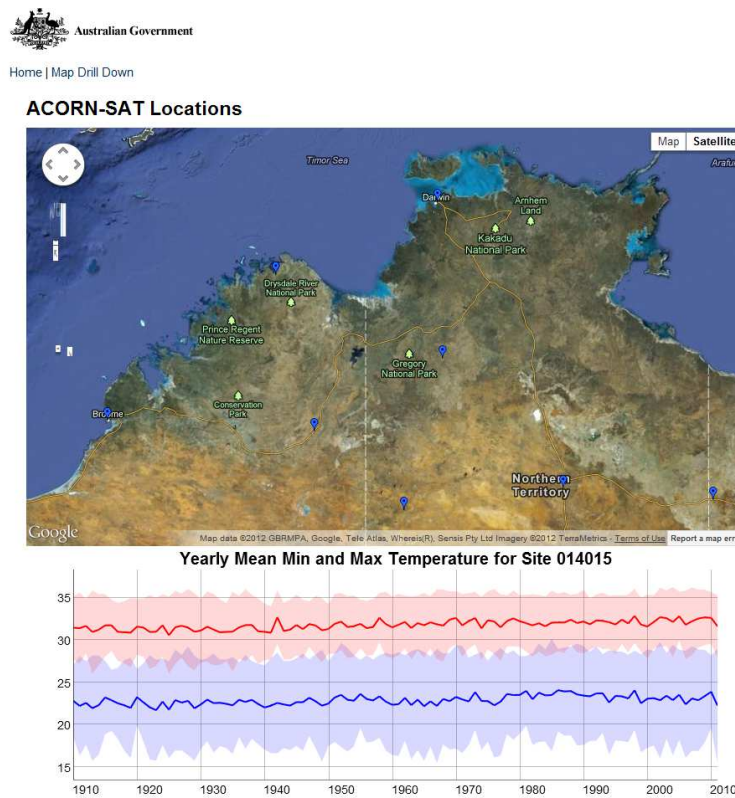


Fig. 11. 100 year min-max temperature time series data (Darwin)

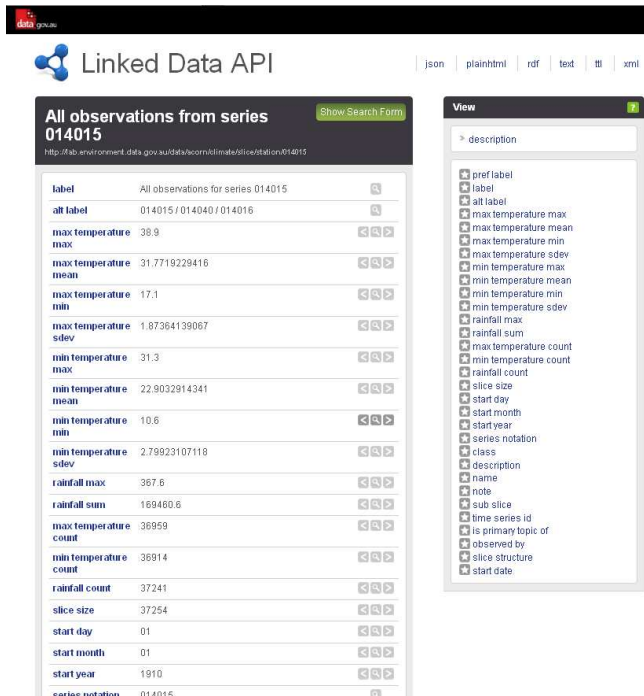


Fig. 12. Access to Darwin 100-year data via the API

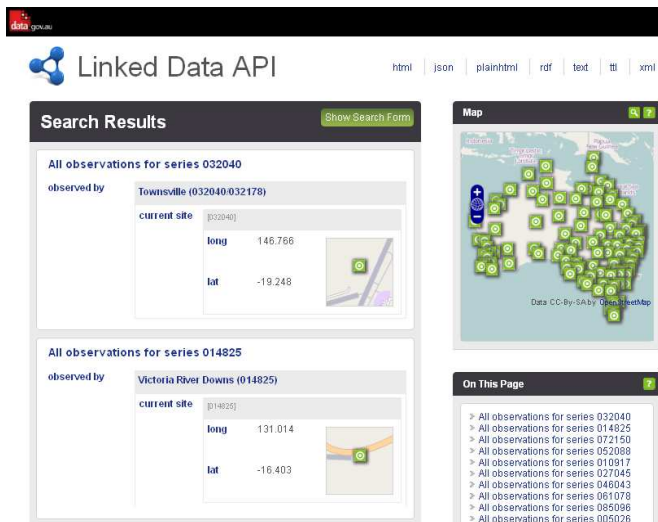


Fig. 13. Access to station metadata via the API (list endpoint)

The common foundations of the URI and API schemes (see Fig. 3 and Table 2 above) and the ability to browse the data are important features for new users wanting to rapidly build on the ACORN-SAT data.

5 Discussion

We have uncovered multiple challenges during this first attempt to apply the SSN ontology and the RDF Data Cube vocabulary to long term climate data time series. The first challenge is the volume and diversity of metadata which needs to be captured beyond the transitions and overlaps between the deployment phases discussed above. The changes in the sensor locations, technologies and observation procedures ([2], [11]) are not the only categories of events which are recorded in the Station Catalogue document [3]. For example, the construction of buildings and the growth of vegetation close to a weather station are also noted because they can have a significant impact on the quality of the observations. There are also more specific inhomogeneities discussed in other publications [17] e.g. the creation of a substantial artificial lake in Canberra (Lake Burley Griffin) about 4 km west of the observation site. Approximately half of the adjustments done on the ACORN-SAT minimum and maximum temperature values [16] are supported by metadata records of which 80% were linked to station moves. With the help of the SSN ontology, we have captured the major moves which, in Australia, have occurred mainly in the 1940s and the 1990s when the sites of observations were transferred from town centres to airports. For long-term time series of extremes measurements such as ACORN-SAT, the change of observation times are also critical. In Australia, before 1964, about 30% of the weather stations used a 0000-0000 day. In 1964, the BoM switched to the current 0900–0900 standard. We have found that the available MET ontologies define the minimum and maximum temperature but not the limits of the 24 hour period used to record them.

The second challenge is to answer the public demand to have a more transparent homogenisation process. The ACORN-SAT peer review [10] states that “a list of adjustments made as a result of the process of homogenisation should be assembled, maintained and made publicly available, along with the adjusted temperature series. Such a list will need to include the rationale for each adjustment”. We have not investigated if it would be beneficial to use the W3C PROV-O²³ ontology or to add new slices to our current data cube structure for this purpose. We know that more ontology work is needed for the definition of breakpoints, of the data structures used by percentile-based transfer functions and the parameters used by these algorithms [18].

The third challenge is to help the climate science community to create consistent and comprehensive climate data resources ([13], [15]). We have learned that the dependencies between the datasets published by national and international organisations are hard to document, partly because the numbering scheme can be a source of confusion. One possible approach is to extend the VoID²⁴ vocabulary to manage these relationships at the level of the slices of a RDF Data Cube rather than at the level of the data cube itself.

²³ PROV-O: The PROV Ontology <http://www.w3.org/TR/prov-o/>

²⁴ Vocabulary of Interlinked Datasets (VoID) <http://vocab.deri.ie/void/>

The fourth challenge is to link the ACORN-SAT data to other datasets. We provide two main ways of how other data sets can link to ACORN-SAT. Like Patni et al [12], we have linked the BoM stations to their associated GeoNames²⁵ features. We also provide temporal slices for each year of observation and consequently, we have 100 temporal slices that could be linked from other temporal data sets.

6 Conclusion

The publication of this data set represents a milestone in e-government in Australia—it is the first linked data published by the Australian Government’s open data sharing initiative known as data.gov.au.

We believe that the explicit support for metadata attachment as offered by the linked data approach presented here is of ongoing importance to the publication of climate data, and may help to enrich the public debate about the scientific foundations for climate science.

By coupling the SSN ontology and the RDF Data Cube vocabulary, we are able to capture the station history and to attach it to the data at the right level of temporal and spatial granularity. The well designed URI and API scheme makes the data cube structure easier to understand and navigate by all both producers and consumers of the linked data.

Acknowledgments

This project was funded by the Australian Bureau of Meteorology under the National Plan for Environmental Information initiative. Part of this research was conducted as part of the CSIRO Water for a Healthy Country National Research Flagship and the Sensor and Sensor Network Transformation Capability Platform.

Thanks to Alexander Coley (UK Environment Agency), Dave Reynolds, Stuart Williams, Ian Dickinson (Epimorphics) for the shared material on the UK Bathing Water project.

Thanks also to Benedikt Kämpgen (AIFB, Karlsruhe Institute of Technology and W3C Government Linked Data working group) for the suggestion to publish sensor data as statistics²⁶ and to couple the RDF Data Cube and the SSN ontologies.

7 References

1. Australian Bureau of Meteorology: Report 3a - ACORN-SAT analysis and results document (2011).
2. Australian Bureau of Meteorology: Report 4 - ACORN-SAT surface air temperature observing methods document (2011).

²⁵ <http://www.geonames.org/>

²⁶ http://www.w3.org/2011/gld/wiki/Data_Cube_Vocabulary/Use_Cases

3. Australian Bureau of Meteorology: Report 5 - ACORN-SAT station catalogue (2011).
4. Compton, M., Barnaghi, P., Bermudez, L., Garcia-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K., Kelsey, W. D., Phuoc, D. L., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K.: The SSN ontology of the W3C semantic sensor network incubator group, *Journal of Web Semantics*. Elsevier (2012).
5. Cyganiak, R., Hausenblas, M., McCuirc, E.: Official Statistics and the Practice of Data Fidelity. In: *Linking Government Data*, Wood, D. (Ed.), pp. 135-151. Springer (2011).
6. Cyganiak, R., Reynolds, D., Tennison, J.: The RDF Data Cube Vocabulary, W3C Working Draft 05 April 2012. W3C (2012).
7. Davidson, P.: Designing URI Sets for the UK Public Sector. UK Chief Technology Officer Council (2010).
8. Della-Marta, P., Collins, D., Braganza, K.: Updating Australia's high-quality annual temperature dataset. *Aust. Met. Mag.* 53(2), pp. 75-93. Australian Bureau of Meteorology (2004).
9. Eskridge, T., Hayes, P., Hoffman, R., Warren, M.: Formalizing the informal: A Confluence of Concept Mapping and the Semantic Web. In: *Proc. of the 2nd Int. Conf. on Concept Mapping*, Vol. 1, pp. 247-254 (2006).
10. Matthews, K., Peterson, T., Wang, X., Wratt, D.: ACORN-SAT Report of the Independent Peer Review Panel, Australian Bureau of Meteorology (2011).
11. Nicholls, N., Tapp, R., Burrows, K., Richards, D.: Historical Thermometer Exposures in Australia. *Int. J. Climatol.*, 16, pp. 705–710. John Wiley & Sons (1996).
12. Patni, H., Henson, C., Sheth, A.: Linked Sensor Data 2010 In: *Proc. Of the Int. Symp. on Collaborative Technologies and Systems (CTS 2010)*, pp. 362-370. IEEE (2010).
13. Peterson, T. C., Manton, M. J.: Monitoring Changes in Climate Extremes: A Tale of International Collaboration, *Bull. Amer. Meteor. Soc.* 89, pp. 1266-1271. American Meteorological Society (2008).
14. Street, R. B., Allsopp, D., Durocher, Y.: Guidelines for Managing Changes in Climate Observation Programmes WCDMP-No. 62. World Meteorological Organization (2007).
15. Thorne, P. W., Willett, K. M., Allan, R. J., Bojinski, S., Christy, J. R., Fox, N., Gilbert, S., Jolliffe, I., Kennedy, J. J., Kent, E., Tank, A. K., Lawrimore, J., Parker, D. E., Rayner, N., Simmons, A., Song, L., Stott, P. A., Trewin, B.: Guiding the Creation of A Comprehensive Surface Temperature Resource for Twenty-First-Century. *Climate Science*, *Bull. Amer. Meteor. Soc.*, 92, ES40-ES47. American Meteorological Society (2011).
16. Trewin, B.: A daily homogenized temperature dataset for Australia. *Int. J. Climatol.* (2012).
17. Trewin, B.: Techniques involved in developing the Australian Climate Observations Reference Network Surface Air Temperature (ACORN-SAT) dataset. Technical Report No. 049. Centre for Australian Weather and Climate Research (2012).
18. Williams, C. N., M. J. Menne, Thorne, P. W.: Benchmarking the performance of pairwise homogenization of surface temperatures in the United States, *J. Geophys. Res.*, 117, D05116 (2012).
19. Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., Zwiers, F. W.: Indices for monitoring changes in extremes based on daily temperature and precipitation data, *WIREs Clim Change*, 2, pp. 851–870. John Wiley & Sons (2011).