Undefined 1 (2014) 1–5
IOS Press

Energy Efficiency Measures as Linked Open Data [1]

Eva Blomqvist [a,*], Patrik Thollander [b], Robin Keskisärkkä [a] and Svetlana Paramonova [b]
[a] *Department of Computer and Information Science, Linköping University, SE-581 83 Linköping, Sweden*
E-mail:{eva.blomqvist|robin.keskisarkka}@liu.se
[b] *Department of Management and Engineering, Linköping University, SE-581 83 Linköping, Sweden*
E-mail: {patrik.thollander|svetlana.paramonova}@liu.se

Abstract. This paper describes an open linked dataset containing data on energy efficiency improvements, i.e., recommendations and measures taken based on energy audits, from both Sweden and the US, i.e., from the Swedish Energy Agency and the US Department of Energy's Industrial Assessment Centers (IAC), respectively. The overall goal of our project is threefold; (i) to facilitate better energy audits through allowing auditors and the organizations themselves to be inspired by information on measures taken earlier, in similar organisational settings, (ii) to allow researchers and policy-makers to search, compare, and assess Swedish energy audit data, and data from the US, in an integrated fashion, and (iii) to facilitate easier building of third-party applications on top of energy audit data by publishing it as Linked Open Data on the Web. The dataset is currently available through both a SPARQL endpoint, a Snorql interface, and a demonstration search interface tailored for human end-users. The data is being updated based on an ongoing manual quality control effort, and future work includes the use of the dataset to perform studies on the effects of using past energy audit data as inspiration for future recommendations for Swedish industry, as well as continuously publishing updates and extensions to the dataset itself.

Keywords: Energy Audit, Energy Efficiency, Linked Data

1. Introduction

Reduced energy consumption in industry has been identified as one of the major paths to reduce global warming [6]. Energy audits are a means of assisting companies and other organizations to improve their energy efficiency, and hence, to reduce their energy consumption. In such an audit, one of the main outcomes, apart from an overview of the current energy usage of the organisation, is a set of recommendations for improvements that the organization may implement in order to improve their energy efficiency. This could be such things as to improve the isolation of buildings to reduce heat loss in the winter, to invest in motion controlled lighting, or to make improvements or modernisations in the production processes in order to reduce energy consumption there. The focus of our dataset is precisely on such recommendations of measures, and data on their implementation. To the best of our knowledge this is the first dataset in the Linked Data cloud that focuses on energy efficiency data on such a detailed level. In the following section we briefly outline the purpose of the dataset, but for more details on the motivation, potential case studies, and more information about the ongoing quality control effort and future work, we refer the reader to [1].

[1] We would like to thank the Swedish Energy Agency for providing the funding and data for completing the projects DEFRAM and DEFRAM-2 leading to the results presented here, and we acknowledge the kind assistance of Prof. Michael R. Muller, Rutgers University, for confirming the reuse of the IAC database. We additionally give acknowledgement to Danica Ilic, Kaihong Sun, Joel Forsberg, and Joel Österqvist for the processing and quality controlling of the EKC-data.
* Corresponding author. E-mail: eva.blomqvist@liu.se.

1.1. Purpose of the Linked Dataset

Energy audits can be performed by the organizations themselves, or by a trained energy auditor. Currently researchers are investigating how to properly support both organizations and auditors in this process, and how to improve the effects of energy audits, e.g., by making sure that more improvements can be suggested, and actually implemented. One barrier to implementation has been found to be the lack of proper information [12]. This could be the lack of information about what improvements could be proposed, but also lack of information about the benefits of an improvement, e.g., how much energy can be saved, what would the improvement cost, and how long would it take for the saving to pay off that investment. Hence, one way to improve the effects of energy audits would be to provide auditors with the proper background information, e.g., what others have already done in similar situations, and how that turned out. This is one of the objectives of our published dataset.

Another possible lack of information is caused by the current coverage of our Swedish data. The data collected by the Swedish Energy Agency originates from audits between 2004 and today, and was originally mainly focused on large industries, while today also other types of organizations are studied. Nevertheless, still today there are several complete industry categories where no audit data exist. Hence, even if data is now being made accessible, organizations within those categories will find no inspiration tailored for their particular conditions. On the other hand, the Department of Energy's Industrial Assessment Centers (IAC) in the US have been collecting similar information from US organizations since the beginning of the 80's, and have information concerning most industry categories. Additionally, studies have shown that this information is indeed relevant also in a European context [2], whereby having access also to this information, in an integrated fashion, might help overcome the information deficiencies for Swedish auditors.

Furthermore, there has up until now not been an easy way for researchers and policy-makers to asses the current audit programs, e.g., in Sweden. Information about their results has been gathered by the Swedish Energy Agency, but published online in various proprietary formats, such as Excel, or even hidden inside PDF documents. Hence, researchers and policy-makers have had to gather this information from the different sources, and manually integrate it, e.g., in their own tailor made Excel-sheet, or internal database system, before being able to analyse it. Having easier access to this information, i.e., in an integrated fashion, on the Web, and in a standardised format, has the potential to greatly improve the work processes of researchers and policy-makers. It is our hope that this will in the end increase the use of the data for research and policy purposes, such as for evaluating the actual effects of the energy audit programs in Sweden.

Finally, another direction of research concerning energy efficiency audits studies the development of appropriate tools for energy auditors. Some tools exist already today, but mainly in the form of "templates" for collecting the right information, and other similarly basic tools. A major shortcoming is that these tools are not giving proper assistance to auditors in the second part of their task, i.e., recommending measures to be taken. Rather, most tools focus on gathering information about the current situation, and then leaves it up to the auditor to figure out what to suggest. However, in order to automate the process of giving suggestions, the tools need to be based on data from real cases, which has not been available to a sufficient extent until now. We envision that the dataset we have published will be used as the basis for such tools in the future.

In summary, the overall goal of our project is threefold; (i) to facilitate better energy audits through allowing auditors an organizations themselves to be inspired by information on measures taken earlier, in similar organizational settings, (ii) to allow researchers and policy-makers to search, compare, and assess Swedish energy audit data, and data from the US, in an integrated fashion, and (iii) to facilitate easier building of third-party applications on top of energy audit data by publishing it as Linked Open Data on the Web. More in detail the technical challenges of publishing this data are (a) to make data previously unavailable in machine-readable form available for reuse on the Web, (b) to transform data not currently published using open standards into RDF, (c) to make our data part of the Linked Data cloud, through reusing vocabularies and linking to existing datasets, and (d) to investigate and demonstrate the feasibility of linking and integrating energy audit data from Sweden and the US.

1.2. Dataset Usage

To illustrate the intended use of the dataset we have developed some use cases for each of the three categories of users mentioned above, i.e., (1) energy auditors and organizations being audited, (2) researchers and policy-makers, and (3) tool researchers and devel-

Table 2
Data sources used for our Linked Data (as of late 2012).

	PFE	EKC	IAC
No. of org. audited	93	225	15 570
No. of recommendations	1 256	1 438	116 960
No. of planned/impl. measures	1 216	968	54 474
Audited during time	2004-2005	2011-2012	1981-2012
Implemented during	2004-2011	2011-2014	1981–

opers. Due to space limitations we will only exemplify this with one (informal) use case for each user category. The use cases are briefly described in Table 1 as questions that may be asked by the user, which are then transformed into an example query that can be ran over the dataset.

2. Source of the Data and Topic Coverage

The dataset published consists of data from three sources; (i) the Swedish national auditing program, called PFE, managed by the Swedish Energy Agency, which mainly involves energy audits and measures in industrial organizations, (ii) a financial incentive, denoted EKC, provided by the Swedish Energy Agency for co-financing energy audits, especially for smaller organizations, where the organizations in return for the funding are required to report their achievements, and (iii) the IAC database of recommendations. Table 2 summarizes the characteristics of the data sources, and in Table 3 some additional details of the content of the sources is described (note that some information has been omitted since it is not relevant for this project, e.g., the IAC database contains a lot more information than what is mentioned here, but since this had no counterpart in the Swedish data it was not considered in our present project). For more information about the PFE, please see [10], about the EKC, please see [11] and [8], and finally, the IAC database is described by IAC on their website[1].

Names of organizations are codified in the IAC data. In a similar manner we have codified this information from the Swedish datasets. This allows for the possibility of identifying that a measure is implemented within the same organization as another one, but not for identifying the actual organization. Location of a facility is present in parts of the Swedish data, and is then included in the dataset published, however, locations for IAC data are simply given as "US" in our resulting data. Size of the audited facility is given in square meters both in IAC and EKC data and included as is in the resulting data. Industry classifications are present in all datasets, but using different standards[2]. This discrepancy has been addressed in our harmonization process, which is described in Section 3.2. PFE contains a field containing the year the audit program was started, while IAC includes the year of the actual audit, which are both reflected in our resulting data. In EKC this information is implicit from the year the report was submitted, but this is currently not made explicit in our resulting data. The type of measure is given by both PFE and IAC, but in two different classification systems, while EKC originally only had a textual description of the measure. This information has been amended in our published data, by reclassifying all data into a

Table 3
Comparison of the content of the IAC, PFE and EKC data sources (X indicating that the information is present in the source, and (X) that it can be derived or is partly available).

Data	IAC	PFE	EKC
Name of audited organization	(X)	X	X
Location of audited facility	(X)	-	X
Size of audited facility	X	-	X
Industry classification (SNI-2002)	-	X	-
Industry classification (SNI-2007)	-	-	X
Industry classification (SIC)	X	-	-
Program start year	-	X	-
Assessment year	X	-	(X)
Type of measure/recommendation	X	X	-
Free text description of measure	(X)	X	X
Reduction of electricity usage	X	X	X
Reduction of other energy sources	X	-	X
Investment cost	X	-	X
(Straight) pay back time	X	X	-
Planned implementation year	-	-	X
Actual implementation year	-	X	-
Implementation status	X	-	-
Method of verification of savings	-	X	-

[1] http://iac.rutgers.edu/database/

[2] The Swedish industry classification system called SNI has been available in several versions. SNI-2002 is an old version, while SNI-2007 is the current one, which is the national version of the NACE Rev. 2 standard prescribed by the European Union.

Table 1

Use cases and example queries (queries can be copied and ran as they are written here in the Snorql query interface). Prefix "data1:" indicates the initial release of data, other prefixes are listed in Table 4. Queries are slightly simplified, not including the retrieval of labels, nor retrieving measures with missing data elements.

User cat.	Use case	Example SPARQL query

1	An energy auditor seeks inspiration for recommendations to a given organization: Given a detailed industry classification, what measures have other such organizations implemented, within what unit process, what savings did those measures give, and what did they cost in terms of an investment?	SELECT DISTINCT ?s ?type ?description ?reduction ?unit ?carrier ?cost ?currency ?year WHERE {
		?s a energy:Measure.
		?s energy:inOrganisation ?org.
		?org org:classification <http://data.scb.se/terms/sni/10890>.
		?s energy:hasEffect ?e.
		?e energy:reduction ?reduction.
		?e energy:hasUnit ?unit.
		?e energy:reductionOf ?carrier.
		?s energy:hasInvestment ?i.
		?i energy:investmentCost ?cost.
		?i energy:hasUnit ?currency.
		?s energy:hasMeasureType ?type.
		?s dc:description ?description.
		?s energy:implementationYear ?year.
		}

2	A policy-maker seeks to evaluate how much energy was saved during one year due to measures of a certain category: What is the average saved energy in kWh for measures implemented in a certain unit process of the organizations, during one specific year?	SELECT DISTINCT ?type (COUNT(?measure) AS ?no_measures) (AVG(?saving) AS ?avg_savedkWh) WHERE {
		?measure a energy:Measure.
		?measure energy:hasMeasureType ?type.
		?measure energy:hasEffect ?e.
		?e energy:reduction ?saving.
		?e energy:hasUnit data1:kWh_perYear.
		?measure energy:implementationYear '2009'^^xsd:gYear.
		}

3	A system developer implements a function in an energy audit tool for showing the user the savings potential by implementing measures of a certain type in an organization of the same type as the current user's: What is the maximum saved energy, and of what energy carrier, for the different unit processes, concerning a specific organization type?	SELECT DISTINCT ?type (MAX(?saving) AS ?max_saving) ?unit ?carrier WHERE {
		?measure a energy:Measure.
		?measure energy:inOrganisation ?org.
		?org org:classification <http://data.scb.se/terms/sni/22>.
		?measure energy:hasMeasureType ?type.
		?measure energy:hasEffect ?e.
		?e energy:reduction ?saving.
		?e energy:reductionOf ?carrier.
		?e energy:hasUnit ?unit.
		}
		GROUP BY ?type ?unit ?carrier
		ORDER BY ?type

joint classification schema of measure categories. Reduction of energy usage, investment costs, and payback time are all included as they are, but making explicit the different units of measures that are used in the source data. Implementation year is also included, while the implementation status from the IAC data is currently not included since it has no counterpart in Swedish data, but this will be considered for future versions of the dataset. Finally, the information of verification method, in the PFE data source, is retained in our published data.

3. The Linked Dataset – Creation, Structure and Maintenance

The linked dataset published was created through a combination of manual editing and tool-based transformations, to achieve the resulting RDF file(s), which are available through our SPARQL endpoint and additional user interfaces. In the following sections the process of creating the dataset is described, as well as the resulting data itself, and its maintenance (both present activities and future plans).

3.1. Vocabulary Selection and Creation

In accordance with Linked Data best practices we have tried to reuse as much of existing vocabularies as possible. However, for a certain part of our data we were not able to find any existing vocabulary, and hence had to create our own vocabulary[3], which links to and reuses parts of other common vocabularies. In the modelling process we have used several common modelling patterns [4], such as the n-ary relation pattern (used for both the `Reduction` and `Investment` classes), and the Multi-Lingual Literal pattern (for labels in our vocabulary). Our vocabulary has also been annotated using the VOAF metadata vocabulary[4], hence it fulfils the criteria of a four-star vocabulary [7].

In the process of creating the vocabulary we found that for general concepts, such as organizations, locations, descriptions, provenance etc., there were plenty of vocabularies, and vocabulary elements, to reuse. While, for the energy- and audit-related concepts we did not find any sufficiently detailed vocabulary. There exist several energy-related linked datasets, such as Enipedia[5], Open EI[6], and Reegle[7], as well as general data including energy statistics from EUROSTAT[8] and the World Bank[9]. However, these all publish data on a higher level of abstraction, such as the energy usage of regions and countries, or general energy efficiency improvement projects (c.f. Reegle). In the end we were therefore not able to reuse any elements from these vocabularies, except for the `reegle:ProjectOutput`, which is aligned to our `energy:Measure` concept.

Table 4 presents the elements used from existing vocabularies, as well as the namespace prefixes for all the vocabularies, while our complementary vocabulary is illustrated in Figure 1. Two detailed examples of vocabulary usage are also given in Figures 2 and 3, related to Section 3.4.

[3]Current version of the OWL vocabulary is available at http://www.ida.liu.se/projects/semtech/schemas/energy/2013/09/efficiency.owl

[4]http://purl.org/vocommons/voaf

[5]http://enipedia.tudelft.nl/

[6]http://en.openei.org/

[7]http://data.reegle.info/

[8]http://eurostat.linked-statistics.org/

[9]http://worldbank.270a.info/.html

3.2. Data Transformation, Cleaning, and Re-classification

The data transformation is currently done more or less by hand, using the RDF plugin[10] for Google/Open Refine[11], by loading the source data as Excel sheets and exporting it into RDF after applying a set of transformations, for harmonizing as well as cleaning the data. As for cleaning, the source data sometimes contains information in different syntax, e.g., different date formats, misspellings etc. Such errors and discrepancies can to some extent be corrected through the standard functionalities of a tool like Google/Open Refine, e.g., by grouping data and identifying outliers.

As for harmonization, we have chosen to, for instance, remove all the codified energy carriers present in the IAC data, and replace them with a link to a node in the data representing that energy carrier, and to add an explicit link to the unit used to measure a reduction of such an energy carrier. For instance, in the IAC database the code "EC" denotes "Electrical consumption" with the implicit unit kWh, while in our dataset this code has been replaced by a link to a URI representing the concept of "Electricity" and the unit is explicitly provided as well. Another example is the use of a uniform metric prefix for the numeric data in the dataset, e.g., kWh per year, instead of a mix between MWh and kWh as in the original data. Such harmonization was not absolutely necessary for publishing the data, however, it was deemed valuable enough for the future users of the data in order to warrant the manual effort it required.

Once such cleaning and harmonization was completed, the data was also amended in two ways. First, as could be noted in Table 3 previously, all the three data sources used their own classification of measure types (or no classification at all in the case of EKC data). To be able to find measures of a certain type, or concerning a certain aspect of the organization (unit process), is a highly relevant use case for both energy auditors and researchers, hence, this problem needed to be addressed. Therefore we created a complete mapping between the respective categories used in the IAC and PFE data and a unified classification proposed by Söderström et al. [9]. Then EKC data was manually classified into the same category system.

Second, to be able to find measures proposed or implemented at an organisation similar to the one at

[10]http://refine.deri.ie/

[11]http://openrefine.org/

Table 4
Vocabularies (and datasets) from which we have reused concepts, relations or linked to data.

Name (prefix)	URI	Reused elements
DBpedia Ontology (dbpediaont:)	http://dbpedia.org/ontology/	dbpediaont:Country
DBpedia (dbpedia:)	http://dbpedia.org/resource/	dbpedia:Sweden
		dbpedia:The_United_States
Dublin Core (dc:)	http://purl.org/dc/elements/1.1/	dc:description
Energy Audit Ontology (energy:)	http://www.ida.liu.se/projects/semtech/	
	schemas/energy/2013/09/efficiency.owl#	(Our own vocabulary)
FOAF (foaf:)	http://xmlns.com/foaf/0.1/	foaf:Organization
Geo (geo:)	http://www.w3.org/2003/01/geo/wgs84_pos#	geo:location
Geonames (gn:)	http://www.geonames.org/ontology#	gn:Feature
		gn:parentFeature
Org (org:)	http://www.w3.org/ns/org#	org:classification
		org:hasSite
Provenance (prov:)	http://www.w3.org/ns/prov#	prov:wasDerivedFrom
		prov:Entity
		prov:wasGeneratedBy
		prov:Activity
RDFS (rdfs:)	http://www.w3.org/2000/01/rdf-schema#	rdfs:label
Reegle (reegle:)	http://reegle.info/schema#	reegle:ProjectOutput
SCB LKF (lkf:)	http://data.scb.se/terms/lkf/	(specific instances of Swedish
		municipalities and counties)
SCB SNI (sni:)	http://data.scb.se/terms/sni/	(specific instances of SNI-codes)
SKOS (skos:)	http://www.w3.org/2004/02/skos/core#	skos:Concept
		skos:inScheme
		skos:ConceptScheme
		skos:prefLabel
Unit (unit:)	http://www.w3.org/2007/ont/unit#	unit:Unit

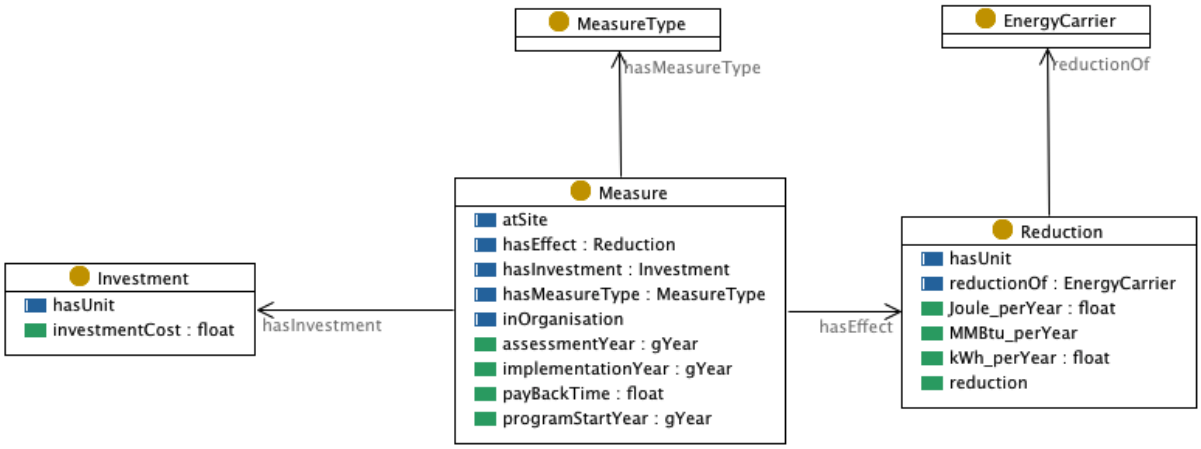

Fig. 1. An illustration (created through the ontology editing tool TopBraid Composer, using their UML-like visualisation format) of the part of the vocabulary created specifically for this dataset. Some details have been omitted due to readability reasons, e.g., the mapping of the Measure class to `reegle:ProjectOutput`.

hand, is another important use case for the data. Hence, industry classifications are an important part of the data. However, IAC use the SIC industry classification, while Swedish data use two different versions of the national SNI classification maintained by Statistics Sweden (SCB). There does not exist a one-to-one mapping between SIC and SNI-2007, but between SNI-2002 and SNI-2007 such a mapping is provided by SCB. Since we wanted to keep our Swedish data fully classified, we therefore decided to reclassify all the Swedish data into the SNI-2007 system, and then also reclassify the SIC entries in the IAC data to an SNI-2007 code. The latter unfortunately meant "loosing" some fraction of the IAC data, i.e., the part that did not have a clear mapping to the SNI-2007, which is then published without a SNI-classification, but we found that the effort of manually reclassifying the remaining IAC data was simply too large, due to the extensive amount of records (see also discussion in Section 5).

3.3. Links to Other Datasets

Before publishing the data also links to other datasets were added. We primarily link to two Swedish linked open datasets hosted by SCB. The first set is concerned with the SNI-2007 classification system, where SCB provides URI:s for each industry classification code, and provides data about the code, such as its description and preferred label. The second set is a geographical dataset, providing URI:s for all the municipalities and counties of Sweden. The latter dataset also in turn links to external sources, such as DBpedia and Geonames, but primarily provides an authoritative source of location information for the Swedish administrative geography. For the subset of audit results where the detailed location of the audited organization is given, we have linked to this dataset, using the URI:s for the municipality and county. However, for the cases where no detailed location information was available, we have instead simply linked to the DBpedia entries for Sweden and the US, respectively.

3.4. Results and Data Access

Two examples of resulting data are given in Figures 2 and 3. At the top of the figures the vocabulary, in terms of used classes, can be seen, while in the middle the use of properties is illustrated through some example data. At the bottom we illustrate what data is given as literal values in the data. For readability reasons most `rdfs:label` values have been omitted in the figures. Data is available as an RDF dump[12], through a SPARQL endpoint[13], through an adapted Snorql graphical query interface[14], and through a demo search interface[15], and is catalogued through its page on datahub.io[16]. The dataset is available as open data, under a Creative Commons attribution license[17].

3.5. Maintenance, Versioning and Quality Control

Data is currently being updated through a follow-up project to the original project publishing the data. Hence, during 2014 new versions of the dataset are being published continuously. Primarily these updates concern the improvement of data quality (for details on quality issues discovered, and how they have been addressed, see [1]), and addition of new data that has become available since 2013. The work on data quality has been focused on the Swedish data sources only, and also resulted in ideas for updated methods for quality assurance already at the data source owner, i.e., the Swedish Energy Agency.

In the future we are planning to maintain the dataset by adding data from recent energy audits in the coming years as far as our financing permits, and we will also together with the Swedish Energy Agency investigate the possibility of partly automating the data publication process. The first dataset release has also been made available through Svensk Nationell Datatjänst[18], a national data archive for long-term preservation of research data. Nevertheless the currently available endpoint and interfaces will be kept online also in the foreseeable future.

When an update of the data is made the data in the triple store gets a new URI, representing the new version of the data. The version information is encoded in the URI, using the year and month of the update. For instance, the first release of the data used the URI http://www.ida.liu.se/projects/semtech/energy/effektivisering/2013/09/ while the current release uses the following URI http://www.ida.liu.se/projects/semtech/energy/effektivisering/2014/06/. Also the vocabulary is being versioned in the same way, i.e., by changing the URI and replacing the year and month with the current ones.

[12] http://www.ida.liu.se/projects/semtech/energy/Energy_201309.zip
[13] http://www.ida.liu.se/projects/semtech/openrdf-sesame/repositories/energy
[14] http://www.ida.liu.se/projects/semtech/energy/snorql/
[15] http://www.ida.liu.se/projects/semtech/energy/demo/
[16] http://datahub.io/dataset/energy-efficiency
[17] http://creativecommons.org/licenses/by/4.0/
[18] http://snd.gu.se/sv/catalogue/study/SND0935

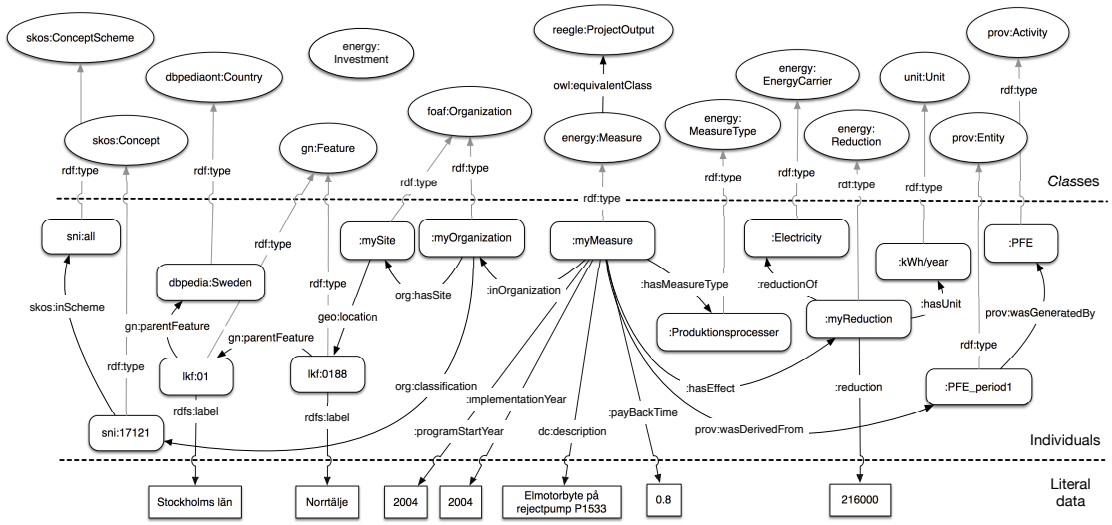

Fig. 2. Example data illustrating one measure originating from the PFE project as represented in our dataset, using our vocabulary (energy:). Entries given without namespace prefix are defined in the relevant version of the local data namespace.

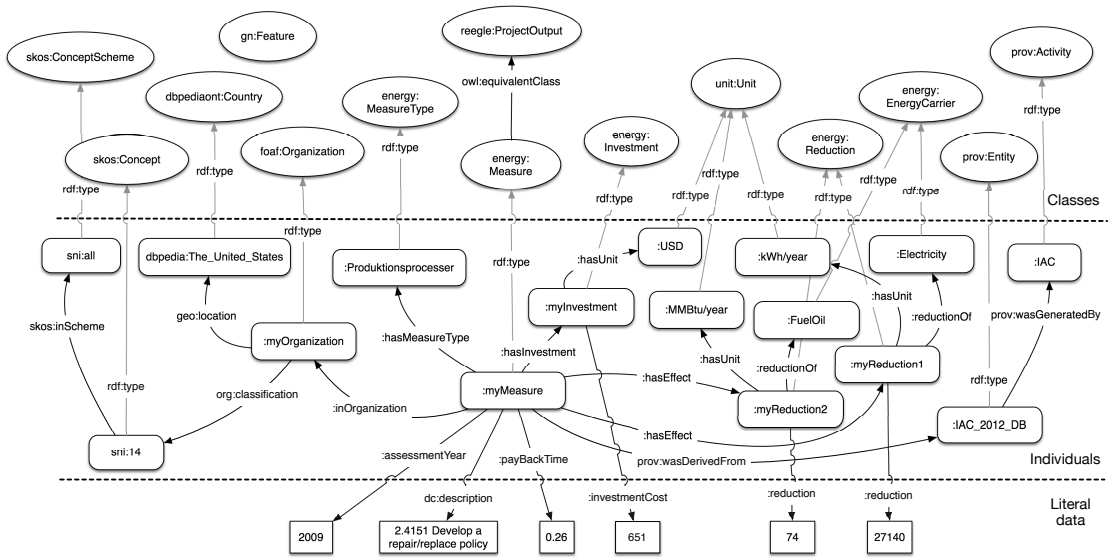

Fig. 3. Example data illustrating one measure originating from the IAC as represented in our dataset, using our vocabulary (energy:). Entries given without namespace prefix are defined in the relevant version of the local data namespace.

4. Dataset Usage

The Swedish part of the data, has so far been used in a governmental ex-ante study on the effect of a future Swedish technology procurement policy program for industry [5], in ongoing research, national and international [13], e.g. in an IEA-project, ex-post evaluation of the Swedish energy audit program, and studies on district heating potential in Swedish industry [3]. It should be noted that the cited work used the Swedish data, thus primarily showing the usefulness of the data as such, and also the usefulness of having a uniform categorization of data, which enabled program comparison [5]. However, in e.g., energy program evaluation, implemented measures could now, with our created dataset be compared from a bottom-up perspective, something which up until now has not been pos-

sible in policy program evaluation. The linked dataset is also being used by the Swedish Energy Agency as a demonstration of feasibility for publishing open data and linking data between US and Sweden. Future plans also include the utilization of the data in a software application for supporting energy auditors.

5. Lessons Learned and Known Issues

During our work with this dataset we have encountered some interesting issues that had to be addressed. Several issues concerned the harmonization and reclassification that was applied to the data. Before the project started we envisioned to spend more time on this than what was actually necessary. However, we avoided the task of creating a complete mapping between the lower level classifications of SIC and SNI-2007, which would have taken much more time (if even possible to find a complete one-to-one mapping). In fact, we believe that a more feasible route might be to describe the two classifications, using for instance SKOS, and then use the links, e.g., `skos:relatedMatch`, instead of a complete reclassification, but this is still future work.

Also multilinguality is an issue that is not completely addressed today. For instance, the measure types are currently only described in Swedish. While a translation to English now exists, it is not part of the dataset today. Also other things, such as the textual descriptions of proposed measures, are currently only available in one language, i.e., Swedish or English, not both. While RDF and Linked Data caters for an easy way to express language information, e.g., through language tags, the actual data still needs translation.

Finally, in this project we have focused on data itself, and not on the services surrounding it. Hence, the demonstration search interface is a "quick-and-dirty" prototype, which makes it less user friendly than it could be, and we cannot guarantee uptime, nor response time. Transferring the data and interfaces to a production environment is still future work. Additional future work also concerns to include all the IAC data, i.e., not only the part that has a Swedish counterpart, and include more recent data from IAC (from 2012 and onwards).

References

[1] Eva Blomqvist and Patrik Thollander. An integrated dataset of energy efficiency measures published as linked open data. *Forthcoming*.

[2] Enrico Cagno, P. Trucco, Andrea Trianni, and G. Sala. Quick-e-scan: A methodology for the energy scan of SMEs. *Energy*, 35(5):1916 – 1926, 2010.

[3] D. Djuric Ilic and L. Trygg. Economic and environmental benefits of converting industrial processes to district heating. *Energy Conversion and Management*, 87:305–317, November 2014.

[4] Leigh Dodds and Ian Davis. *Linked Data Patterns - A pattern catalogue for modelling, publishing, and consuming Linked Data*. (Online book), 2012.

[5] P-Å. Franck, A. Åsblad, P. Thollander, and P-E. Nilsson. Teknikupphandling inom industrisektorn [technology procurement within the industrial sector]. [in swedish], CIT (Chalmers Industrial Management), Gothenburg, 2014.

[6] IPCC. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers. Retrieved October 8, 2007, from: http://www.ipcc.ch/SPM0405 07.pdf, 2007.

[7] Krzysztof Janowicz, Pascal Hitzler, Benjamin Adams, Dave Kolas, and Charles Vardeman II. Five stars of Linked Data vocabulary use. *Semantic Web Journal*, 5(3):173–176, 2014.

[8] M. Karlsson, P. Thollander, and P. Rohdin. Evaluation of the swedish national energy audit program for small and medium sized industries. In *Proceedings of the ECEEE Industry Summer Study, 11-14 September, 2012*, 2012.

[9] Mats Söderström, Per Carlsson, and Ola Hall. ENHETSPROCESSER - Ett sätt att strukturera industrins energianvändning. Technical report, Linköping University and Sydkraft Konsult, 1994.

[10] Christian Stenqvist and Lars J. Nilsson. Energy efficiency in energy-intensive industries—an evaluation of the Swedish voluntary agreement PFE. *Energy Efficiency*, 5(2):225–241, 2012.

[11] Patrik Thollander and Erik Dotzauer. An energy efficiency program for swedish industrial small- and medium-sized enterprises. *Journal of Cleaner Production*, 18(13):1339–1346, September 2010.

[12] Patrik Thollander and Jenny Palm. *Improving Energy Efficiency in Industrial Energy Systems - An Interdisciplinary Perspective on Barriers, Energy Audits, Energy Management, Policies, and Programs*, volume 978-1-4471-4161-7. Springer-Verlag London, 2013.

[13] Thollander P., et al. International study on energy end-use data among industrial smes and energy-end-use efficiency improvement opportunities. *Submitted for publication in Journal of Cleaner Production*, 2014.