DG ENERGY

Final Report

Maintenance and update of the EU Building Stock Observatory

ENER/C3/2016-547/01

June 2020

Issue and revision record

Revision	Date	Originator	Checker	Approver	Description
1.1	06/04/20	DC	JQ	DC	Initial draft
1.2	07/04/20	JQ	DC	DC	Working draft
1.3	08/04/20	DC	JQ	DC	Working draft
1.4	10/04/20	DC	JQ	AW	Draft final
1.5	11/04/20	JQ	AW	DC	Draft final
1.6	28/05/20	DC	JQ	DC	Final report

RICS
Parliament Square, London,
SW1P 3AD, United Kingdom
t +44 (0)20 7695 1500
f +44 (0)20 7695 1501
e contactbcis@rics.org
w www.rics.org

Contents

Executive Summary

This Final Report outlines the activities of the delivery Consortium over the life of the service contract (ENER/C3/2016-547/01) and summarises progress overall to deliver the tasks associated with the maintenance and update of the EU Building Stock Observatory (BSO). The report details the activities undertaken in order to achieve the project objectives. Predominantly, this included activities associated with data collection, database population and the revamping of the website and associated web tools. The report concludes with options for the future operation of the BSO based on the experience gained from delivering the service contract during the three years from mid-2017 through to mid-2020.

The overall objective of this project was for the continuation, improvement and expansion of the BSO. The contract required the delivery of five specific tasks, as follows:

Task 1: Maintenance and update of the EU BSO website;

Task 2: Data validation, quality control and gap filling;

Task 3: Acquisition of data;

Task 4: Revamping the existing website and database;

Task 5: Feasibility study for launching an EU Buildings Big Data Initiative.

Most tasks were delivered as per expectations however the delivery of Task 4 was subject to significant delays, largely due to programme delays and issues with the QA process, the platform finally went live in early 2020. The remaining tasks were very much focussed on addressing as many of the data gaps in the BSO as possible, based on the availability of data and within the constraints of the existing service contract, and providing a view as to the way forward for any subsequent phases of the BSO. Data acquisition via Task 3 was not taken forward by ENER during this contract. Below we provide a summary of progress made against the key objectives that guided the delivery of this phase of the BSO service contract:

- Ensure the maintenance of existing webpages of the EU BSO, in particular regarding the correct functioning of the three main tools that were developed under the previous contract: **Completed and delivered via Task 4**;
- Update the database and website tools every quarter, by collecting all relevant data and generating the required indicators using the templates developed in the first contract, and / or other templates / methodologies as appropriate: **Completed and delivered via Task 1**;
- Special focus should be given to addressing the data gaps, in particular regarding nondomestic buildings, building renovation, nearly-zero energy buildings, technical building systems, energy certification, energy poverty and financing schemes statistics aggregated at national / regional levels: **Completed subject to data availability via Task 2**;
- Additional data gaps will be created because data sets produced by European projects (e.g. ZEBRA2020, EPISCOPE / TABULA, ENTRANZE, INSPIRE, etc.) will not be updated, therefore it will be necessary to come up with specific measures to address these additional data gaps: **Proposed indicator rationalisation and data acquisition via Task 3 was not taken forward**;
- Upgrade the existing website and tools by enhancing the analytical capabilities of the tools, the navigation of the website and the user-experience. The introduction of new tools should also be considered, in particular tools that allow for further analysis of the data, by linking different datasets: **Completed and delivered via Task 4**;
- Assess the costs, benefits of developing, launching and managing a platform for largescale collection of building characteristics and building performance data to produce statistics and to produce quality statistics and benchmarking tools for decision-makers. **Completed and delivered via Task 5**.

During the life of the project the Consortium sought to harmonise and rationalise the existing data indicators, to ensure confidence in the data contained within the BSO database. The report outlines in some detail the Consortium's actions to address the existing data gaps. While outlining the scale and nature of those gaps,

highlighting existing data sources and the potential of big data to address some of the remaining persistent data gaps going forward.

However, in our opinion access to relevant and reliable data on an ongoing basis is likely to remain a significant issue for the future operation of the BSO. We are aware of the recent Horizon 2020 Framework Programme (H2020) research calls B4E-6 (Big data for buildings) and B4E-7 (European building stock data 4.0) which may help with future data population once delivered. However, there is likely to be a significant lag between now and when any data will become available from those sources.

Our recommendations to improve the data position of the BSO suggest the adoption of a phased approach going forward, more specifically:

Short-term: Until future access to reliable data sources can be guaranteed, on an ongoing basis, we believe the best option for the BSO is for the adoption of statistical/modelled/estimated approaches to supplement data population, cost optimal modelling for instance.

Longer-term: Relevant longitudinal data could become available via wide-scale adoption of smart meters and associated technology across EU Member States, however this may have to be mandatory to be successful.

Whilst describing past activities to deliver the second phase of the BSO this report also provides an assessment of potential future options delivered as part of Task 5. More specifically, the report assesses the cost and benefit of developing, launching and managing an EU Buildings big data platform. In this context, the big data initiative is a large-scale collection of building characteristics and building performance data to produce statistics, quality indicators and benchmarks. The feasibility study sets out a cost benefit appraisal that provides ENER with a view of what is possible from a technology and data stance and assesses a number of options in accordance with the value vs. cost and required resource input for delivery and maintenance.

As the BSO is already in place, using big data technology is the natural next step in its evolution. The current overarching purpose of the existing BSO is to monitor and measure the effectiveness of energy efficiency policy across the EU 28-member states via historical data. One of the significant benefits of the big data initiatives as proposed is the potential it would allow for prediction via forecasting. For the EU Buildings big data initiative, we believe that it will be crucial for building owners / users to be able to predict future energy use of their assets and whether renovation activities would potentially make adequate returns.

In our view, one of the best potential routes to the provision of future quality data, fundamental to the success of this initiative, is to leverage smart meter technology that is currently being installed in domestic and non-domestic properties across the EU. Mandating data collection through Regulations vs. asking suppliers to send meter readings voluntarily to a secured endpoint is considered. We also clearly acknowledge that the future success of any big data initiative will to a large extent be dependent on its capability to ingest data from connected devices and to use that data for its analytical objectives.

The future options considered are flexible, scalable, cost effective and capable of evolving further, as data from a greater number of technologies is incorporated in the coming years and beyond the scope of the development timeframe. Based on our experience and expertise in this area, we believe that the future adoption of a big data initiative that is automated as far as possible would provide the best value for ENER and provide security of data provision going forward. In our opinion, the launch of an EU Buildings big data initiative is certainly feasible from the technology standpoint.

However, the technology platform is only part of the potential solution. The other main component of the initiative is access to relevant data on an ongoing basis and this, we believe, will likely require input from the European Commission on a number of fronts to ensure that the data position significantly improves going forward.

Further, ENER currently takes on the responsibility for the operation and data collection for the BSO, via service contracts with consultants. This centralised function might not be the best route to ensuring adequate data population of the initiative. Based on our experience we believe that the data collection responsibilities could and perhaps should be delegated to individual Member States as they are likely to have a better understanding of relevant national data availability and sources. This approach could be delivered via individual service contracts or mandated through supplements to existing EUROSTAT data collection exercises.

.

1. Introduction

1.1 Background

EU policy makers have for a long time recognised the importance of building performance in the effort to mitigate climate change, starting with the Energy Performance of Buildings Directive (EPBD) and the Energy Efficiency Directive (EED), but capturing the energy saving potential of the existing building stock has posed a significant challenge. While the efficiency of new buildings has steadily improved over time, most of Europe's existing building stock has yet to be affected by energy performance requirements and exhibits low renovation rates.

In order to evaluate the potential energy saving opportunities and improvement in energy efficiency across the EU building stock, to increase the effectiveness of building policies and to be able to evaluate if these policies had the intended result, reliable data on the composition of the building stock was required.

The EU Building Stock Observatory (BSO) is an initiative of the European Commission's Directorate General for Energy (ENER) as part of the *Clean Energy for All Europeans* package in November 2016. It was established following an evaluation of the Energy Performance of Building Directive 2010/31/EU which stated that there was a lack of quality, reliable and consistent data on the actual effect of energy efficiency policies on the building stock across EU Member States (MS).

That evaluation suggested that a better understanding of the effectiveness of policy measures and of market support mechanisms was necessary, to steer an improvement in the depth and rate of buildings' renovation. It was suggested that more transparent information on building stocks would better inform policy makers, supporting the decisions of market players, in particular financial institutions.

Rationale for the BSO

The BSO was developed as an "essential piece" of the EU's building energy efficiency policies. Data provided in the Observatory would contribute to the improvement of the way the building sector is considered in economic modelling of energy efficiency policy options. Further supporting the effect of the EPBD implementation, as well as of relevant articles of the EED i.e. Art. 4-5 and of the Renewables Energy Directive (RED) i.e. Art. 13-14 at national and regional levels and contributing to future policy making and supporting the review process of EU directives.

Therefore, the primary driver for this project was the lack of quality data on a comparable basis on the impact of energy efficiency policies on the building stock across the EU 28 MS and regions. A clearer understanding of the effectiveness of EU policy measures, and of market support mechanisms was deemed necessary to steer an improvement in the depth and rate of building renovations, to enhance energy efficiency across the EU. For these reasons, the European Commission (EC) sought expertise to assist in putting in place the BSO, to monitor and analyse the improvement of energy performance in buildings through the provision of online data and analytics.

BSO development

Launched in November 2016, following development by a Consortium led by the Buildings Performance Institute Europe (BPIE), the BSO had two primary objectives:

- Provide a snapshot of the energy performance of the EU building stock, and provide high-quality data from all MS in a consistent and comparable manner; and
- Set a framework / methodology for the continuous monitoring of the building stock

The 1st phase of the BSO included the:

- development of a list of relevant data indicators;
- a methodology for data collection; and
- a website which contained a database, a DataMapper and factsheets

However, following the 1st phase of the BSO, data gaps and limitations to the data collection, aggregation and comparison across EU MS were identified. Further, while the BSO website was online some features did not work properly. Thus, a second phase of development was required, the intention of the 2nd phase was for the continuation of the BSO but also to:

- address the data gaps;
- provide better data and statistics to its users;
- rationalise the indicators; and
- revamp the website in order to make it more user friendly

Three partner organisations, the: Royal Institution of Chartered Surveyors; Building Services Research and Information Association and Luxoft formed the delivery consortium for Phase 2 with combined expertise in the built environment, energy efficiency and software development.

The key objectives and purpose of the 2nd phase of the service contract was to:

- Ensure the maintenance of existing webpages of the BSO, in particular regarding the correct functioning of the three main tools that were developed under the previous contract;
- Update the database and website tools every quarter, by collecting all relevant data and generating the required indicators using the templates developed in the first contract, and / or other templates / methodologies as appropriate;
- Special focus should be given at addressing the data gaps, in particular regarding non-domestic buildings, building renovation, nearly-zero energy buildings, technical building systems, energy certification, energy poverty and financing schemes statistics aggregated at national / regional levels;
- Additional data gaps will be created because datasets produced by European projects (e.g. ZEBRA2020, EPISCOPE/TABULA, ENTRANZE, INSPIRE, etc.) will not be updated, it will therefore be necessary to come up with specific measures to address these additional data gaps;
- Upgrade the existing website and tools by enhancing the analytical capabilities of the tools, the navigation of the website and the user-experience. The introduction of new tools should also be considered, in particular tools that allow for further analysis of the data, by linking different datasets;
- Assess the costs and benefits of developing, launching and managing a platform for largescale collection of building characteristics and building performance data to produce statistics and to produce quality statistics and benchmarking tools for decision-makers.

1.1 Project Overview

The purpose of this final report is to outline the activities the delivery Consortium have carried out during phase two of the BSO service contract (ENER/C3/2016-547/01), to address the deliverables associated with each task and to provide an audit trail of those deliverables in Annexes to this report. The report draws on the detailed outline of our approach highlighted in the project Inception Report drafted by the Consortium in July 2017 and builds on subsequent progress reports provided by the Consortium in July 2018 and July 2019. The overall objective of this second phase of the BSO was for the continuation, improvement and expansion of the BSO as delivered during the first phase in 2016. The contract required the delivery of five specific tasks:

Task 1: Maintenance and update of the EU BSO website;

Task 2: Data validation, quality control and gap filling;

Task 3: Acquisition of data;

Task 4: Revamping the existing website and database;

Task 5: Feasibility study for launching an EU Buildings Big Data Initiative.

These tasks were interlinked, but discrete work packages composed of further sub-tasks which are shown in Figure 1.1.

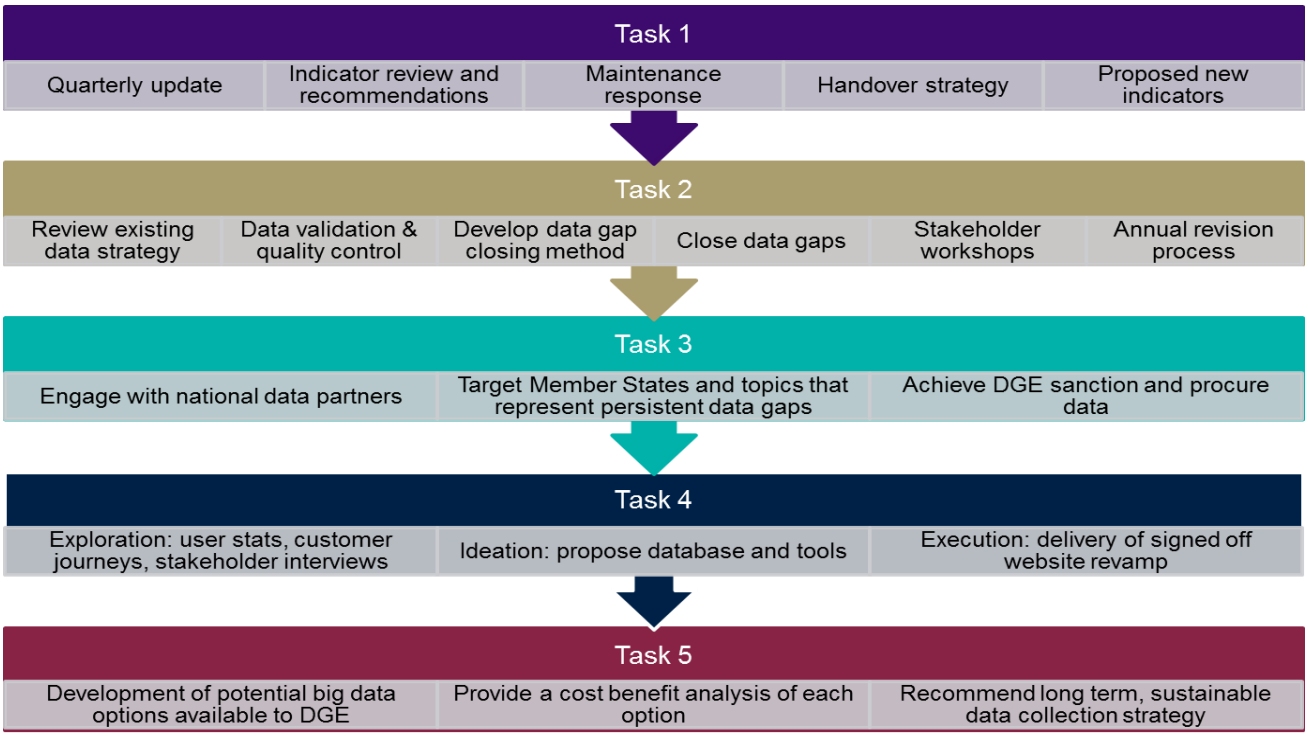

Figure 1.1: Project Tasks and Associated Sub-tasks

1.3 Structure of this Report

This section (Section 1) of the report outlines the projects requirements and describes the objectives to be addressed by the delivery Consortium.

Section 2 provides a high-level programme of the delivery of the five tasks associated with the second phase of the BSO service contract. This current phase was delivered between May 2017 through to June 2020.

Section 3 describes in detail the five tasks that were the focus of this phase of the BSO. Each task is reported individually and presented through the sub-activities which structure each task, incorporating recommended next steps. A large proportion of this report has been dedicated to the methods that the Consortium have explored and implemented to close the existing persistent data gaps, given that this was a significant priority area for both the Commission and the Consortium during the second phase of the BSO.

Section 4 provides an overview of the BSO Stakeholder Workshops, which were held during the contract.

Section 5 describes the various efforts made to promote the BSO via publications (i.e. newsletters) and presentations over the life of the service contract.

Section 6 provides conclusions and recommendations, highlighting potential next steps in delivering future phases of the BSO, including the Consortiums suggestions for ways to take the BSO forward.

2. Project Programme and Deliverables

2.1 Overview

The five tasks that form this service contract were to be delivered between May 2017 through to June 2020. The project programme and deliverables have been scheduled in accordance with the ITT requirements (EU Building Stock Observatory ENER/C3/2016-547/01). This section presents the high-level delivery timetable by task and date.

2.2 High-level Delivery Timetable

Table 2.1 outlines a high-level delivery timetable for the five main tasks that made up the second phase of the BSO service contract, including quarterly data updates, gap filling, data collection strategies, and the revamp of the BSO website. The table also includes stakeholder workshops, meetings, reporting requirements and newsletter releases.

Most tasks were delivered as per expectations however the delivery of Task 4 (website revamp) was subject to significant delays, largely due to a programme delay and issues with the QA process, the platform finally went live in early 2020. Data acquisition via Task 3 was not taken forward by ENER during this contract.

Table 2.1: BSO High-level Delivery Timetable

Activity	Delivery
Task 1: Maintain website (once revamp completed) Including quarterly data updates	March / June / September / December; 2018 / 2019 / 2020
Task 2: Data quality and gaps Including data gap filling Including non-residential built stock survey	 August 2018 – April 2020 May 2018 – Dec 2019
Task 3: Data acquisition	August 2017 – August 2019
Task 4: Website revamp	May 2017 – Jan 2020
Task 5: Feasibility study	July 2017 – November 2018
Workshops	January 2018, January 2019; May 2020
Meetings Kick off 1st progress 2nd progress 3rd progress Final	 11/05/2017 28/09/2017 26/07/2018 24/07/2019 04/05/2020
Reports Draft / final inception Draft / final 1st progress Draft / final 2nd progress Draft / final	 12/06/2017 – 11/07/2017 09/05/2018 – 11/07/2018 13/05/2019 – 11/07/2019 04/04/2020 – 05/06/2020
Newsletter	Drafted for publication following website re-launch

2.3 Schedule of Deliverables

Table 2.2 provides a summary of key deliverables per task. A traffic light system has been adopted to demonstrate clearly where the deliverables have been completed (green), and where deliverables have been postponed or cancelled (red). As can be seen all tasks have been delivered apart from Task 3, where ENER decided not to proceed with the data acquisition element of the project.

Table 2.2: Schedule of Deliverables, by Task

Activity	Deliverable	Delivered
Task 1: Maintain website	Review of existing indicators	08/09/2017 & 10/04/2018
	Rationalisation of existing indicators	08/09/2017 & 10/04/2018
	Gap analysis based on rationalisation of indicators	08/08/2018
	Prioritisation ranking of existing indicators	26/11/2018
	Climate zone definition review	11/10/2018
	Recommendations on new/proposed indicators	June 2020
	Gap analysis based on new/proposed indicators	June 2020
	Quarterly data updates	2017 | 2018 | 2019
	Handover strategy	June 2020
Task 2: Data quality and gaps	Review of existing data collection strategy	19/12/2018
	Gap analysis of indicators	27/10/2017
	Develop a primary data collection strategy	11/04/2018
	Collect, analyse and disseminate results from online non-residential survey	June 2020
	Three Stakeholder Workshops	2018 | 2019 | 2020
	Annual revisions	Complete
	Engage with other EU projects	Complete
	Quantifying the EU NR built stock	June 2019
Task 3: Data acquisition	Acquisition of datasets that populate indicators and/or permit ENER to form sound opinions about the characteristics and energy performance of the built stock	n/a
	Received and evaluated data proposals	10/06/2019
	Upload data to BSO platform once received	n/a
Task 4: Website revamp	Exploration documentation pertaining to investigative activities of existing database and tools	28/09/2017
	Ideation documentation regarding the potential solutions available to ENER	28/09/2017
	Various iterations of the revamped website delivered between Dec 17 and May 18	Dec 2017-May 2018
	Delivery of final revamped website and tools	23/05/2018
Task 5: Feasibility study	Development of potential big data options available to ENER	09/08/2018
	Cost benefit analysis of each option	09/08/2018
	Defined user benefits	09/08/2018
	Technical terms of reference	09/08/2018
	Recommendations	09/08/2018
	Task 5 Report	09/08/2018

3. Progress by Task and Topic

3.1 Overview of Tasks

As outlined previously (Figure 1.1), there are five tasks that make up the BSO service contract. Tasks 1-3 are interlinked and co-dependent. Meaning, the review and rationalisation of the existing indicators in the BSO (Task 1) could only be performed when the data availability, gaps and data collection strategy had been fully critiqued and set out (Task 2). Furthermore, the proposed indicators and data position alongside the defined data collection strategy could then mutually inform the data acquisition plan (Task 3), which required engagement with data partners and the targeting of Topics and MS that represented persistent data gaps.

Task 4 which includes the revamp of the original website focused on the three main tools of the BSO, namely: the database; DataMapper; and the data factsheets. Task 1 directly feeds into Task 4 as the rationalisation of existing indicators highlighted indicators to become inactive within the public domain once the refurbished website had gone live (i.e. due to lack of publicly available data). Moreover, this becomes pertinent in the future phases of the BSO when data becomes more widely available, and additional indicators could be placed within the BSO database and made visible in the public domain.

Task 5 was largely informed from the review and rationalisation of indicators (Task 1) as well as the review of the existing data collection strategy and potential data availability (Task 2). Task 5 looked at automated data collection, artificial intelligence and machine-based learning as a potential approach to close out persistent data gaps. The task required the Consortium to produce a guidance report on the appropriate technology choices to support and enable this type of data capture and generation. The Task 5 output was delivered in August 2018, providing ENER with two defined proposals / options for the future feasibility of launching an EU Buildings big data initiative.

The data structure of the BSO is made up of six broad Topics which encompass the energy efficiency metrics of the EU building stock (Figure 3.1).

Figure 3.1: BSO data topics

The data topics are further presented in the database in the following ten thematic areas: building stock characteristics; building shell performance; technical building systems; nearly zero energy buildings; building renovation; energy consumption; certification; financing; energy poverty; and energy market.

Much of the text that follows, regarding the data position of the BSO, will refer to the specific data topics.

3.2 Task 1: Maintenance and Update of the BSO Website

Task 1 was an ongoing activity that continued to be delivered through to June 2020. This included various sub-tasks in addition to quarterly updates of both database and the individual topic csv files for upload to the website.

There are over 170 main data indicators (made up of over 70,000 individual data points), spanning the data topics. Given that the BSO purported to host a large number of data indicators, combined with the fact that the database had a substantial proportion of missing data it was necessary to perform a review and rationalisation of the indicators in order to understand the large number of persistent data gaps. This activity was originally identified in the first progress report as being a high priority to complete within the first year of the contract. There were a number of reasons why it was suggested that a sizeable share of the database be removed from the BSO, including duplication of indicators, some indicators are simply not measured, publicly available or known.

However, the rationalisation of indicators remained ongoing and ultimately ENER made the decision not to proceed with the rationalisation as proposed in Autumn 2019. This activity was closely linked with the data gap analysis, data gap closing strategy and the review of existing data collection outlined in Task 2 and additionally the proposition of any new indicators (also within Task 1). We recommend that the data rationalisation is taken forward during the next phase of the BSO.

Table 3.2: Task 1 Progress

Activity	Status
Quarterly data updates	Updates of open source data occurred in March, June, September, and December 2017-2020. This activity focused on open source databases (Eurostat, National Statistical Offices, Odyssee-Mure).
Rationalisation of existing indicators	Completed in July 2018 but has remained ongoing (Annex 1). ENER took the decision not to proceed with rationalisation in Autumn 2019.
Gap analysis based on rationalisation of indicators	Completed to inform above.
Prioritisation of existing indicators	Completed to inform the Task 3 data acquisition proposal (Annex 2).
Climate zone definition review	Revised and completed.
Maintenance response	To be handled in-house by DIGIT.
Handover strategy	Technical handover document prepared (Annex 5).
Propose new indicators	Not considered appropriate given the remaining data gaps. Should only be considered based on the availability of data.

3.2.1 Task 1: Activities to Achieve Deliverables

Over the contract lifetime the Consortium performed several key activities associated with the delivery of Task 1 and these included:

- The quarterly updates were undertaken in the months of March (2018 - 2019 - 2020), June (2018 - 2019), September (2017 - 2018 - 2019) and December (2017 - 2018 - 2019). It has become apparent over the past 12-months that the largest and most significant data release from open sources (i.e. Eurostat) occurs between September and November each year, making Q4 updates

the most fruitful. During this process the established data validation and quality assurance processes are performed (Task 2);

- Rationalisation of existing indicators within; Topic 1: Building Stock Characteristics; Topic 3: Certification; Topic 4: Finance; Topic 5: Fuel Poverty and Social Aspects. The final proposal was submitted by the Consortium in June 2018,
- Rationalisation of existing indicators within Topic 2: Technical Systems, this activity was undertaken by BSRIA and finalised with ENER in July 2018;
- A gap analysis of the BSO database was completed with the proposed rationalisation of indicators taken into account. This sub-task was to show the extent of the data gap closure once the rationalisation had been realised and was completed in August 2018;
- A ranking system to prioritise the existing indicators within the BSO was undertaken, to highlight which indicators should be the focus of data population routines. This activity was carried out in November 2018;
- A review of the existing climate zone definitions was produced, this was to inform the primary data collection strategies that the Consortium had implemented. This activity was finalised in August 2018;
- An initial discussion of Topic 6 occurred in September of 2018 with ENER but the development of this Topic remains unresolved.

3.2.2 Quarterly Data Updates

One of the main activities associated with Task 1 was the quarterly data updates from open source databases, this research occurred during March, June, September and December of each year. It must be noted that the majority of relevant datasets for the BSO were typically released during the latter stages of each calendar year, therefore Q1-3 data updates weren't usually as valuable as the Q4 update. It should also be noted that these exercises weren't directly related to data gap closure exercises (Task 2), as the open source databases maintained and updated the existing populated indicators within the BSO. Although, if data was discovered that could populate a previously unpopulated indicator, it was utilised within the BSO database and contributed towards data gap closure.

The entire BSO contains 74,008 individual data points and the open source databases which are regularly exploited are those of Eurostat, Odyssee-Mure and National Statistical Offices. Through this exercise Eurostat populated 4,657 complete data points (6% of the entire BSO), Odyssee-Mure populated a further 1,534 data points (2% of the BSO), and associated National Statistical Offices filled 60 individual data points, which equated to just less than 1% of the database. Table 5 demonstrates the level of contribution that each open source database made to the BSO by the main topic areas (excluding Topic 6).

Table 3.3: Open Source Database Contributions to the BSO

	BSO	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Eurostat	**6%**	9%	2%	1%	<1%	62%
National Statistics	**<1%**	<1%	0%	<1%	0%	0%
Odyssee-Mure	**2%**	6%	2%	0%	0%	0%

3.2.3 Rationalisation of the Existing Indicators

The rationalisation process associated with the data topics predominantly focused on the availability of quality data. In the final report submitted by the previous Consortium, the indicators not available in the public domain were highlighted. We reviewed the previous Consortium's assessment. The Consortium engaged with academic and research institutions and with other actors, such as the CA EPBD, to fully understand the issues related with the availability of data. Specifically, to highlight if there was any alignment between the BSO and the Key Implementation Decisions (KIDs) database, however while there is some overlap it is limited at this stage.

The proposed rationalisation of the existing indicators that make up the BSO database was a priority sub-task of Task 1. It followed directly from another associated sub-task which was the annual revision of the indicators, which was an ongoing activity. The rationalisation process focused on indicators associated with Topics 1-5. In

the final report submitted by the previous Consortium the indicators not available in the public domain were highlighted and the present Consortium received this assessment. However, due to some of the activities associated with this phase of the service contract (i.e. primary data collection strategies) some of the previous Consortiums recommendations for the rationalisation of indicators have subsequently been superseded.

The main focus of the rationalisation process was to suggest certain indicators to become inactive within the public domain on the live website. This in turn will reduce the amount of visible persistent data gaps, resulting in a more consistent database with higher user satisfaction and confidence with the data available within the database. The rationalisation process predominantly focused on the inherent data gaps in the database, and the quality, availability and frequency of open source data was also considered. There were several characteristics that the Consortium deemed appropriate for the recommendation to hide certain indicators from the public domain, namely; the data is not presently known, measured, or publicly available; or the data is highly unlikely to be measured in the near future (see Annex 1).

As mentioned previously, there are 74,008 individual data points in the BSO, and as a result of the rationalisation of indicators activity, the Consortium have suggested that 51,244 data points (approximately 69%) of the BSO be deactivated and removed from the public domain. This translated to just over 1,700 sub-indicators that currently exist within the BSO to be hidden, resulting in a substantial increase in visible data population across the entire database. This activity would contribute to significantly reducing the proportion of persistent data gaps from 87% to approximately 64%. Consequently, the amount of populated data would increase to 36% and 23% of the existing persistent data gaps would be closed, as summarised in Table 3.4 below.

Table 3.4: Impact of the rationalisation as proposed

	BSO	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Original data population	**13%**	20%	26%	4%	<1%	62%
Indicators to be hidden	**69%**	61%	19%	74%	93%	37%
Proposed data population	**36%**	42%	31%	4%	6%	97%

It was originally suggested in the First Progress Report, that the sub-task of rationalising the indicators was a priority which should be concluded within the first year of this phase of the project (i.e. December 2018). The original rationalisation proposal was submitted to ENER in July 2018, nevertheless this activity remains unresolved, although in our opinion the rationalisation should be taken forward in the next phase of the BSO.

3.2.4 Prioritisation of the Existing Indicators

The prioritisation of existing indicators was a sub-task falling under Task 1. This activity involved categorising each individual indicator into priority gradings, based on the priority of populating each indicator. Table 3.5 below establishes the prioritisation rankings set out by the Consortium.

Based on the prioritisation below there was a fairly even split between all of the gradings with the number of indicators present within each category, which facilitated the process moving forward. The purpose of this activity was to better inform discussions with potential data partners, to display clearly where the Consortiums data population priorities lay.

As anticipated the majority of indicators which make up the top 2 prioritisation grades (1) are from Topic 1 (Building Stock Characteristics) which has been identified as a priority Topic for data population since the beginning of this service contract, with a primary focus on the non-residential sector. Some of the examples included the non-residential built stock by building type, non-residential building floor area, and EPC indicators (Topic 3). This activity was due to directly inform Task 3, where the Consortium specifically identified indicators within topics that could make use of the data acquisition budget available for this project (see Annex 2).

Table 3.5: BSO Indicator Prioritisation Levels

Rank	Prioritisation Grade	Description
1	Very high priority	Indicators that needed to be populated first and the focus of the primary data collection strategies and acquisition budget (Task 3)
2	High priority	Indicators that needed to be populated, but as a second priority to those graded "1"
3	Medium priority	Indicators that should be populated if the data becomes measured or known
4	Low priority	Indicators that should be a focus at a later date during a later phase of the BSO
5	Lowest priority	Indicators that have been recommended to be hidden from the public domain
0	Not a priority	Indicators that were regularly and wholly populated with high quality, reliable data

3.2.5 Climate Zone Definition Review

Throughout the rationalisation process and the recurrent review of the BSO database, it became apparent that there weren't clear definitions of the climate zones previously held within the BSO. Thus, a review of the climate zones was required, the Consortium distinguished seven climatic zones with transparent definitions which align with the Inspire project, Table 3.6 below summarises the definitions.

Table 3.6: BSO Climatic Zones

Climate Zone	Definition
Southern Dry	Warm to hot dry summers, with winters that tend to be cold. Higher altitudes than other regions of southern dry/semi-arid, and larger daily ranges in temperatures.
Mediterranean	Characterised by hot dry summers, and cool wet winters. Annual temperature ranges are generally smaller than those found in marine/oceanic/temperate regions.
South-continental	Warm to hot (and often humid) summers and cold (more severe the further north) winters. With even distribution of precipitation across the annum.
Continental	Generally characterised by strong contrasts between the cold winters and hot summers. The continental nature of the climate becomes more pronounced as you move West to East, with the East suffering more defined changes, whereas to the West it is also influenced by oceanic climate influences.
Oceanic	Predominant in most western parts of Europe. These regions experience cool, but not severely cold winters and warm summers, but cooler summers than continental climates. Annual temperature range of these region is relatively narrow. Precipitation is also evenly distributed across the year.
North-continental	Climate zone with a mild summer and a cold winter not so extreme as the Nordic climate. It includes Countries at the South of the Baltic Sea (Denmark, Poland, and Lithuania).
Nordic	Coldest winters of the EU countries, also the least amount of daylight hours during winter months. The parts of Estonia and Latvia which are coastal will suffer less extremes in climate, compared to the landlocked areas of the countries.

3.2.6 Task 1 Recommendations for Future Phases

1. The consortium's view is that finalising the rationalisation of indicators, which would then be reflected within the live BSO website, should be a priority for any subsequent phase of the BSO.

2. In addition, defining Topic 6 on building codes and national definitions would be useful to be incorporated into the BSO. Population of the topic should be based on the availability of high-quality reliable data and potential collaboration with other EU projects.
3. A further sub-task under Task 1 included the quarterly data updates from open source data (i.e. Eurostat, Odyssee-Mure, and associated National Statistical Offices) which contributed towards maintaining and updating the BSO website. This data was fed manually directly into the BSO database, but in future phases it should be explored how this process could become automated.
4. Finally, gap analyses should follow each quarterly update as an associated activity within Task 2 (see next section).

3.3 Task 2: Data Validation, Quality Control and Gap Filling

Task 2 was delivered on an ongoing basis through to the end of the service contract. In addition to the Task 2 activities outlined below, data gap analysis and quantifying the building stock have been included as more prominent activities for this Task. Table 3.7 identifies the sub-tasks associated with Task 2 and the status of each activity.

Furthermore, the same data validation and quality control processes have remained unchanged throughout this phase of the service contract.

Table 3.7: Task 2 Progress

Activity	Status
Review existing data strategy	The existing data strategy was continually reviewed, and the calculation methodologies were assessed during the early stages of this service contract. Open source data collection methods remain unchanged, and the engagement with other EU projects to identify potential data sources and gap closures should still be pursued during future phases
Data validation and quality control	Developed QA / validation process and approach which is reflected in the database and live website
Develop data gap closing method	Developed 5-stage data collection strategy to be implemented throughout contract lifetime
Close data gaps	Remained ongoing activity until contract end
Data gap analysis	A gap analysis on the original database was carried out in late 2017, and a revised gap analysis was carried out after each quarterly data update
Quantifying the non-residential built stock	Completed, the Consortium were provided non-residential building stock estimates from Hotmaps project (EASME)
Stakeholder workshops	Three workshops successfully delivered
Annual revision process	Activities above repeated annually

3.3.1 Task 2 Activities to Achieve Deliverables

The Consortium performed several key activities in addition to those outlined above which were associated with the delivery of Task 2 and included:

- Quarterly data updates occurred regularly and during this process the established data validation and quality assurance processes were applied (March/June/September/December)
- A detailed analysis on the data gaps present within the BSO database which was originally submitted in October 2017 with further analyses completed after each data update
    - Analysis by topic
    - Analysis by building sector (residential and non-residential)
    - Analysis by MS
- Critique of the existing data collection strategy and availability of previously utilised data sources (submitted December 2017)
- Designed a 5-stage data population plan to close as many persistent data gaps as possible
- During the earlier stages of the service contract two primary data collection strategies were designed, the online self-completion non-residential building stock survey and a series of focused interviews carried out by BSRIA. These approaches were launched in May 2018, with the series of interviews reaching completion in February 2019 whilst the online survey remained open until early 2020

- The Consortium were provided with EU building stock estimates (residential and non-residential), including floor areas from the EU project Hotmaps. These data were utilised to populate certain indicators within the BSO, along with data obtained from the NZEB study
- Progress made during the service contract were presented at stakeholder workshop events on an annual basis.

3.3.2 Data Validation and Quality Control

Through the quarterly data updates (Task 1) that occurred in March, June, September and December 2017 - 2020, the data validation and quality assurance processes were maintained, which were an important element of Task 2. The previous Consortium outlined and implemented a data validation / quality assurance scoring system, which was applied to each data source and is displayed on the BSO website next to the associated data. In the interest of consistency, the current Consortium maintained the same scoring approach (Table 3.8).

Table 3.8: Data Quality Scoring System

Quality score	Description	Comments
5	Official statistics	Official European (i.e. Eurostat) or national statistical office data (e.g. DESTATIS) were considered more reliable and a higher quality of data above research from project results.
4	EU projects	Statistics developed in various research projects, such as; Tabula / Inspire / Entranze and countless others. Another chief example was data collected from Odyssee-Mure, which is already a secondary data source.
3	Literature / research projects	Literature, or research projects aimed at specific topics. These reports often contained data based on multiple sources and/or own analyses, in output reports the quality of this data is already assessed and tailored for the relevant context. BSRIA WMI data, some of which was used to populate Topic 2 indicators, also falls under this category.
2	Model output / calculations	Results of modelling based on statistical data and expert assumptions. Data from both primary data collection strategies falls under this category.
1	Assumptions / estimations	Expert assumption; no analytical modelling involved.

3.3.3 Data Position

As mentioned previously, multiple gap analyses were carried out over the course of this phase of the service contract. The entire BSO consisted of 74,008 individual data points, and during the first data gap analyses (submitted in October 2017) it was revealed that only 9,668 (13%) individual data points within the entire EU BSO database were populated with data. Figure 3.9 (below) demonstrates the current data position. The final quarterly data update of 2020 was carried out in March and subsequently a data gap analysis was performed, the number of populated data points has increased to 13,893 (19% of BSO) and the data gap proportion has decreased accordingly to 60,115 (81%).

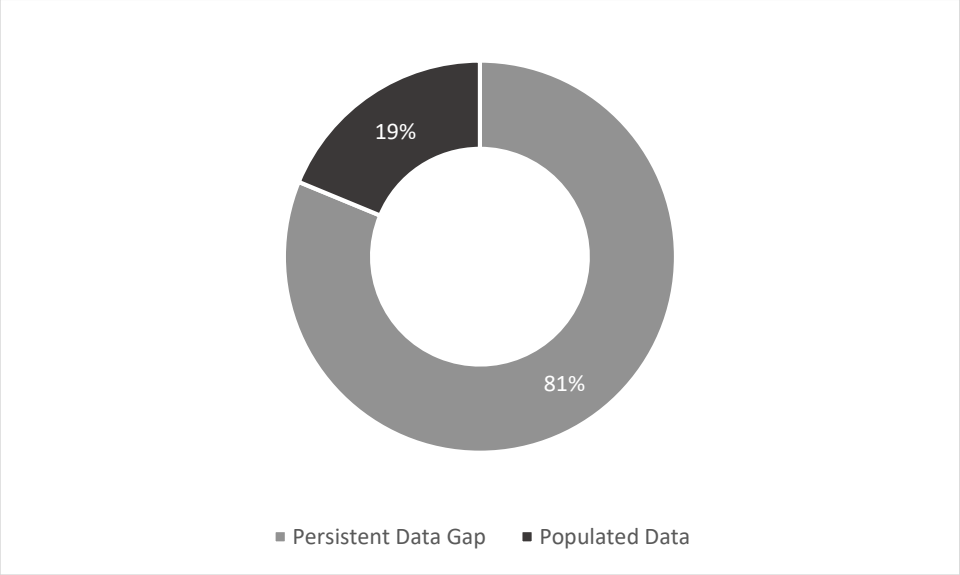

Figure 3.9: Current Data Position of the BSO

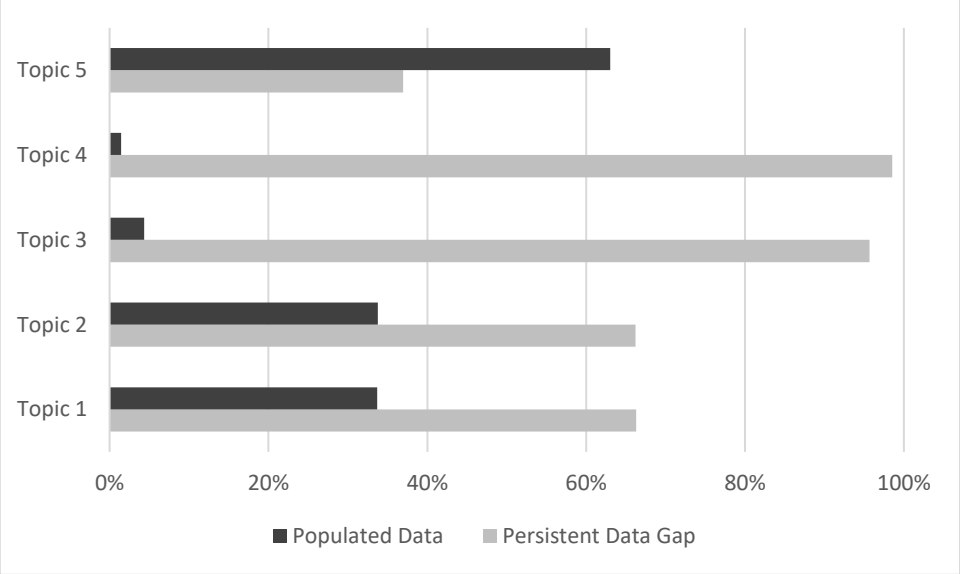

Figure 3.10: Current Data Position of Topics 1 – 5

As displayed above in Figure 3.10, the most noteworthy data gaps remain in Topic 3 (Certification) and Topic 4 (Finance). Topic 3 consists of 10,324 individual data points with only 449 containing data, resulting in 96% of missing data. Topic 4 includes 27,985 data points, only 410 of which were populated with data equating to 99% of this Topic being devoid of data. Conversely, Topic 5 (Fuel Poverty and Social Aspects) was relatively well populated with data when compared to the remaining topics. Topic 5 consisted of 3,393 individual data points with 2,139 being populated through Eurostat, which resulted in 63% of data population for Topic 5. Topic 1 which was recognised as a priority topic held 23,461 individual data points, 7,907 (34%) of which were populated with data. The second priority Topic Technical Systems (Topic 2) had 8,845 data points where 2,988 data points contained data, which represented 34% of the Topic.

In terms of **data population by building sector**, the residential building stock section of the BSO had 34,800 individual data points, with 8,880 (26%) of these populated with data, whilst the non-residential building stock was made up of 33,756 data points with only 4,600 containing data (14% data population). The remaining 9,483 data points belonged to an “All buildings” category, where only 50 individual data points have been filled to date. The comparatively low data population rate within the non-residential built sector is one of the major reasons

why the Consortium's primary data collection strategies were geared towards gathering data on the non-residential building stock.

The BSO also provides **data at the EU 28 MS level** where available, individual MS each have 2,356 data points, and each distinct MS has 19% - 24% of populated data points. Individually, France displayed the highest number of populated data points out of all the MS, with 557 (24%), conversely the Netherlands had the least amount of populated data points with only 453 (19%) individual points. Table 3.11 below highlights the current data position per MS.

Table 3.11: Data Position by Member State

MS	Populated Data Points	%	Persistent Data Gaps	%
EU 28	484	21%	1,872	79%
Austria	556	24%	1,800	76%
Belgium	506	21%	1,850	79%
Bulgaria	557	24%	1,799	76%
Croatia	494	21%	1,862	79%
Cyprus	494	21%	1,862	79%
Czech Republic	557	24%	1,799	76%
Denmark	522	22%	1,834	78%
Estonia	478	20%	1,878	80%
Finland	496	21%	1,860	79%
France	567	24%	1,789	76%
Germany	555	24%	1,801	76%
Greece	501	21%	1,855	79%
Hungary	500	21%	1,856	79%
Ireland	491	21%	1,865	79%
Italy	504	21%	1,852	79%
Latvia	482	20%	1,874	80%
Lithuania	481	20%	1,875	80%
Luxembourg	479	20%	1,877	80%
Malta	486	21%	1,870	79%
Netherlands	453	19%	1,903	81%
Poland	500	21%	1,856	79%
Portugal	492	21%	1,864	79%
Romania	507	22%	1,849	78%
Slovakia	488	21%	1,868	79%
Slovenia	476	20%	1,880	80%
Spain	525	22%	1,831	78%
Sweden	521	22%	1,835	78%
UK	545	23%	1,811	77%

Further analysis was carried out to determine the **breakdown of sources** which populated the 19% of the data within the BSO database (13,893 data points). Of the proportion of populated data 6,253 data points (approximately 8.5%) was sourced through high quality and reliable databases (i.e. Eurostat). The remaining 6,059 data points (approximately 8%) was obtained from various EU projects, which have all now reached completion without any confirmed replacements. The implication with the EU project data is that the data won't be updated in the near future, thus the existing populated data points will eventually become outdated. Table

3.12 reveals the BSO data population by source and then by Topic. During the first phase of the BSO the previous Consortium populated numerous data points through calculations and estimations, however the methodologies were not provided in the handover documentation. Consequently, these calculations cannot be duplicated, and the data points therefore cannot be populated using the previous approach.

Table 3.12 BSO Database by Data Source

	BSO	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Persistent Data Gap	**81%**	66%	66%	96%	98.5%	35%
Eurostat	**6.32%**	9%	2%	1%	< 1%	65%
National Statistics	**< 1%**	< 1%	0%	< 1%	0%	0%
Odyssee-Mure	**2.32%**	7%	2%	0%	0%	0%
EU Project	**8%**	17%	29%	3%	1%	0%
Calculation	**< 1%**	< 1%	0%	0%	0%	0%

3.3.4 Data Population: by Topic

The BSO database holds over 170 energy efficiency indicators, 2,500 sub-indicators, and approximately 74,000 data points. Which cover key energy efficiency indicators across all EU 28 countries (including an EU overview), with 2,500 data points per MS.

The database is structured with five broad data topics; T1 Building Stock Characteristics; T2 Technical Systems; T3 Certification; T4 Finance; T5 Fuel Poverty and Social Aspects (see Annex 3 for detailed Topic factsheets).

In order to address the significant proportion of persistent data gaps within the BSO the Consortium devised a data population plan to close as many of the data gaps as possible within the constraints of this service contract. The approach represented a substantial segment of Task 2, as well as Task 3, and remained open until the end of the service contract. Part of the data population plan consisted of implementing the rationalisation of indicators to deactivate certain ones from the public domain on the website. Whilst this isn't a gap closing exercise it allowed the users of the BSO to be more confident in data availability in the database and presented with less indicators that were reported as missing. The main bulk of the data population plan was to engage Task 3 and endeavour to collaborate with other EU projects to identify synergies between databases to close data gaps.

The Consortiums data population plan also consisted of two primary data collection strategies, an online non-residential building stock survey and a series of focused interviews (carried out by BSRIA). As previously stated the gap analysis highlighted the persistent data gaps between the two building sectors, and the rationalisation of existing indicators emphasised the need to prioritise Topics 1 and 2. As a result both the primary data collection tools were designed to address persistent data gaps from the non-residential building sector in Topic 1 (Building Stock Characteristics) and Topic 2 (Technical Systems).

3.3.4.1 Topic 1: Building Stock Characteristics

Topic 1 was regarded as a priority because the indicators gave an overview of the scale and make-up of the building stock, knowing how a building was composed allowed for further calculations to estimate possible energy efficiency improvements. The topic contained useful energy efficiency indicators, which could contribute to demonstrating energy efficiency trends across the EU building stock. Which in turn allowed specific building types, of a certain construction period to be highlighted as the best pathway to increasing the energy efficiency of the EU's buildings. Topic 1 is made up of 23,461 individual data points, which span three sub-topics. 7,907 of these data points were populated with data (34%):

Topic 1.1, Building Stock characteristics illustrated the EU building stock to building type (residential and non-residential), disaggregated to construction period, location and climatic zone, floor area, annual construction and demolition rates, and energy renovation rates. Additional indicators highlighted post-renovation energy efficiency gains.

Topic 1.2, Energy Consumption, included indicators which covered a buildings energy consumption, which were broken down into energy use (i.e. total use, electricity, space heating, space cooling, water heating, ventilation, and appliances), disaggregated to building type.

Topic 1.3, Fuel Supply Mix, highlighted energy use, by energy carrier (i.e. Total electricity production, of which solid fuels, etc.).

High quality, reliable data (i.e. Eurostat, and national statistics) provided the highest proportion of this topic's data, with a share of over 9%. Odyssee-Mure data also covered nearly 6% of the populated data points, leaving EU projects and one calculation to fill nearly 5% of this topic.

In mid-2019 the Consortium were provided with data from the EASME project Hotmaps that had recently completed. The data covered many of the building stock characteristic indicators for Topic 1 population. Furthermore, data from an additional completed service contract was made available to the Consortium in the Autumn of 2019, which focused on Nearly Zero Energy Buildings (NZEB) and energy renovation across the EU building sector, undertaken by a consortium led by Ipsos (previously Navigant). Both projects contributed towards data population of the EU BSO with the Hotmaps data being particularly valuable given its focus on the scale of the building stock across EU Member States.

Observations/recommendations

The most practical and cost-effective route to maximise data gap filling for Topic 1 is to utilise any new EU project data concerned with the building stock, therefore all relevant future service contracts commissioned by ENER should consider the architecture of the database and ensure that any data collected aligns closely with the indicators within the BSO.

Furthermore, cross-sectional stock data (Hotmaps) already present in the database could be used to provide a longitudinal dataset if used as a baseline and subsequently updated via calculations/modelling using stock replacement rates to update.

3.3.4.2 Topic 2: Technical Systems

Topic 2 indicators covered metrics that could provide a high-level view of key energy efficiency indicators, which are focused on technical systems, and the buildings fabric. Namely; U-values, including windows, roofs, doors etc.; degree of window glazing; on-site renewable energy generation; heating systems disaggregated to boiler type and heat pumps; and the technical systems behind a buildings space cooling system, water heating systems, ventilation, lighting, cooking, and appliances. The indicators within this topic are important, as they provide an understanding of the building fabric, and the technical systems within a building which allows for analysis to be made on specific aspects of technical energy efficiency trends across the EU building stock.

Nevertheless, only 34% of Topic 2 indicators are populated with data. However, unlike Topic 1, the main contributor of data to this topic were EU projects, many of which have now reached completion, accounting for approximately 25% of the data within this topic area. Higher quality, and more reliable data had been provided by Eurostat, national statistics, and Odyssee-Mure databases which accounted for a combined 4% of data points within Topic 2. Consequently 66% of Topic 2 indicators remain without data.

Observations/recommendations

Topic 2: Technical Systems for non-residential buildings had an exhaustive set of indicators in the BSO where the key challenge at the outset had been to identify the largest data gaps and develop a methodology to efficiently gather as much reliable data as possible to reduce the number of data gaps.

BSRIA adopted a two-stage approach of collating their historic data and gathering primary data from interviews after rationalising the indicators to populate a significant number of data points. Going forward, BSRIA made the following key recommendations for the future continuation of the BSO:

- The indicator list and indicator definitions need to be segregated for residential and non-residential buildings as they fundamentally differ in their operational needs.
- Average efficiency of a system in a non-residential building is very difficult to capture as the asset

managers or facility managers operate the building to fulfil their business requirements and not to achieve a certain operational efficiency. BSRIA recommends that the indicator should either be amended to capture the system's capacity or removed from the public domain.

- Combi Boilers are often not used in non-residential buildings, as they have limitations that make them unsuitable for this type of buildings, hence this indicator should be removed from the public domain in the non-residential data sheet.

Energy consumption for water heating, space cooling, appliances, and ventilation, are indicators present in Topic 2 (Technical Systems) and are examples of some of the indicators within the BSO that are seen as beneficial in providing an overview of energy efficiency gains in buildings. Nevertheless, these indicators have been labelled as persistent data gaps, as the data isn't widely collected or known, it has been suggested that the data required to populate these indicators won't be readily available until smart meters are widely adopted across the EU.

Topic 2 is the most complex data topic to populate given the inherent technicalities involved, we suggest that ENER will need to explore partnerships with technical experts to populate going forward, although it should be noted that the data is often proprietary and therefore difficult to access.

3.3.4.3 Topic 3: Certification

Topic 3 focused primarily on data concerning Energy Performance Certificates (EPCs) of buildings. Indicators cover total number of EPCs, by year, floor area, size of unit, and the impact of EPCs on the value of properties. This topic also includes indicators which reveal rates of alternative energy certificates, such as; Passivehaus, Minergie, LEED, BREEAM, DGNB, and HQE, which provides a more complete view of energy related certification schemes across the EU. Theoretically, EPCs are very useful and provide a snapshot of how a building performs, although these certificates don't reveal the reason behind energy efficiency / performance gains, they do allow for trends in energy efficiency improvements to be measured for individual buildings.

Topic 3 contained 10,324 individual data points, 449 (4.3%) of which were populated with data. When this phase of the EU BSO service contract commenced, Topic 3 held 440 data points meaning, nine data points had been added to this Topic.

However, Topic 3 had historically been one of the least populated with data despite the data existing, the main issue is agreeing access to the EPC data at an MS level. The majority of the populated data points have been provided from completed EU Projects and from sources produced from the alternative certification schemes mentioned above. Eurostat data, national statistics, and the Odyssee-Mure database account for less than 1% of Topic 3 with high quality, reliable data.

Observations/recommendations

Despite numerous attempts to access EPC databases via the EPBD the Consortium have been unsuccessful in their attempts to secure the relevant datasets. Currently, there isn't a uniform system across the EU that obliges MS to record and update EPC databases. Because of this, EPC data isn't publicly available across a number of EU countries, which results in this topic having significant data gaps of over 95%. This is an area of concern, as unlike some other topics the data for Topic 3 exists but is not currently available pan EU. Therefore, we recommend that national EPC databases should be structured/formalised across MS to allow for future use in the BSO. In addition, future public access to EPC data should be mandated by the EC to make it available more widely.

3.3.4.4 Topic 4: Finance

Topic 4 indicators aim to illustrate trends in energy renovation investments, including patterns in government led financial incentives. Topic 4 is regarded as an important topic within the BSO, as it demonstrates how EU and individual MS financial schemes impact the rate of energy renovations across the EU. This is significant, as it is known that government driven financial incentives hugely influence the rate of investment into energy efficiency improvements. Thus, data within this topic had the ability to reveal which financial schemes had the

biggest impact in encouraging the uptake of energy efficiency / performance improvements, which in turn provided an idea of which schemes should continue and could be developed in the future.

Topic 4 is made up of 27,985 individual data points, of which 409 (1%) data points contained data. Through the March 2020 data update, 112 individual data points were updated taking the total population to 522.

Similarly, to Topic 3, Topic 4 had major data gaps, with only 0.6% of the entire topic populated with data. Eurostat data regularly populated 0.4% of this topic, with EU projects providing 0.2% of the topic with sporadic / temporal data entries.

Observations/recommendations

There are a number of relevant studies exploring the different financial mechanisms in individual MS and during the contract effort was made unsuccessfully to gain access to the data available from these sources. Going forward there is a need to explore synergies with relevant data sources to access any relevant data. More specifically, exploration of partnerships with 'green banks' is an option for populating the indicators within this topic.

3.3.4.5 Topic 5: Fuel Poverty and Social Aspects

Fuel poverty and social aspect indicators make up Topic 5 and were concerned with providing a detailed view of the key reasons why some of the population was at risk of, or in fuel poverty. For instance; the proportion of the population that can't keep comfortably warm in winter, and cool in summer and total expenditure on household energy bills, were examples of indicators from Topic 5. The topic further provided an overview of the energy market, with views on switching rates, energy prices per kWh, etc. The importance of this topic was that it demonstrated what percentage of the European population fell into the fuel poor bracket, which in turn allowed target groups to be identified for further assistance.

Topic 5 was the topic with the highest data population, approximately 64% of indicators had data largely sourced from Eurostat. This topic had the most populated data points, and also the highest representation of high quality, and reliable data.

Topic 5 was the smallest of all the BSO Topic's with only 3,393 individual data points, this topic has also historically held the largest proportion of data with 2,183 (64%) populated data points. During the last data update of 2020, 35 (0.1%) data points were updated with data from Eurostat. The remaining data that had previously been sourced through Eurostat were not updated, as no updates to the datasets have yet been made publicly available.

Observations/recommendations

Given that this Topic is solely populated through Eurostat data establishing an API link to automatically upload the data would make sense going forward. Moreover, the Energy Poverty Observatory (EPOV) almost mimics this Topic and there should be opportunities to share data across initiatives.

3.3.4.6 Topic 6: Building Codes and National Definitions

Topic 6 was not well defined during the first phase of the BSO, although potential indicators were outlined. Initial discussions during phase two for the finalisation of Topic 6 began in September 2018 although they remain unresolved.

3.3.4.7 Data population summary

As previously mentioned, the BSO phase 1 was inherited by the Consortium with over 170 main indicators, 2,500 sub-indicators, and approximately 74,000 data points; which covered the 28 EU MS, and the EU as a whole. However, only 13% of the BSO database contained any data, of which 9% of these populated data points were sporadic, and temporal across MS and yearly data entries. Ultimately, this meant that only 4% of the BSO, was populated and regularly updated from high quality and reliable data sources (i.e. Eurostat data, national

statistics, and Odyssee-Mure). The current data position has improved with 19% of the indicators now populated with data, however data collection and population are likely to remain an ongoing issue for the BSO.

Data collection and by extension population of the BSO proved to be a significant challenge during delivery of the contract. Extensive efforts were made by the Consortium to address as many of the data gaps as possible. However, many of the persistent data gaps within the BSO don't currently have data that is either widely measured, known or available which hampered our ability to populate the database and close all the data gaps.

Approach to addressing the data gaps

Improved data population was a requirement and to address the data gaps the Consortium developed a 5-stage strategy, agreed with ENER, to:

1. Rationalise the original set of indicators within the BSO;
2. Prioritise the collection of non-residential building data where possible (via surveys);
3. Collaborate with other "live" EU funded projects that were capturing useful and relevant data (i.e. the renovation rate / NZEB uptake service contract);
4. Collaborate with other stakeholders to gain access to relevant data (i.e. CA EPBD and MS for EPC databases; JRC; EASME; etc);
5. Utilise the project specific data acquisition budget (Task 3)

The Consortium estimated that the five approaches combined would have led to a 39% data population position, an increase of 26% over the original Phase 1 data population. However, two of the five activities (the proposed indicator rationalisation and utilisation of the data acquisition budget) were not completed during this phase of the BSO.

Proposed rationalisation

In order to address some of the persistent data gaps the consortium devised an indicator rationalisation plan. The rationalisation identified indicators with certain characteristics which we proposed were removed from the public domain, those indicators could be "hidden" until data became available i.e. from smart metering etc

Some of the characteristics included:

- Duplication of indicators (e.g. Number of Dwellings)
- Data not available to the public (e.g. EPCs per year)
- Data not known or measured (e.g. EPCs per size of building)

The proposed rationalisation recommended that 69% of the existing indicators were hidden from the public domain, removal of the indicators as proposed improved the data population position. However, ENER decided not to proceed with our recommendations during this contract.

Task 3: Data Acquisition

There was a separate data acquisition budget made available for this project to source additional data. However, ENER decided not to proceed with this element of the contract during the second phase (see Section 3.4 for further detail).

Surveys

In order to address some of the existing data gaps in Topic 1 the Consortium designed an on-line non-residential building stock survey. The survey was launched on six separate occasions, the first release in early 2018 was aimed at a convenience sample of over 2,000 respondents who were familiar with the facilities management of buildings across Europe. Following the second release of the survey, which only gathered 49 datasets, it was deemed necessary to adapt the survey and focus on a wider population. However, the introduction of GDPR legislation in May 2018 necessitated a different approach to data collection. Thus, the second version of the survey utilised an RICS opted in respondents list of European members who were qualified building professionals, with a sample size of approximately 7,000 respondents. The survey remained open until late 2019 and during that time 261 completed datasets were returned.

In addition to the RICS on-line survey and the use of BSRIA proprietary data, BSRIA undertook a series of interviews with stakeholders across the EU, to gather data on technical systems installed in non-residential buildings. To maximise the data gathering opportunity the focus was on typical space heating and cooling, domestic hot water supply, metering and controls, lighting, and energy performance.

However, during 2019 the Consortium were provided with a dataset from the EU project Hotmaps, which provided statistics on the EU residential built stock and estimations for the non-residential built stock, by building type and by building age. Additionally, these datasets also provided measurements for the EU building stock floor area, ownership and occupancy of the unit, U-values and construction materials. The data provided from this source duplicated much of what had been collected via the surveys and largely resulted in the survey data being superseded. The Hotmaps data was given priority following a review of the coverage and approach/methodology and has been used to populate a number of indicators within the BSO database.

Collaboration

To further contribute towards data gap closures, the consortium sourced data from National Statistical Offices, Eurostat and Odyssee-Mure databases on a quarterly basis. In addition, potential synergies between data partners was anticipated to address a further proportion of the BSO data population, however this proved a challenge following the introduction of GDPR legislation and the relative lack of engagement from the data partners. The Consortium also had various discussions with CA EPBD, EASME, JRC regarding access to relevant data sources which seemed encouraging but ultimately proved overall to be unsuccessful.

3.3.4.8 Observations/recommendations

The existing data gaps, and temporal data entries within the database are an issue for any future phase of the BSO. The main concern is that the indicators which represent persistent data gaps don't have data that is widely known (e.g. Space heating by construction period, Topic 1), or publicly available (e.g. EPCs publicly displayed, Topic 3). The significance of this is that these indicators can't currently be populated and will most likely continue to remain unpopulated throughout future phases of the BSO unless the EC can secure access to the required data sources. Furthermore, the indicators which hold temporal data entries were mainly sourced through EU funded projects, which have all now reached completion. The implications of this are that these indicators are unlikely to be populated by further data entries and will therefore be regarded as persistent data gaps in future phases of the BSO.

In addition, some of the data entries in the database were populated through calculations, the methodologies of which haven't been provided to the current Consortium through handover documentation from the original consortium delivering phase 1. Consequently, these approaches can't be duplicated which means that these indicators haven't been updated during this phase of the BSO. If alternative calculations were to be applied, it would result in inconsistencies between the former data points and the data produced from the new methodologies, making the data conflicting and unreliable.

In hindsight and in our opinion, the BSO database should have been created with existing and available energy efficiency data in mind. The list of indicators within each topic in the database appears more of an aspirational list of energy efficiency indicators, rather than realistic key energy efficiency metrics which could have been populated through available data. Additional indicators could have been added when actual new data was documented and made publicly available. The significance of this is reflected in the indicator rationalisation we have undertaken and proposed, in which we recommend that approximately 69% of the existing indicators were removed from the database due to the lack of available data both currently and predicted future data availability for those indicators.

Going forward, open source data availability will likely remain unchanged and engagement with other relevant EU projects to identify any potential data sources to address data gaps should continue during future phases of the BSO. Further, in Section 3.6 of this report (see Annex 6) we briefly outline a future view for the BSO based on the feasibility of launching an EU Buildings big data initiative to address some of the current data access issues.

3.4 Task 3: Acquisition of Data

Task 3 was based on the utilisation of a separate €300k data acquisition budget to allow the Consortium to source further data to populate the BSO. Table 3.13 (below) identifies the sub-tasks associated with Task 3 and describes the status of these activities.

Table 3.13: Task 3 Progress

Activity	Status
Engage with National data partners and other EU projects	Engaged in summer 2017 with limited success. Alternative approach to utilisation of data allocation budget prepared and submitted. Started new engagements with other EU projects late 2018 and potential data partners in early 2019
Target member states and Topics that represent persistent data gaps	Based on indicator prioritisation (Annex 2)
Data requirement documents produced	Indicator prioritisation developed to inform data procurement exercise with tenderers
Data acquisition proposals received and evaluated	Recommendation made to ENER regarding preferred offer
Achieve ENER sanction and procure data	ENER decided not to proceed with Task 3 based on offers received

3.4.1 Task 3 Activities

Before the data acquisition plan was created, the Consortium first wanted to assess the enthusiasm and ability of the existing national data partners to provide new data to the BSO. Each data partner was contacted via email to introduce the Consortium and then contacted again with a request for proposals to address the data requirements (i.e. the full list of indicators and supporting data points) by indicator and topic for the relevant country.

The previous Consortium had a list of 20 contributing national data partners of which, three responded to the current Consortium's request to provide proposals. All three proposals would have equated to over a third of the total data acquisition budget (+€100,000) and data coverage for the three countries would have still largely been disconnected and incomplete. As previously described, the Consortium established additional networks with other communities in the EU whose primary objective is to describe the scale, nature, and characteristics of both the residential and non-residential built stock in terms of energy efficiency. These communities include research and academic communities, such as; University College London, University of Chalmers and other organisations such as Climate KIC.

The Consortium undertook a number of key activities associated with the delivery of Task 3, these included:

- Made contact with the 20 national data partners from the 1st phase of the BSO (August 2017) to introduce the Consortium and assess appetite of data partners to contribute to the 2nd phase of the EU BSO
- Created data requirements documents issued to data partners
- First round of communication to data partners (August 2017)
- Second round of communication to data partners (September 2017)
- Telecommunication with 3 of data partners (September – October 2017)
- Receipt of fee proposals from 3 data partners, price of data from 3 partners represents over a third of the total data budget (€300k) for 3 out of 28 MS which we believe does not represent good value for ENER
- Lack of national data partner engagement – this was likely due to the lack of quality data available to address the indicators once EU projects had closed
- Updated data requirement documents issued to potential new data partners in November 2018, which highlighted where the Consortiums priorities lie for indicator data population;

- Face-to-face meetings with two separate potential data partners have occurred (January and March 2019);
- Telecommunication with 3 other data partners (October 2018 – January 2019);
- Continued to reach out to potential data partners, many of whom presented themselves at the second EU BSO Stakeholder Workshop;

3.4.2 Collaboration

Over the past 12-months two supplementary presentations about the BSO were given, the first was at the EPBD committee in September 2018 and the second was at the EPBD plenary in November 2018. These two events encouraged potential data partners to engage with the Consortium about potential future data partnerships. In addition, at the second BSO Stakeholder Workshop more potential data partners approached the Consortium to explore possible future collaborations. Navigant who were working on the EC service contract for a comprehensive study on renovation rates and NZEB uptake across the EU, and CUES B.V. both expressed an interest in the potential provision of data for the BSO. Both entities were asked to prepare proposals for the provision of "new" data to address some of the persistent data gaps within the specific data allocation budget for this project.

3.4.3 Data acquisition

Based on the rationalisation of existing indicators and prioritisation of indicators (undertaken as part of Task 1), Topic 1 Building Stock Characteristics and Topic 2 Technical Systems were both identified as the priority topics to focus any potential data collection, with several indicators included from Topic 3 Certification and Topic 4 Finance. This analysis initially informed where the data allocation budget would be best directed.

The terms of reference (ToR) for the *Task 3: Acquisition of Data* requirement were taken directly from the original EU BSO service contract tender documents drafted by ENER, as follows:

"For important indicators, in order to improve data quality and to bridge all the data gaps, the acquisition of data should be considered from commercial suppliers, national/regional agencies, experts, etc. There should be the highest possible level of disclosure of this data to the public, even if at an aggregated level. Data sources need to be indicated on the website. The offer should include a detailed plan, prospective sellers, estimated budget for acquisition of data and the data needs that would be addressed. The cost of buying data will be reimbursable up to a maximum amount of €300,000".

Two proposals were received to address the terms of reference as outlined above, they were based on the data indicator prioritisation documents as circulated by RICS in mid-May 2019 and various discussions with both parties prior to submission on 4th June 2019. A summary of both proposals is provided below:

Navigant provided a reasonable offer that covered some of the requirement and could provide longitudinal data. However, the reliance on survey-based data collection (1,372 interviews) introduced an element of risk in terms of potential response rate and actual data received/offered within the timescale available. In addition, there appeared to be a future focus adopted rather than addressing the immediate data requirements of the BSO and the ToR. Furthermore, the proposal made reference to the use of data collected for the nZEB service contract which should be made freely available to the BSO, the value of its inclusion in this context was therefore questionable.

CUES B.V. provided a clear and concise offer that detailed exactly what and how the data would be provided within the constraints of the task. The method was largely based on existing cross-sectional data which would be extrapolated, the modelling assumptions were clear and transparent with a detailed task description provided. The approach provided good coverage in terms of missing data indicators across most topic areas for a single year (2017) and would therefore make only a partial a contribution to closing existing data gaps within the BSO.

Based on our evaluation of the two proposals received, combined with our extensive knowledge of the existing BSO data position, the most economically advantageous offer was provided by CUES B.V. However, despite the fact that our recommended solution provided a significant contribution towards addressing the remaining high priority data gaps within the BSO database ENER later decided not to proceed with utilising the data

acquisition budget for this phase of the BSO. We suggest that this should be taken forward in a subsequent phase of the BSO.

3.5 Task 4: Revamping the Existing Website and Database

Task 4 was concerned with an initial review of the existing database and tools with a view to enhancing the interlinking of data, the navigation of the website and the user-experience. Based on the feedback from the first stakeholder workshop (January 2018) and subsequent client feedback on the revamped website, concerning usability and interface, we adapted our design to address the issues highlighted.

3.5.1 Revamping the existing website

During the kick-off meeting (11/05/2017) ENER outlined that the revamp of the existing website was one of the priorities of the service contract. This encompassed:

- A revamped website to be live six months from the project kick-off
- The revamped website must be in line with the proposed EUROPA design which all EU websites are migrating to (UI requirement / stipulation)
- Reaching a new target market (i.e. the EU citizen)

During the first client meeting regarding Task 4 (29/05/2017) ENER outlined the key priorities for Task 4 as follows:

1. Revamp of the existing website to include:
    a. Mapping tool
    b. Interlinking of datasets
    c. Ability to layer data
    d. Permit country comparison
2. New target user groups to be identified:
    a. Integrate user feedback
    b. Make site accessible to diverse group of stakeholders (from policy makers through to citizens while continuing to provide the granularity of data that technical users require)
    c. Ensure site is accessible

3.5.2 Methodology

The Consortium approached Task 4 utilising a methodology consisting of three stages; exploration, ideation and execution outlined in Table 3.14 below.

Table 3.14: Task 4 progress

Activity	Status
Exploration documentation regarding existing website and tools	Completed with results presented Sept 2017
Ideation documentation regarding potential solutions provided	Completed with results presented Oct 2017
Execution and delivery of revamped website and tools	Completed with final handover in May 2018

Task 4A: Exploration

The desk review, conducted by the Consortium, included competitor analysis, key evaluation questions (terms of reference for the website review: branding, typography, UI layout, UI interactive elements, responsiveness, current code and platform), data, insight data collection instruments, sampling strategy and limitations to the evaluation (Figure 3.15).

Figure 3.15: Exploration Activities

The first phase consisted of interviewing stakeholders/users to gain an understanding of their use of the existing website and any problems/issues that they faced.

The interview strategy included a 3-week campaign of telephone and email communications. The UX interview concept was task based, with a duration of 1 – 1 ½ hour and enabled the Consortium to understand day to day usage, discuss user requirements in more detail and show users wireframes of a potential new UI. The Consortium conducted nine stakeholder interviews and two online usability surveys (Microsoft desirability test and demographic data – not public). The second round of the UX approach included for the testing of prototypes with an additional two users (test items from different navigation areas, such as top-level/second items, header/footer etc. and performed A/B testing of image placement).

Key to our approach was to recruit stakeholders/users who could/would become long term partners in the process of creating the best user experience possible for the website. We believe the stakeholders/users who took part in the study to be representative of the 'real users' and potential 'new users' of the BSO website and therefore our target market of subject matter experts and non-experts (i.e. EU Citizens). In addition, the Consortium reviewed existing user statistics (web analytics provided by ENER). The user interviews conducted by the Consortium established stakeholder business needs and provided evidence on the existing user experience.

The UX approach determined user expectations for a new or modified product feature, functionality or piece of content and was converted to requirements (quantifiable, relevant and detailed). The definition of each was recorded and prioritised as per their impact (user experience, customer experience, onsite conversion rate, download rates, page download speed are just a few examples of valid impact factors). These issues were translated into a requirements list (Annex 9) on the 24th October 2017 and presented to ENER during Task 4b on the 30th November 2017.

During the second phase the Consortium established and confirmed the client requirements for Task 4a. This activity progressed over several months beginning 31st May 2017 and culminating in a presentation to the client on the 28th September 2017. The requirement analysis included current limitations and opportunities for the service. As part of the analysis of requirements, the Consortium reviewed key information, data and

access/permissions and created tracking of all documents reviewed and/or required. Table 3.16 provides a list of the documents submitted in support of the delivery of Task 4.

Table 3.16: Deliverables Submitted

Deliverable	Comments
EC Digital Guidelines	Providing overview access to web tools, access to Wiki etc.
Wireframes (EUROPA)	Details on menu, layout
Visual Identity Manual	Brand, UI guidelines
Web tools (accessed via ECAS account)	Build information
Website Governance (accessed via ECAS account)	Procedure for delivery
User statistics	Google Analytics

From a software evaluation perspective, the following were analysed and reviewed by the Consortium:

- Database architecture;
- Per-country and general gap identifications;
- Gap notification module development;
- Data collection robot development (in case we have application programming interfaces (APIs) for third-party databases);
- Integration with EU data portals (RESTful);
- Data validation;
- Data source validation;
- Website UX improvements;
- Website redesign;
- Analytical tools and graphs;
- Document downloading;
- Social media tools;
- Historic data maintenance; and
- Archive setup and maintenance.

The existing Drupal module consisted of more than 8,000 lines of code and multiple functions with different concerns (query, presentation, helper etc.). Although this was not relevant to user experience, it would have affected the ability to maintain the code base in the future.

The Consortium split the existing single module into multiple modules to remove ‘technical debt’, improve performance and allow for multiple team members to work on the code base in parallel.

The existing pages made several http requests for each page interaction (for example clicking a link in the navigation), this affected usability and performance. Some of these requests triggered lengthy SQL queries on static data (calls to the backend database). We recommended optimisation, such as implementing caching (dependent on co-ordination with DIGIT).

Some of the requests resulted in full page reloads and others made only AJAX calls. This behaviour was not consistent and effected user experience. The consortium recommended the AJAX requests to return HTML code, replacing whole blocks of data on the screen. This would not affect the user experience significantly but increases the amount of data transferred. In our opinion it's preferable to return JSON formatted data and pass it to JavaScript functions, which subsequently manipulate the code. This was something we explored during the Ideation phase.

Phase 4 consisted of the heuristic analysis, executed by usability experts and the salient points raised were as follows:

- User profiling/personas
- Consistency (coherence)
- Features exposure
- Focus
- Context
- The principle of grammar
- Metaphor
- Visibility of system status
- User control and freedom
- Error prevention or recovery
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help; and respect to human limitations.

To complete Task 4a, the Consortium evaluated the 'as is' user journeys (expert evaluation of current system/usability review, heuristic analysis), the client 'do and do not like' document (provided by ENER 18/07/17) and validated wireframes and prototype designs with feedback by users/non-users.

Deliverables

These were the two key deliverables from Task 4a:

- Stakeholder interview results and evaluation report;
- Expert evaluation report.

These were presented to ENER on the 28th September 2017.

Task 4B: Ideation

The Consortium developed documentation that will aid and define future development of the platform (post Task 4).

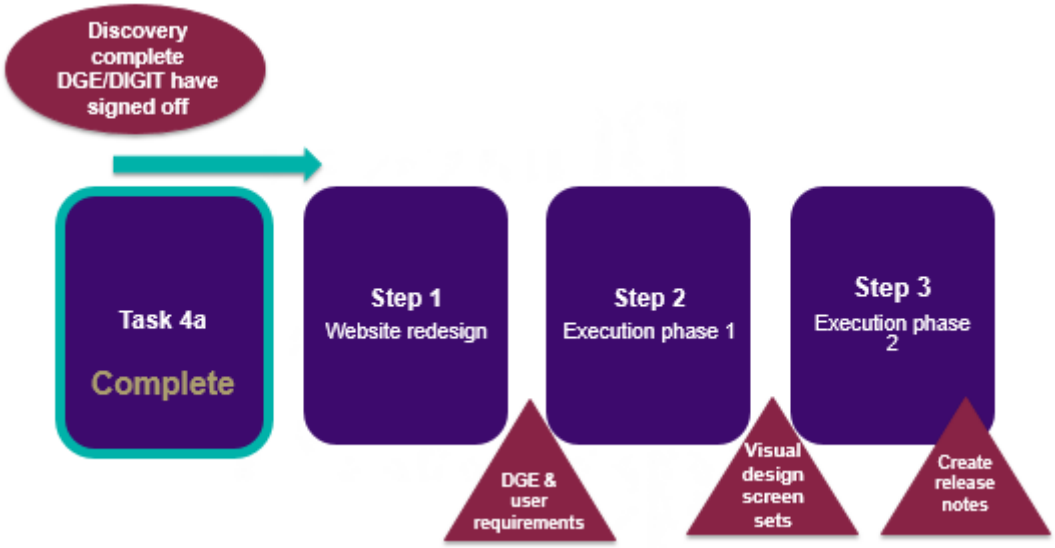

Figure 3.17. Ideation Phase

Activities

The tender specification ENER/C3/2016-547/01 was used as an initial set of requirements for the backlog. The high-level scope was re-evaluated and elaborated further in the Inception Report for this project.

Further activities included:

- Information architecture analysis;
- Information architecture presentation to customer;
- Low-fidelity wireframes creation;
- Iterative creation of high-fidelity wireframes;
- High-fidelity wireframes presentation and discussion with customer;
- Approval of information architecture and wireframes;
- Forming the final backlog of requirements regarding revamping of the existing website and database.

The Consortium prioritised requirements and created backlog items according to documentation from the previous phases.

The Consortium created two different UX designs for three web tools, one design as per client request and one design as per the Consortium's recommendation based on UX methodology (GOMS = Goals, Operators, Methods, Selection Rules) presented 28/09/17

Despite the fact, that the change to the EUROPA design version (drop down menu) seem insignificant, they actually impact not only the web version of the site, but also the mobile version.

The Consortium compared the effectiveness of the two design options using the GOMS methodology. The version recommended by the Consortium, performed better, due to greater clarity, standard use of components and because it was faster/performed better. That meant that to perform the same task using both versions side by side, our recommended version allowed the user to achieve the task 16% faster.

Deliverables

- Final backlog of requirements regarding revamping of the existing website;
- Revamped website information architecture confirmed;
- High-fidelity wireframes created.

Task 4C: Execution

The Consortium conducted a Discovery workshop, where backlog items were converted into 37 user stories. The team implemented necessary changes in the BSO website in line with the documentation created from the previous phases.

Activities

The Consortium identified and implemented iterative release bundles comprising of backlog items on revamping the existing website.

The following investigations took place:

- Code quality;
- Testing and acceptance;
- Solution completeness;
- Documentation; and
- Release preparation (working closely with DIGIT)

Deliverables

Key deliverables included:

- Visual design screen sets;
- Impact analysis for website updates/improvements;
- Update of the code base;
- Release notes outlining changes to the code base (to be provided once website final);
- Functional and integration test results; and
- Maintenance and continuation of the BSO once launched

3.5.3 Summary of main deliverables for Task 4

1. **Revamped website now has a new UI/layout that reflects user expectations**

The Consortium recruited its own pool of experts and delivered an experience that was positively received at the first stakeholder workshop held on 31st January 2018. The Consortium executed a bespoke approach and methodology according to the client's branding and brief.

- Brief: Target EU Citizen, while the Observatory is generally used by technical experts, it was felt that there is generally a lot of interest in energy efficiency from members of the public and therefore the site should be just as easy to use for the average EU citizen as it should be for someone that fully understands the data
- Tools: Competitor review, Comparable review (other EU tools)

- Outcome: Drop-down demonstrated in responsive design did not work. Drop-down demonstrated as less efficient than Consortium recommended design using left hand navigation.
- Delivered new UI – Database / DataMapper / Factsheet

2. **Optimal view and interaction experience across a wide range of devices**

Support for desktop, laptop, tablet and smartphone has been optimised. User Testing and implementing GOMS UX methodology during development of UI design delivered fully responsive designs for all devices.

3. **Responsive design**

As measured during our interactive session at the first workshop. Responsive design was tested in Task 4A Exploration and Task 4B Ideation stages, with users and by experts. Prototypes initially created for user testing in Task 4A were then further developed with 2 designs being presented at the progress meeting on 28th September 2017. 3rd iteration prototypes delivered during the client and user workshop in January 2018.

4. **Improved navigation, added functionality**

Following comprehensive UX design we delivered tools that were considered easy to use by the attendees at the workshop for which the majority were BSO users.

5. **Flexibility for potential introduction of new tools**

Further development to the code base has improved the stability and performance of the BSO website thus facilitating the introduction of new tools as/when required.

6. **Flexibility to allow for the introduction of new data and produce new indicators for regional data and new MS**

Ability to add regional and MS data is feasible if/when new data becomes available.

7. **Two new data visualisation tools for data and indicators in the form of graphs and related tables**

The ability to download data in different formats has been provided. All graphs and tables allow for the data to be downloaded in a number of formats. Original / current webtools are downloadable as per: Database, DataMapper and Factsheets: pdf, image formats (jpeg, png, vector).

8. **The presentation and analysis of time series is now flexible regarding the selection of frequency**

The Consortium has provided the ability to select multiple years and display the associated data. Integration with navigation – Items / Country / Years. Thereby permitting the display of current annual data and historical data.

9. **Additional**

The Consortium has improved performance across all three tools. Specifically, the DataMapper and Factsheet tools now work correctly.

3.5.2 EUROPA Platform Upgrade

Initial delivery of the Task 4 output was on the 15th December 2017. Following feedback received from ENER the Consortium implemented some suggested changes. The delivery of the final version of the revamped existing website occurred on 23rd May 2018 with a view to being re-launched by mid-July 2018 following a formal QA process.

However, on the 18th June 2018 DIGIT informed the Consortium that the EUROPA platform would be upgraded (to Drupal version 7, 2.5) and that this was expected to be complete by the 28th June 2018. The Consortium were therefore instructed to ensure that the delivery on the 23rd May 2018 would be compatible with this upgrade. The Consortium proceeded to test and correct any breaking changes following this upgrade.

The EUROPA platform upgrade was delayed, and further instruction was received to deliver the changes on the 9th July 2018 ready to start the QA process.

3.5.3 QA Process

The QA process began on the 9th July 2018; however, the platform upgrade was not deployed to the production environment until later in July and therefore changes were not initially pulled from the repository until the 9th August 2018.

An initial review of the deliverable asked for a number of points to be resolved including:

- Coding style violations,
- Removal of commented code,
- Installation process for ENER Building (hook_update),
- Issue with the site.make file.

The Consortium made the necessary changes and created a new pull request on the 5th September 2018 which was then rejected due to other issues that were identified on the 6th September 2018. Following further changes, the Consortium delivered a new pull request addressing these issues on the 10th October 2018.

The QA process continued for a period of time due to the code base not passing automated phpcs and drone tests. However, the Consortium did not have the phpcs configuration that DIGIT were using or access to so were unable to be proactive at the time which severely delayed addressing the issues highlighted during the QA process. The Consortium then asked for a conference call with the QA team on the 18th October 2018 in an attempt to resolve any outstanding issues in a timely manner, however this did not occur until the 30th January 2019 and in the meantime, the Consortium continued to resolve issues on a rolling basis.

Following the conference call on the 30th January 2019, it was agreed that the remaining issues would be accepted as is and the deliverable deployed to the playground. The deliverable was deployed on the 31st January 2019 and following an initial test it was found that the playground was not an exact copy of the production environment and therefore there were issues with supporting scripts. These issues were resolved by DIGIT on the 5th February 2019 and testing then began in earnest.

During the testing process, the Consortium identified some issues with the look and feel of the tools as a result of the EUROPA platform upgrade. Due to other projects requiring deployment to production, the EU BSO was pulled from the playground on the 7th March 2019 for these issues to be resolved. However, further issues were subsequently identified via the QA process and it transpired that there were revised coding guidelines issued by the EC to accompany the new EUROPA platform. However, these guidelines were not shared with the Consortium until the final week of April 2019 and had a significant impact on our ability to finalise the website revamp.

Overall, the QA process for the BSO website deployment had been ongoing for almost 18 months in which time the Consortium made multiple fixes to the code of varying complexity. The final iteration of the revamped website was finally made available to the public in early 2020. A summary document detailing our engagement with DIGIT during the QA process is available in Annex 4 of this report.

3.5.1 Summary of Enhancements

Based on the feedback following the interactive session with stakeholders at the 1st BSO Workshop, the following enhancements were made to the delivery of Task 4:

Table 3.18: Enhancements to Existing Website and Database

Tool	Enhancement
Database	Refined the navigation to make it clearer what indicators / countries / years have been selected.
	Improved the indicators hierarchy providing the user with a clearer understanding of where they are within the hierarchy.
	Improved the view of selected indicators / countries / years and provided a clearer separation of selected items.
	Provided an EU-28 average where the selection of indicators / countries / years makes sense.
	Improved the ability to remove one or more selected indicators / countries / years, with a clearer understanding of what will be removed.
	Provided a default selection of indicators / countries / years when the user first accesses the tool allowing the user to more easily understand how to work with the tool.
	Improved the styling of tabular data.
DataMapper	Refined the navigation to make it clearer what indicators / countries / years have been selected.
	Improved the data visualisations, specifically ensuring that where a lot of data is being visualised, the labels are clearly visible.
	Improved the indicators hierarchy providing the user with a clearer understanding of where they are within the hierarchy.
Factsheets	Improved data visualisations.

3.5.4 Observations/recommendations

The issues and delays that effected Task 4 had a significant impact on the Consortiums delivery programme and resource allocation schedule for the entire project. Given that the role of DIGIT was pivotal in the QA and approval process for the launch of the BSO website revamp and that the Consortium had to wait while multiple websites were being tested following the EUROPA platform upgrade it makes sense for any future developments to be undertaken by DIGIT.

Going forward it is recommended that the EC use their in-house resources who have both the expert knowledge of the EUROPA platform and an intimate knowledge of the EC's internal processes to maintain/develop the BSO. It is our opinion that this will prove far more efficient than working with an external technology provider given the nature of the platform and the tools that are currently developed in-house of which the BSO is but one.

3.6 Task 5: Feasibility Study for Launching an EU Buildings Big Data Initiative

Throughout the initial stages of the BSO service contract the delivery Consortium had been examining the issues of data collection and ways of addressing the persistent data gaps that existed within the current iteration of the Observatory. It is our view that the adoption of a 'big data' approach could secure its future operation and relevance.

This section evaluates the feasibility of developing an EU Buildings big data initiative as a way of enhancing the Observatory through access to existing, alternative and additional data sources. These data sources could include the physical attributes, operational characteristics and energy consumption of actual buildings.

Our understanding of the requirement for the 'big data' initiative was that it should be able to: store; aggregate; analyse; and present data from individual buildings and dwellings (as well as other potential data sources) and therefore suggested to us a potential linkage to data obtained from 'smart meter' technology among other things, i.e. BIM and other property data such as EPCs, FM systems, BMS, etc.

Currently the data requirements of the EU Building Stock Observatory are addressed in several ways, including both secondary and primary data collection. However, we are aware of many of the issues currently facing ENER in the operation and maintenance of the BSO. Of particular relevance is the future access to data sources both old and new to guarantee the relevance and accuracy of the BSO going forward. Given that a census of the energy use in the built stock across EU MS is currently practically unachievable, in the sections that follow, we outline the issues that we believe based on our own experience will be important considerations for ENER in the specifying and successful implementation of a big data initiative.

It is our view that the future collection of EU building stock data could be enhanced through accessing different data sources on building characteristics and building energy consumption, this data could then be used to produce bottom-up statistics and other information that could be used to monitor and evaluate progress towards climate change goals and supplement the BSO.

This section looks at the feasibility of developing a database and big data tools to tap into different data sources including data on the physical and operational characteristics and energy consumption of actual buildings. Broadly speaking the report explores different formats for the EU Buildings big data initiative based on current best practice in big data collection and analysis, an assessment of models for running the approach with a cost benefit appraisal of the different approaches is also provided. Finally, the Consortium provides a recommendation for the adoption of a specific approach which includes the technical terms of reference and the outline planning for the implementation of that approach. Table 3.19 identifies the sub-tasks associated with Task 5.

Table 3.19: Task 5 Progress

Activity	Status
Development of potential Big Data options available to ENER	Concluded that a big data initiative was feasible from a technology standpoint, but noted that securing data would remain an ongoing challenge (Annex 6)
Provide a cost benefit analysis of each option	Provided analysis of two options: one for immediate procurement and one more long-term and aspirational (Annex 6)
Recommended a long-term sustainable data collection strategy	Completed to inform procurement options for any future phases of the BSO (Annex 6)

3.6.1 Task 5 Activities to Achieve Deliverables

The full Task 5 report (see Annex 6 of this report) provided recommendations for an approach to address the future data requirement for the BSO via the use of "big data" and artificial intelligence, which included the main elements of: visual interface; technical requirements of the database; analytical capabilities; assessment of the feasibility of adopting similar approaches outlined in the case studies reviewed; and included proposals for the

validation of any data and issues related to sensitivities around data collection and protection. The report output provided the following coverage:

- Assessment of the best available options to address the requirement;
- Cost / benefit appraisal of those options;
- Recommended solutions to address the requirement based on the cost / benefit appraisal;
- Proposed technical terms of reference for the chosen solutions; and
- Proposed implementation plan based on the chosen solution.

3.6.2 Feasibility Study for Launching an EU Buildings Big Data Initiative

The current overarching purpose of the existing BSO is to monitor and measure the effectiveness of energy efficiency policy across the 28 MS via historical data. One of the potential benefits of the big data initiative is the potential for prediction via forecasting. Forecasting energy use / demand would be possible via the initiative as real-time data would allow near real-time updates of energy use, frequency and intensity of use of a building. This will allow through the use of predictive analytics combined with weather modelling, future forecasting for energy demand across the built stock. This type of forecast data would permit policy makers and building owners / occupiers to make informed investment decisions about the stimulation and targeting of specific building types by age, size, location in advance of future reductions in the energy efficiency performance of those buildings or peaks in their energy demand requirements. It could also be used to inform owners / occupiers of the potential payback periods of any energy efficiency initiatives that they may be considering adopting.

Big data collection and analysis

At the time of writing, there is currently no international standard for big data collection and analysis, although there are various standards for data more generally including ISO: 8000-8: 2015 (Data quality). Best practices therefore vary from sector-to-sector and business-to-business, as requirements and regulations differ. However, most organisations tend to use government policy and regional legislation as guidance on the collection and storage of data.

3.6.3 Large-scale collection of building characteristics and building performance data to produce energy efficiency statistics, quality indicators and benchmarks

The Consortium is aware of many of the issues regarding the collection of relevant data pertaining to the energy efficiency of buildings. For the past year we have been working on the service contract ENER/C3/2016-547/01 to *Maintain and Update the EU Building Stock Observatory (BSO)*, a large part of that contract involves understanding the availability of energy efficiency data related to the building stock of EU Member States (MS) and collecting that data where available to populate the BSO. Furthermore, as an organisation RICS understands many of the data issues currently facing ENER in the operation of the existing BSO. We provide online data services to our subscribers and we are subject to many similar issues, in particular ongoing / future access to relevant data sources and persistent data gaps.

Based on the experience outlined above it is our view that access to relevant data sources regarding the energy efficiency of the built stock will remain challenging in the near-term which is likely to impact on the feasibility of the *Big Data* initiative as proposed. Indeed, our engagement with other organisations / initiatives operating in this field suggests that there is a common theme of a lack of available / relevant data concerning quantifying the energy efficiency of buildings, with several initiatives currently operating to try and bridge the data gaps (Building Passport Initiative; KIDs; EPBD; IEA; and various academic institutions).

In terms of any proposed data structure for the initiative we suggest that a review of relevant / available data sources is crucial upfront, and that any database is structured around the availability of data rather than identifying data indicators that would be useful and then attempting to populate those retrospectively as is the case with the current iteration of the BSO. The latter approach is likely to lead to significant data gaps and could draw into question the reliability of the initiative as proposed.

As we see it there are a number of further questions for ENER to consider prior to commissioning:

- Is the initiative needed immediately?
- What is the definition of value for ENER from the initiative?
- Does the initiative have to be standalone owned by the EC or is collaboration / partnership possible?
- What is the maximum budget available to fund?
- What does success / failure of the initiative look like and what would be the impacts of either?
- Are there other routes worth exploring (i.e. further development of the BSO)?

Cost / benefit assessment of options

Given that the success of the big data initiative, as proposed, would largely be dependent on the availability and ongoing access to relevant data, we believe that this is a fundamental issue for ENER to consider. However, there are options to populate the initiative and improve the data position in relation to the energy efficiency of the building stock. In Table 3.20 below we provide a cost / benefit assessment of four potential approaches of direct relevance to the feasibility of the big data initiative.

Table 3.20: Cost / benefit assessment of alternative options for the big data initiative

	Primary Data Collection	**Secondary Data Collection**		**Modelled Approach**
	Census survey of the built stock / energy efficiency across EU MS	3rd party data acquisition via MS	Data mining (web-based via smart meters, etc)	Elements of: primary data collection; statistical sampling; estimation; modelling
	*Indicative rating 1–5 (**5** complex / highest / greatest to **1** simplest / lowest / easiest)*			
DATA				
Initial cost	5	4	3	3
Ease of access	4	3	3	4
Formats	4	3	2	2
Structure	5	3	1	2
Confidence	4	3	1	3
Sustainability (control)	5	3	2	3
Ongoing cost	5	4	3	3
TECHNOLOGY				
Initial costs	5	5	4	4
Ongoing costs	5	5	4	3
IP/LEGAL	3	3	5	4
QUALITY ASSURANCE	5	3	2	3
DATA PROCESSING OPTIONS				
Manual	5	3	2	3
Machine-based learning	N/A	2	1	2

Summary	Cost	Quality	Benefit	Feasibility
Primary data collection/Census	Very high	High	High	Low
Secondary data collection via desk research	High	Low	Low	Low
Secondary data via web-based data mining	High	Medium	Medium	Medium
Modelled approach	Medium	Medium	High	High

Based on the analysis above the two options with the highest perceived benefit are the Census survey and the Modelled approach. However, it is our view that the scope and scale of conducting a regular census of the energy efficiency of the building stock across all buildings in EU MS is not a feasible option given the costs / time involved, furthermore this would need to be a repeatable activity to ensure ongoing access to relevant data. The most feasible options are the web-based data mining and modelled approaches. However, the web-based data mining option would only be feasible if access to suitable data sources can be guaranteed and at the present time this is not the case (see Section 3.3 for more on this).

Our analysis therefore suggests that the adoption of a combined approach including elements of: primary data collection; statistical sampling, estimation and modelling would likely provide the highest benefit compared to cost and hence best value for ENER and includes potential security / sustainability of data provision / approach going forward.

3.6.4 Options for populating the big data initiative

In our view the most feasible option in terms of access to relevant data comprises a synthesised approach incorporating both bottom-up and top-down data collection as highlighted in Figure 3.21.

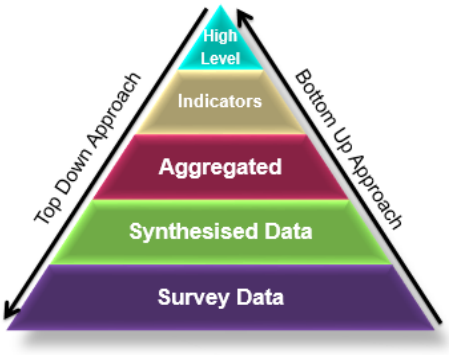

Figure 3.21: Proposed data collection framework

Relevant top-down data should be accessed / collected from publicly available sources including Eurostat, Odyssee and MS National Statistical agencies. The data collection from these sources should be automated as far as possible and machine learning techniques could be established to provide quarterly data collection / upload to the proposed initiatives database (see Annex 6 for more on this).

Given our understanding of ENER's requirements and knowledge of the current position in terms of data availability regarding the energy efficiency of the building stock we believe that the Modelled approach outlined below is currently the most feasible method. However, sometime in the future the use of Smart Meter data could feed directly into the *big data* initiative if data access can be secured (see Annex 6).

Our proposal to address the 'big data' initiative is therefore to focus on the use of real, example and synthesised data. We believe that a crucial element of any initiative would be to understand the quantum of the built stock

across MS and by extension the EU as a whole. Only once the scale and nature of the built stock is defined can the energy performance of the stock be understood.

The approach, as proposed, would require an element of primary data collection to supplement the top-down data already available. We suggest an MS based approach to data collection which could be delivered via individual service contracts commissioned by ENER or mandated through supplementary Eurostat data collection exercises. Given the need to gross any results statistical sampling techniques would be required with random sampling the preferred option.

We believe that the initial stages of the Big Data initiative should concentrate on the provision of relevant strategic data indicators (as listed below), further indicators could be added at a later stage if necessary (it is easier to add indicators than remove them at a later stage due to insufficient data).

The approach as proposed would include the ability to:

- Identify the size, scale and nature of the building stock at MS level;
- Identify the energy demand / efficiency of the MS and by reference building type and segmentation;
- Identify the CO_2 burden by energy carrier for each MS and by reference building type and segmentation;
- Identify the energy efficiency improvements that could ensure a MS delivered to 80% below 1990 baseline in terms of Greenhouse Gas (GHG) emissions;
- Provide a RoI and simple payback periods for lifecycle interventions (this would be useful information for the EU citizen)

The data collected from primary sources would be used to create stochastic models to define the parameters of both the residential and non-residential built stock. This proposed method of characterisation is widely recognised in leading universities as a promising approach to defining and describing the volume of the building stock across Europe.

This classification would produce a building stock typology framework which could be used on a regional, national or pan European scale. The stochastic model would produce a segmented data tree structure defined by the location, type, size and age of the building stock. Then other data topics could be aligned to this structure which would permit the assessment of the energy efficiency of the building stock.

The purpose of the framework would be to quantify the energy consumption of the existing stock across the 28 MS and would include establishing:

- A building stock typology that is harmonised for the 28 MS and identifies additional building characteristics to capture (i.e. geometry), including establishing defined climatic zones;
- A consistent / harmonised methodology for establishing the quantity of built stock across the EU;
- Three methods of data collection and generation (real, example and synthesised) to populate the building typology framework; and
- A methodology for delivering the building typology structure

Using recognised statistical modelling techniques (including establishing a Dirichlet distribution, creating a Bayesian inference and the use of Markov Chain Monte Carlo method for a posterior distribution) it would be possible to generate example buildings from the built stock model that can be used to perform further analysis such as cost optimal modelling. This method is defined through EU guidance notes[1] .

The approach as proposed would permit a snapshot view of the thermal efficiency of the built stock, energy consumption by carrier type and by energy use (i.e. heating, cooling, lighting etc.) across MS and because of

[1] Guidelines accompanying Commission Delegated Regulation (EU) No 244/2012 of 16 January 2012 supplementing Directive 2010/31/EU of the European Parliament and of the Council on the energy performance of buildings by establishing a comparative methodology framework for calculating cost-optimal levels of minimum energy performance requirements for buildings and building elements (2012/C 115/01)

the data structure it would be possible to assess portions of the built stock that are performing poorly and / or would not achieve the required energy efficiency standards even with renovation.

We propose that Lifecycle renovation strategies are created and tested on the reference buildings generated from the built stock model. Reference buildings would be assessed using cost optimal techniques pre and post renovation to determine energy efficiency gains. It would be possible to assess the change in CO_2 burden as well based on energy carrier split and split of energy uses. It would also be possible to assess the level of investment required to reduce the current CO_2 burden of the built stock and determine simple payback periods.

Once established, the method and data structure as proposed could be updated using fewer data points than those required for the existing iteration of the BSO and once smart metering is established more widely the proposed data structure would permit big data capture and analytics.

The approach and methodology as proposed is described in detail in the Task 5 report which is available in Annex 6 of this final report.

3.6.5 Option appraisal

The proposed option outlined would provide a synthesised and consistent / harmonised view of the building stock in terms of archetype, size, age and physical attributes across the EU. While the use of synthesised data could introduce a level of uncertainty into the understanding of the building stock the ability to capture the required level of raw data to be able to provide the same output is not realistic in terms of scale, complexity of activity, access to commercial data, data protection issues, and the cost and time required.

Table 3.22: Assessment of proposed approach for the big data initiative

Advantages	Disadvantages
Potential in future to be connected to any big data initiative that collects real metered data from real buildings	New approach has never been tried or tested on this scale previously
Consistent method to quantify the built stock, both residential and non-residential	User's resistance to change – stakeholder management and consultation required
Quantify the performance of the built stock according to fabric and technical systems	Requires statistical expertise from ENER to agree appropriate confidence intervals – i.e. resource requirement
Identify scale and cost of renovations to achieve GHG emissions targets from built stock	Approach has been used on a small number of countries at a time by academics
Easier to access and interpret data because of parent data structure – i.e. built stock	Requires a GIS mapping tool to make accurate building surface area calculations from
Standardised, repeatable and transparent methodology	Requires significant modelling and assumptions to be made
Built in QA processes	
Based on limited primary and secondary data to create model but post creation model is easily updatable	
Cheaper operational costs – due to automation of updates and reduced data requirements	
Increased data comparability due to standardised method applied to all locations	
Data structure permits cross analytics and more streamlined navigation	

Ability to forecast energy consumption – forward looking indicators so can make interventions	
Ability to forecast energy consumption and monitor progress against EPBD targets	
Based on 'Guidelines accompanying Commission Delegated Regulation (EU) No 244/2012	
The statistical model would provide defensible confidence intervals.	

Once the quantum of the building stock by archetypes has been derived it would be possible to model energy consumption cost to retrofit, to bring up to the relevant code by those specific archetype buildings. The approach can then be further applied to the building stock numbers to establish a view of the energy consumption by country and region from aggregation. A model of this type would make it possible to assess different methods and the potential costs of bringing the stock up to the energy efficiency codes to meet policy objectives.

The approach is widely used by academics to undertake energy efficiency modelling of regions and cities and a summary of the advantages/disadvantages of the proposed method is provided above in Table 3.22.

The approach as outlined would allow for large-scale collection of building characteristics and building performance data to produce energy efficiency statistics, quality indicators and benchmarks across EU MS. In our opinion, this method is the most realistic / feasible option for ENER in developing the big data initiative given the current data position vis the energy efficiency of the building stock across the EU MS.

Estimated cost/time to deliver

Our estimate of the outline costs to deliver the modelled approach as proposed are provided below, these include:

- Database development (including data structure, modelling routines, testing) = €400,000 to €500,000
- Data processing (including data collection, data hosting, data analysis/reporting, testing) = €1,100,000 to €1,500,000
- Personnel costs for ongoing annual operation/management = €250,000
- Total cost to deliver = **€1,750,000 to €2,250,000**

Given the potential costs involved it is our opinion that the initiative as proposed should be delivered / operated and funded by the EC. A centralised function would be important to maintain the initiatives integrity and longevity going forward.

We estimate that the solution as proposed would take approximately 18 - 24 months to develop and would include the following broad activities:

- Data analytics and algorithm development; data analysis, machine learning, testing
- Architecture design, platform build, cloud development, UI / UX development, testing

The option described above could be procured immediately, in Annex 6 to this report we provide an examination of a potentially fully automated approach for the big data initiative and suggest potential data approaches that could be utilised in future if access to those data sources can be secured by the EC.

4. Stakeholder Workshops

The Consortium hosted three annual Stakeholder Workshops, which provided the opportunity to present the progress made over the previous phase of the service contract. The first BSO Stakeholder Workshop took place on 31st January 2018, the second on 31st January 2019 and the third on 27th May 2020.

The stakeholder workshops were delivered to an audience representing the following organisations:

- EU Commission Directorates;
- Non-governmental organisations;
- Professional institutions;
- Technical and trade bodies;
- Research and academic institutes.

The presentation given at the third event is provided in Annex of 7 this report. There were several key suggestions highlighted by the stakeholders mostly with regards to how the revamped website could improve transparency with user journey, data sources and definitional consistency and clarity across indicators.

In addition, methodologies were recommended for quantifying the building stock, including the potential future use of GIS mapping and satellite imagery as a way of overcoming the significant problems surrounding quantifying the existing building stock across MS. These suggestions could be applied in future phases of the BSO, to build upon the dataset received from the Hotmaps (EASME) project.

5. BSO Promotion

5.1 Overview

During the contract promoting the BSO was recognised as an important pathway to increasing the databases visibility within the wider energy efficiency and energy / buildings community. To satisfy this requirement the Consortium produced several thought pieces which have aimed to raise the profile of the BSO, through various mediums. This section outlines the promotion of the database, and supplementary presentations that were given throughout the contract.

5.2 Written Promotion of the BSO

EU Building Stock Observatory Newsletter

In July 2018 the Consortium along with ENER decided that regular releases of an BSO newsletter could be a successful strategy to disseminate progress and news updates concerned with the activities of this phase of the service contract, as additional and more regular material to the annual stakeholder events.

During the second progress meeting in Brussels (July 2018) the first draft of the newsletter was presented to ENER and a final version was agreed prior to the second stakeholder workshop (January 2019). This version was disseminated to all confirmed attendees ahead of the workshop, but not to a wider audience at that time. Since the release of the first draft of the newsletter, the Commissions communication unit revealed that the original format of the newsletter did not comply with the Commissions official communication rules. Thus, a second version of the newsletter was created in line with the official guidelines and with slightly amended content.

Articles

During the latter part of the contract contacts have been made with the UK Green Building Council (UK GBC), who regularly release blog posts and related news articles concerned with the sustainable built environment. Benefitting from this working relationship the Consortium produced a blog style article that was then uploaded onto the UK GBC 'News & Opinion' section of their website, the general demographic of UK GBC website visitors was considered a suitable audience for news relating to the BSO.

Furthermore, introductions were created with the BUILD UP Portal (BUP) which is the European Portal for Energy Efficiency in Buildings. This portal is the epicentre for news, events, and articles concerned with the energy efficiency of buildings across Europe and has thousands of members worldwide. The Consortium saw this as a chance to reach a large potential audience with news on the BSO.

In addition, to the blog style articles another route was taken through RICS Future of Professions publication, whereby an article was published mid-2018. This again was deemed a significant publication to take advantage of to promote the visibility of the BSO, as the subscribers tend to be professionals within the built environment.

5.3 Supplementary Presentations

During the contract lifetime a lot of interest surrounding the BSO has been generated through collaborations with other EU projects and the EPBD, which saw the Consortium being invited to present a summary of the BSO at other related events. The first additional presentation was given in September 2018 at the EPBD Committee and the second in November 2018 at the CA EPBD plenary meeting. The Consortium saw these as opportunities to promote the BSO to a wider audience which included policy makers, the purpose of these presentations was to provide brief summaries of this phase of the service contract, including progress against deliverables, and to report on the BSO's data position.

6. Conclusions and Recommendations

During the life of the project the Consortium sought to harmonise and rationalise the existing data indicators, to ensure confidence in the data contained within the BSO database. The report outlines in some detail the Consortium's actions to address the existing data gaps. While outlining the scale and nature of those gaps, highlighting existing data sources and the potential of big data to address some of the remaining and persistent data gaps going forward.

Whilst describing past activities to deliver the second phase of the BSO this report also provides an assessment of potential future options delivered as part of Task 5. More specifically, the report assesses the cost and benefit of developing, launching and managing an EU Buildings big data platform. In this context, the big data initiative is a large-scale collection of building characteristics and building performance data to produce statistics, quality indicators and benchmarks. The feasibility study sets out a cost benefit appraisal that provides ENER with a view of what is possible from a technology and data stance and assesses a number of options in accordance with the value vs. cost and required resource input for delivery and maintenance.

As the BSO is already in place, using big data technology is the natural next step in its evolution. The current overarching purpose of the existing BSO is to monitor and measure the effectiveness of energy efficiency policy across the EU 28-member states via historical data. One of the significant benefits of the big data initiatives as proposed is the potential it would allow for prediction via forecasting. For the EU Buildings big data initiative, we believe that it will be crucial for building owners / users to be able to predict future energy use of their assets and whether renovation activities would potentially make adequate returns.

In our view, one of the best potential routes to the provision of future quality data, fundamental to the success of this initiative, is to leverage smart meter technology that is currently being installed in domestic and non-domestic properties across the EU. Mandating data collection through Regulations vs. asking suppliers to send meter readings voluntarily to a secured endpoint is considered. We also clearly acknowledge that the future success of any big data initiative will to a large extent be dependent on its capability to ingest data from connected devices and to use that data for its analytical objectives.

The future options considered are flexible, scalable, cost effective and capable of evolving further, as data from a greater number of technologies is incorporated in the coming years and beyond the scope of the development timeframe. Based on our experience and expertise in this area, we believe that the future adoption of a big data initiative that is automated as far as possible would provide the best value for ENER and provide security of data provision going forward. In our opinion, the launch of an EU Buildings big data initiative is certainly feasible from the technology standpoint.

However, the technology platform is only part of the potential solution. The other main component of the initiative is access to relevant data on an ongoing basis and this, we believe, will likely require input from the European Commission on a number of fronts to ensure that the data position significantly improves going forward. Our recommendations to improve the data position of the BSO suggest the adoption of a phased approach going forward, more specifically:

Short-term: Until future access to reliable data sources can be guaranteed on an ongoing basis, we believe the best option for the BSO is for the adoption of statistical/modelled/estimated approaches to supplement data population, cost optimal modelling for instance.

Longer-term: Relevant longitudinal data could become available via wide-scale adoption of smart meters and associated technology across EU Member States, however this may have to be mandatory to be successful.

In conclusion, it is our belief that data collection and population of the database is likely to remain a significant challenge going forward. In hindsight and in our opinion, the BSO database should have been created based on the availability of energy efficiency data. The list of indicators within each topic in the database currently appears to be more of an aspirational wish list of energy efficiency indicators, rather than realistic key energy efficiency metrics which could have been populated through available data. The database is currently not really functioning as a database, more like a list of indicators where data is not currently available hence the preponderance of data gaps. To improve the position, we suggest that the rationalisation of data indicators, as

proposed by us, is undertaken as soon as practicable. Below we provide further options for potential subsequent phases of the BSO.

6.1 Recommendations for the future

This report has provided several broad options for the future delivery of the BSO, a proposed EU Buildings big data initiative, a modelled approach that could be procured immediately and a potentially fully automated future option that could be progressed once access to the proposed data sources has been secured by the European Commission.

The recommendations for the big data initiative as provided could not only work in harmony with the existing approach, but also on a standalone basis. The proposed solutions are flexible, scalable and cost effective, capable of evolving further, as data from a greater number of technologies is incorporated in the coming years and beyond the scope of the development timeframe.

What we have recommended would allow for both the business intelligence and data science / predictive analytics necessary for retrospective reporting, and prospective analysis. Prospective analysis will be essential for forecasting, prediction and benchmarking in terms of energy consumption or the predictive maintenance/adaptation of buildings.

In terms of the potential costs / benefits as we see it, the adoption of the big data initiative could supersede the current iteration of the BSO. Thus, any indicative costs for the big data initiative provided in this document could be offset against the potential cost savings of not procuring the Observatory on a regular basis in future. The establishment of the big data initiative is also likely to involve an initial investment that will not need to be repeated, although there are likely to be ongoing maintenance and operations costs to consider. Further, we believe that our recommended solutions as outlined, will provide a better picture of the energy efficiency of the building stock incorporating both top-down and bottom-up data.

Regarding the user benefits of the approach as proposed, we believe that the big data initiative will provide greater support for building owners/users/citizens to identify and quantify the potential for cost-effective improvements to the built stock to enable energy efficiency savings. For the policymaker the ability to have a bottom-up statistical view with greater granularity will provide a more robust assessment of the EU's and individual Member States progress towards energy efficiency targets and wider climate change goals.

6.2 Immediate options

We appreciate that the options described in the section above will take some time to realise and are at best a longer-term approach. However, in our view there are several important issues that ENER should consider immediately to improve the BSO in the shorter-term. In the section that follows we discuss several broad options based on our experience of delivering the second phase of the BSO contract.

1. Continue as is, i.e. maintain the status-quo with both the database and web platform and proceed with sporadic data population as and when data becomes available, however this option is not discussed in detail as a practical option going forward;

2. Start again from scratch, for instance re-design a database based on the current availability of data and rather than revamp the existing website completely re-develop the web platform, this option would require significant investment by the EC and writing-off the existing initiative and is therefore not considered as a viable alternative;

3. Maintain the current platform and improve/increase data population, this option would involve an attempt at improving the data position of the existing BSO via various routes as outlined below.

Maintenance of the BSO website

Regardless of the decision made to maintain the status-quo or completely re-develop the database and web platform, it is recommended that the Commission's IT directorate (DIGIT) utilise their in-house resources who have both the expert knowledge of the EUROPA platform and an intimate knowledge of the EC's internal processes. It is our opinion that this will prove far more efficient than working with an external technology

provider given the nature of the platform and the tools that are currently developed in-house of which the BSO is but one.

Rationalisation of indicators and proposed new indicators

It is our view that the BSO database be restructured based on readily available data from high quality sources (i.e. Eurostat). One of the major obstacles the BSO currently faces is a lack of publicly available data, in addition a large proportion of the indicators within the database aren't currently measured.

We have proposed rationalisation of the existing indicators within the BSO based on persistent data gaps. Specifically, if an indicator does not hold any data, where it isn't publicly available or unlikely to be measured in the near future, those indicators should be removed/hidden from the public domain.

In our opinion the activity of proposing new indicators should not be started until the current state of relevant publicly available data has been assessed. Otherwise the likelihood is that more persistent data gaps will be introduced into the database. In addition, a revision of the unique codes for individual indicators should be performed as duplications have appeared (See Topic 1.2 code "sbconmf" used for both indicators of biofuel consumption for residential multifamily houses and non-residential buildings) creating confusion.

Furthermore, we would also propose that in any future form of the BSO, the units of measure are standardised (i.e. throughout the database measurements are considered in millions of meters squared to thousands and even meters squared), in addition there should be a move away from using equivalent units (such as mtoe) and adoption of international standard units (i.e.Kwh or Kw in the case for energy).

Proposed data gap filling strategy

<u>Open source databases and EU projects</u>

Access to data is key to delivering a robust BSO and at present it's the biggest obstacle. Data population from existing / future EU funded projects is currently disjointed with no clear coherent strategy for how any data collected could easily be made available to the BSO, this area should be considered during future tender exercises for relevant EU funded projects.

Quarterly data updates are currently performed from open source databases, mainly Eurostat, national statistical offices and Odyssee-Mure. Currently, the quarterly data update is labour intensive and time consuming as data mining these sources isn't straight forward. Moreover, given that Eurostat is currently the primary data source for the BSO, it is suggested that establishing an Application Programming Interface (API) link to enable automatic data uploads in future would save time and any data uploaded could be reflected in the live website more quickly.

Furthermore, other potential data sources (i.e. CA EPBD or future EU funded projects) need to be formatted / structured in a way that has the BSO database structure in mind to allow efficient upload of any data into the BSO. Alternatively, the BSO database could always be adapted to suit other potential data sources.

Current H2020 research calls B4E-6 (Big Data for Buildings) and B4E-7 (EU Building Stock Data) are very similar and with significant overlaps. The expected budget for both combined is €6m which might be enough to develop a pan-European platform (although this sum represents just €214k per MS), however the initiatives are currently budgeted for separately and in our opinion, this is likely to dilute their potential impact.

Another direction would be to investigate the utility of building passport type initiatives as potential future data sources. If successful it is expected that these initiatives could present a bottom-up view of energy efficiency building data in a central database.

<u>Modelling and / or big data</u>

Cost-optimal modelling is a well-founded method that compensates for the lack of available reliable data. However, it should be noted that a statistically significant initial dataset would be required for this pathway to produce meaningful future data for the BSO.

See the Task 5 recommendations (p.35) for our view of the potential Big Data Initiatives currently available to ENER for the future operation of the BSO, however in our opinion this proposal is someway off unless the EC can tackle the issues surrounding access to real-time data.

Primary data collection strategies

During the second phase of the BSO a primary data collection strategy was launched, consisting of an online survey and a series of targeted interviews. In our opinion, a reliance on this type of approach is unlikely to satisfy the future long-term data requirements of the BSO. Nevertheless, this approach could work in parallel with the cost-optimal modelling option described previously.

Data gap filling strategy by Topic

The existing data gaps, and temporal data entries within the database are an issue for any future phase of the BSO. The main concern is that the indicators which represent persistent data gaps don't have data that is widely known (e.g. Space heating by construction period, Topic 1), or publicly available (e.g. EPCs publicly displayed, Topic 3). The significance of this is that these indicators can't currently be populated and will most likely continue to remain unpopulated throughout future phases of the BSO unless the EC can secure access to the required data sources. Furthermore, the indicators which hold temporal data entries were mainly sourced through EU funded projects, which have all now reached completion. The implications of this are that these indicators are unlikely to be populated by further data entries and will therefore be regarded as persistent data gaps in future phases of the BSO.

In addition, some of the data entries in the database were populated through calculations, the methodologies of which haven't been provided to the phase two Consortium through handover documentation from the consortium responsible for delivering the first phase. Consequently, these approaches can't be duplicated which means that these indicators haven't been updated during this phase of the BSO. If alternative calculations were to be applied, it would result in inconsistencies between the former data points and the data produced from the new methodologies, making the data conflicting and unreliable.

In hindsight and in our opinion, the BSO database should have been created with existing and available energy efficiency data in mind. The list of indicators within each topic in the database appears more of an aspirational list of energy efficiency indicators, rather than realistic key energy efficiency metrics which could have been populated through available data. Additional indicators could have been added when actual new data was documented and made publicly available. The significance of this is reflected in the indicator rationalisation we have undertaken and proposed, in which we recommend that approximately 64% of the existing indicators are removed from the database due to the lack of available data both currently and predicted future data availability for those indicators.

Going forward, open source data availability will likely remain unchanged and engagement with other relevant EU projects to identify any potential data sources to address data gaps should continue during future phases of the BSO. Further, in Section 3.6 of this report (see Annex 6) we outline a future view for the BSO based on the feasibility of launching an EU Buildings big data initiative to address some of the current data access issues.

The Consortium have prepared six Topic factsheets outlining where the existing data gaps lie (see Annex 3 of this report). These highlight issues with data gap filling and recommendations for populating each individual data topic within the BSO. The section that follows provides a brief summary of our recommendations for future data population by data topic.

Topic 1: Building Stock Characteristics

The most practical and cost-effective route to maximise data gap filling for Topic 1 is to utilise any new EU project data concerned with the building stock, therefore all relevant future service contracts commissioned by ENER should consider the architecture of the database and ensure that any data collected aligns closely with the indicators within the BSO.

Furthermore, cross-sectional stock data (Hotmaps) already present in the database could be used to provide a longitudinal dataset if used as a baseline and subsequently updated via calculations/modelling using stock replacement rates to update.

Topic 2: Technical Systems

Topic 2 is perhaps the most complex data topic to populate given the inherent technicalities involved, we suggest that ENER will need to explore partnerships with technical experts to populate going forward, although it should be noted that the data is often proprietary and therefore difficult to access.

Topic 3: Certification

Despite numerous attempts to access EPC databases via the EPBD the Consortium have been unsuccessful in their attempts to secure the relevant datasets. Currently, there isn't a uniform system across the EU that obliges MS to record and update EPC databases. Because of this, EPC data isn't publicly available across a number of EU countries, which results in this topic having significant data gaps (+95%). This is an area of concern and a potential quick win if it can be addressed, as unlike some of the other topics the data for Topic 3 exists but is not currently available pan EU. Therefore, we recommend that national EPC databases should be structured/formalised across MS to allow for future use in the BSO. In addition, future public access to EPC data should be mandated by the EC to make it available more widely.

Topic 4: Finance

There are a number of relevant studies exploring the different financial mechanisms in individual MS and during the contract effort was made unsuccessfully to gain access to the data available from these sources. Going forward there is a need to explore synergies with relevant data sources to access any available data. More specifically, exploration of partnerships with 'green banks' is an option for populating the indicators within this topic.

Topic 5: Energy Poverty and Social Issues

Given that Topic 5 is solely populated through Eurostat data establishing an API link to automatically upload the data would make sense from a practical perspective going forward. Moreover, the Energy Poverty Observatory (EPOV) almost mimics this Topic and there should be opportunities to share data across initiatives when available.

Topic 6 [Future]: Building Codes and National Definitions

Population of this data topic would likely involve in-depth desk-top study across MS that utilises both qualitative and quantitative data sources.

Finally, ENER currently takes on the responsibility for the operation and data collection for the BSO, via service contracts with consultants. This centralised function might not be the best route to ensuring adequate data population of the initiative. Based on our experience we believe that the data collection responsibilities could and perhaps should be delegated to individual Member States as they are likely to have a better understanding of relevant national data availability and sources. This approach could be delivered via individual service contracts or mandated through supplements to existing EUROSTAT data collection exercises.