A data strategy to promote the clean growth of UK industries

The role of data in accelerating positive change for industry and the environment

March 2020

Jonathan Norman, John Barrett,
Alice Garvey, Peter Taylor,
Justin Goodwin, Mark Gibbs,
Richard German and Lucy Garland

About this report

Recent legislation commits the UK to reducing greenhouse gas emissions to net zero by 2050. In order to achieve that goal, good quality data will be needed to implement this transformation and inform credible decision-making. This report explores whether the existing publicly available data can meet these new challenges.

Reference

This report should be referenced as:

Norman, J., Barrett, J., Garvey, A., Taylor, P., Goodwin, J., Gibbs, M., German, R. and Garland, L.. 2020. A data strategy to promote the clean growth of UK industries: The role of data in accelerating positive change for industry and the environment. Centre for Research into Energy Demand Solutions, Oxford.
ISBN: 978-1-913299-02-6

Authors:

- John Barrett
- Jonathan Norman
- Alice Garvey
- Peter Taylor
- Justin Goodwin
- Mark Gibbs
- Richard German
- Lucy Garland

Contents

Key messages **5**

1.1 Motivation 7

1.2 Report objectives and approach 8

2.1 Stakeholder needs 10

2.2 Industrial datasets review 13

2.3 Key issues and the need for a data strategy 16

3.1 Data users and needs 19

3.2 Data management 19

3.3 Data quality 22

3.4 Communicating data 25

Annex 1. Stakeholder consultations and engagement **33**

A1.1 Stakeholder interviews 33

A1.2 Stakeholder workshop 33

Annex 2. Dataset assessments **35**

A2.1 Datasets evaluated 35

A2.2 Data Assessment: 38

Annex 3. Policy questions, indicators & datasets identified **42**

Annex 4: Project team **45**

References **47**

Abbreviations

BEIS	Department for Business, Energy and Industrial Strategy
CCA	Climate Change Agreements
CCC	Committee on Climate Change
CCS	Carbon Capture and Storage
CHP	Combined Heat and Power
CRC	Carbon Reduction Commitment (Energy Efficiency Scheme)
CREDS	Centre for Research into Energy Demand Solutions
CRISP	Core Resource for Industrial Symbiosis Practitioners
DECC	Department for Energy and Climate Change (now part of BEIS)
Defra	Department for Environment, Food and Rural Affairs
DUKES	Digest of UK Energy Statistics
EA	Environment Agency
ECA	Enhanced Capital Allowances
ECUK	Energy Consumption in the UK
EPC	Energy Performance Certificate
E-PRTR	European Pollutant Release and Transfer Register
ESOS	Energy Saving Opportunity Scheme
ESS	European Statistical System
EU ETS	European Union Emissions Trading System
EUROSTAT	Statistical Office of the European Commission
FITs	Feed-in Tariffs
FOI	Freedom of information
GHG	Greenhouse Gas
IDBR	Inter-departmental Business Register
INSPIRE	Infrastructure for Spatial Information in Europe
LCP	Large Combustion Plant
LSOA and MSOA	Lower and Middle Super-Output Area respectively – levels of the census geographical system
NACE	Statistical Classification of Economic Activities in the European Community
NAEI	National Atmospheric Emissions Inventory
NFR	Nomenclature for Reporting
NISP	National Industrial Symbiosis Programme
NOMIS	Official Labour Market Statistics Portal
NRW	Natural Resources Wales
ONS	Office for National Statistics
PRODCOM	Database of UK manufacturers' sales by product
QA/QC	Quality Assurance and Quality Control
RDCO	Registered Dealers in Controlled Oil
SDG	Sustainable Development Goal
SEC	Specific Energy Consumption
SECR	Streamlined Energy and Carbon Reporting
SEPA	Scottish Environment Protection Agency
SIC	Standard Industrial Classification
SSEA	System of Environmental Economic Accounting
UED	Usable Energy Database
UKERC	UK Energy Research Centre
UKPIA	UK Petroleum Industry Association

Key messages

Climate change is increasingly being treated as a crisis in both policy and public life, and the UK has signalled its commitment to radical greenhouse gas reductions in recent legislation for a net zero 2050 target.[1] This scale and pace of change requires significant industrial decarbonisation across all sectors. Good quality data will support action to implement this transformation and inform credible decision-making.

In this report, we review whether the existing publicly available data can meet these new challenges. We found that data are not currently good enough for the demands of providing robust and timely analysis while monitoring progress over time. There are serious concerns related to:

- inaccessibility and proprietary ownership limiting the availability of data to users;
- poor transparency in dataset methods, underlying sources and assumptions;
- inconsistency, incompleteness, and incomparability in the use of classifications and taxonomies, units and formats;
- a lack of continuity with irregular updates to datasets; and
- poor detail, particularly technological detail, in existing datasets, limiting the quality and scope of representing industrial sub-sectors.

We conclude that there is an urgent need for a public data strategy which gathers linked data on emissions, technologies, and related environmental, social and economic impacts. We identify that high-quality data and indicators for the monitoring and management of industrial energy should respond to the needs, both present and future, of stakeholders. We further conclude that this ideal dataset(s) should be readily updatable, open access, and independently managed, whilst ensuring consistency in concept, terminology, and classifications. We also suggest that the Department for Business, Energy and Industrial Strategy (BEIS) and the Department for Environment, Food and Rural Affairs (Defra) would be well placed to help establish a central public body that works with key agencies (for example the Office for National Statistics and Committee on Climate Change) with the aim of delivering an improved evidence base on industrial energy, emissions and technologies.

Our ten key recommendations for optimising the quality of datasets in any field fall into four categories and are as follows:

10 recommendations	
Data users and needs	**1.** Identify stakeholders, their roles and responsibilities, and their data needs.
	2. Develop a list of strategic data related questions and indicators.
	3. Establish a list of required datasets and indicators.
Data management	**4.** Appoint a public body (such as ONS) to manage open data access.
	5. Maximise data collection and consider the long-term value beyond a single project.
Data quality	**6.** Ensure consistency of terminology and classifications between datasets.
	7. Standardise the presentation of data, documentation on assumptions and uncertainty and ensure transparency.
	8. Improve data granularity in the industrial sector, in terms of sub-sector and spatial disaggregation, and technology information.
Communicating data	**9.** Disseminate datasets in a consistent, searchable format.
	10. Ensure that all datasets collected using public funding become open access.

1. Introduction

1.1 Motivation

The UK Government has committed to reducing greenhouse gas (GHG) emissions to 'net zero' by 2050[1]. The recent strengthening of the UK's commitment beyond the 80% target to net zero by 2050 clarifies the need for every sector to significantly reduce their GHG emissions as rapidly as possible. UK industry is no exception.

The UK Government wishes to achieve this rapid transition following the most cost-effective pathway. This requires a range of new, and strengthened existing, policies to drive innovation, foreclose highly carbon-intensive options and align cross-departmental objectives. For example, for industry it is important that the Industrial Strategy[2] aligns with the Clean Growth Strategy,[3] the Resources and Waste Strategy[4] and the 25 Year Environment Plan,[5] to name a few. Different strategies pulling in opposite directions will undoubtedly lead to a failure in delivering climate objectives.

Effective partnerships between Government and industry are essential to achieve this rapid transition. There is evidence that this is starting to happen, with the creation of a £144 million fund to drive resource productivity in UK industry comprising a £66m contribution from the UK Government and a further £88m from UK industry. Further to this a £170m Government fund is being made available under the 'Industrial Decarbonisation' challenge of the Industrial Strategy Challenge Fund (Wave 3) to deliver industrial energy efficiency, decarbonisation and carbon capture through the formation of net zero industrial clusters.[6] The Industrial Strategy Challenge Fund is part of a broader £4.7bn government research and development investment scheduled over four years.[7] Additional sector-specific funds include the recent £250m Clean Steel Fund.[8]

The UK no longer has the opportunity to gradually introduce new technologies and allow for the moderate uptake of low carbon options. Instead, the UK Government acknowledges it needs to take an active role in driving innovation and correcting market failures. Such an approach requires robust evidence to ensure that environmental, social and economic goals are met.

From an economic and social perspective, UK industry needs to create jobs while becoming increasingly efficient, making it possible to minimise costs and dependencies. From an environmental perspective not only do climate targets need to be met but other priorities exist to reduce local air pollution and improve resource efficiency. This requirement for improved data is recognised in publications such as the Resources and Waste Strategy.[5]

Policies need to be informed by high quality, robust and transparent data. Government is both a key user and provider of data. There is a need to track the various Government programmes to ensure that they are individually and collectively achieving their desired outcomes. There is a requirement for the data to be open, allowing robust scrutiny from government advisory bodies, academics and non-governmental organisations. Data management must be coordinated across departments, and in line with business and industrial aspirations, to catalyse change. Such data also has wider benefits, enhancing engagement of businesses, communities and the public, helping to harness private investment, enhancing transparency and allowing independent scrutiny of decision-making processes, and helping the UK measure progress against international obligations.

The report, jointly produced by the Centre for Research into Energy Demand Solutions (CREDS) and Aether Limited, provides recommendations for a data strategy to deliver this vision.

1.2 Report objectives and approach

This report provides a comprehensive assessment of the current suitability of industrial energy and emissions data in the UK to meet environmental decision-making needs. More specifically the report:

- Explores the shortcomings of the current datasets in meeting the needs of stakeholders to make robust, informed and timely decisions;
- Outlines what a future data strategy would look like to meet these needs, making 10 clear recommendations for realising a successful strategy;
- Examines the benefits to government, business and civil society of this data strategy; and
- Identifies barriers to delivering a strategy.

Our approach to achieve this was to:

- **Conduct one-to-one interviews** with key data users and suppliers, both within and outside of government, to understand their data needs and experience;
- **Undertake desk research** with a systematic review of 68 publicly available and confidential datasets investigating the accessibility and relevance of datasets, and recording key features such as coding systems, classifications used and granularity (among others); and
- **Carry out workshops** with stakeholders from the Department of Business, Energy and Industrial Strategy (BEIS), Department of Environment, Food and Rural Affairs (Defra), Office of National Statistics (ONS) and the Committee on Climate Change (CCC).

Further details on the approach taken and findings from the in-depth analysis of industrial data are provided in Annexes 1 and 2.

2. Industrial energy data: Current status and problems

A successful data strategy needs to be informed by an understanding of the data users' needs (their strategic questions and the type of answers they need) as well as the data providers' ability to produce such data.

2.1 Stakeholder needs

Stakeholders are individuals and organisations that either provide and/or use the data and indicators. Extracts from recent government strategies (Figure 1) suggest the importance of reliable and accessible data to stakeholders.

"We need to make data more available to support processes such as industrial symbiosis"

Waste and Resources Strategy

"We want robust, effective and transparent systems to collect and report data"

Waste and Resources Strategy

"Turning the vision into reality requires solid foundations: comprehensive, reliable data"

25 Year Environment Plan

"Governments must practice what they preach about the opportunities of the data revolution. That means being innovative in how we collect and analyse economic data"

Industrial Strategy

Figure 1: Extracts from recent strategy documents on the role of data in policy.

Stakeholders relevant to industrial energy demand decision-making were identified (Table 1).

Table 1: Industrial energy demand data stakeholders

Stakeholder	Role and interest in industrial energy progress
Department for Business, Energy and Industrial Strategy (BEIS)	Data needed for policy development and tracking as well as international reporting. They collect, prepare and hold some data on policies and schemes, national energy statistics etc.
Department for Environment, Food and Rural Affairs (Defra)	
Office for National Statistics (ONS)	Recommended data holders (see Recommendation 4).
Environment Agency (EA)/Scottish Environment Protection Agency (SEPA)/Natural Resources Wales (NRW)	Data on energy consumption/efficiency could potentially be derived from, and linked to, permitting.
Trade Associations	They would benefit from energy efficiency information, if not already collecting this. They could have a role in providing data, especially on technologies within their sectors.
Committee on Climate Change (CCC)	Role in analysis of policy, and are users of data.
Carbon Trust	As users of data they could help drive investment to improve industrial energy efficiency once areas for improvement have been identified.
Consultancy businesses	Are often subcontracted by government to provide evidence for industrial energy policy.
Academics	Assess improvement potential within industry through modelling exercises and similar. This can form the basis of evidence for, or a critique of, policy.

To develop appropriate trusted messaging for decision-makers, data needs to be transparently condensed and combined into indicators, briefs and other reliable forms of communication. Indicators can be used to track progress and ambition to inform policymakers, business and the public. A list of 'Policy Questions' on industrial energy demand that could be answered through industrial energy demand data or indicators are presented in Box 1. These questions were developed with input from stakeholder discussions, CREDS and other academic interest areas. They were used to help to pinpoint important data sources and indicator needs. A more extensive list of questions and indicators relating to industrial energy demand has been compiled (Annex 3). Many of the indicators are relevant beyond industrial energy demand and could be used to track progress against other strategies. For example, some of the indicators identified as relevant to industrial energy demand could also be applicable to the Resources and Waste Strategy.

Box 1: Policy Questions: Industrial energy demand

What are the most cost-effective mitigation pathways to industrial decarbonisation? (i.e. how would funding be most efficiently allocated?)

- For each sector, how much energy consumption can be electrified/ decarbonised?
- Where is the greatest energy inefficiency in industry? (i.e. where are the greatest opportunities for intervention?)
    - Which industrial activities have the largest contribution to industrial energy consumption?
    - What is the specific energy consumption (SEC: energy use per unit of physical output) for each subsector? How has this changed over time?
- How much emphasis should be on replacing existing approaches, how much for innovative ones?
    - What potential demand reduction could be achieved through energy efficiency improvement to existing industrial processes?
    - Which technologies have the most potential to improve energy efficiency?
    - What is the state of the existing technology 'stock' by sector?
    - What is the potential for different technologies to deliver GHG emissions reductions in the industrial sector and by when?
- Where might intervention create systemic change beyond a specific project?
- What savings could be made by switching to best practice?
- What is the potential increase in energy demand from the uptake of CCS?

Does the combination of different approaches to improve industrial energy efficiency equate to the scale of reduction to achieve carbon budgets and/or net zero emissions?

The current practice of stakeholders largely involves bespoke calls for one-off data collection and/or analysis projects in order to fill evidence gaps. This work often occurs on an ad-hoc basis as the gaps are identified and so is in large part undertaken, in isolation, by consultancies or academics outside of government. The work most often involves the updating or adaptation of existing datasets, using modelling or other techniques, rather than primary data collection. The resulting datasets can be very focussed on a specific policy, may have limited or no transparency, and are not typically updated or easily updateable. This approach is also reactive rather than proactive, meaning there is often no overarching, long-term, or sustainable data strategy in place to maximise the value of the funding available and of the data collection process itself. This can result in higher net costs in the long-term as well as overlap between the data collection efforts of different stakeholders.

Implementing a data strategy allows for more careful design of data collection processes and more useful dissemination approaches. This makes data collection more cost and time effective, creates space for greater innovation and public use of the data that results, and opens opportunities for collaboration and engagement to achieve common policy and strategy goals between organisations.

2.2 Industrial datasets review

Table 2 lists datasets identified as relevant to understanding and addressing industrial energy demand. The initial compilation of datasets included any that could be used, even tangentially, to assess industrial energy demand. It also included those which helped answer the policy questions of the stakeholders presented above (in box 1), by providing useful indicators. The compiled datasets were first assessed on their availability, completeness and quality, and then on how well they aligned with stakeholder needs and the 'Policy Questions'. The full assessment of data is outlined in Annex 2.

Table 2: Datasets identified. A full list is available in Annex 2.

Energy consumption	Pollution/GHG emissions	Production/sales
BEIS Energy and emissions projections Digest of UK Energy Statistics (DUKES), and other energy statistics Environmental Accounts LSOA and MSOA electricity consumption	Committee on Climate Change including Element Energy data and inputs and outputs from the ESME model European Pollutant Release and Transfer Register (E-PRTR) Large Combustion Plants (LCP) database UK National Atmospheric Emissions Inventory (NAEI) EU Emissions Trading System (EU ETS): Annual returns and UK results	UK Manufacturers' Sales by Product (PRODCOM) Eurostat statistics on production of manufactured goods

Our review of currently available datasets and their ability to fill stakeholders' requirements identified several issues that hampered policymaking on industrial energy demand reduction as well as other environmental and low carbon domains. The issues are numerous and complex, but broadly fall within a few key themes (Table 3).

.

Table 3: Issues identified in establishing valued and repeatable data flows on industrial energy demand and efficiency, their potential underlying drivers, and examples from the UK industrial sector

Theme	Specific issues	Drivers	Examples from the UK industrial sector
Lack of access and ownership	Data has limited or no availability to the public, or there is a charge for access.	• Data collected in association with specific government policy, with dissemination or long-term use not considered. • Trade associations may collect data on energy consumption and technologies or processes used within their sector which are not available to the public. • Inexperience of individuals with setting up data ownership agreements and data curation.	• Of the datasets assessed, 52% were fully publicly accessible, 8% were open access at an aggregated level, and 8% had no public access. Of the remainder, 8% of datasets were on a subscription basis, 5% had limited access, 3% were archived, and 14% were of unknown access status.
	Users are unaware of what is available, and awareness is dependent on knowing the right contacts. No central library is maintained.	• Data are not publicised or linked-to from public websites. • Data archiving or gathering from government procured research and projects is not always implemented in a consistent, structured and re-usable manner.	• Assessing the available data on industrial energy demand required extensive desk research given most data were not accessible via a central repository or by simple online searches. • Of those datasets that we could access and evaluate, 42% were accessed via a government website or policy page, whilst 14% were in an open access government data repository. 16% of datasets were accessed via a dashboard tool owned by the data provider, and 9% were held by ONS. The remainder were accessed via other external websites.
Poor Transparency	Recycling of older datasets through successive reports increases the remoteness of the original source, assumptions and uncertainties.	• Lack of up-to-date data. • Limited access to original sources.	• Within the Energy Consumption in the UK (ECUK) dataset energy demand in each sub-sector is split by broad end use (e.g. high temperature processes, motors). This is produced as an annual update. However, this split is based on a single study from 2000[9] with year-on-year variation based on the changes in the fuel demand of sub-sectors rather than a reassessment of end use demand within the sub-sectors.
	Methods, data sources and assumptions behind datasets are not clear, resulting in poor use and low confidence.	• Datasets that are updated annually using many separate data sources and assumptions that are not updated annually. • Reporting by industry often lacks transparency.	• In discussion with public stakeholders who constructed national statistics, we found that there was some reliance on 'expert opinion' in making data assumptions. This leads to a lack of transparency as such assumptions are often not clearly documented, or else are not available in a public format.
	Reliance on datasets with limited accountability/ independent scrutiny, and often of exaggerated quality.	• Lack of a standardised quality management and auditing process in data not collected for statistical purposes.	• The lack of alternatives to datasets with poor transparency mean such datasets are relied upon. These are then consistently cited due to the lack of alternatives and can become accepted as a definitive source without their limitations recognised by secondary users of the data.
Theme	**Specific issues**	**Drivers**	**Examples from the UK industrial sector**

Table 3: Issues identified in establishing valued and repeatable data flows on industrial energy demand and efficiency, their potential underlying drivers, and examples from the UK industrial sector

Little consistency and completeness	Poor consistency across datasets, with use of a range of different classifications, geographical scales, time scales, sector splits and definitions, units and data formats.	• Data collection is designed for a specific purpose only and without a wider understanding of established classifications or nomenclatures. Data collection is undertaken by separate bodies and is not coordinated.	• Approximately half of the datasets evaluated used a national geographic resolution, compared to around a third reporting point location/site level data. • In terms of industrial classifications, various conventions are used between datasets. For example, NACE is used in EUROSTAT datasets, NAEI uses IPCC and NFR codes, whilst many UK based schemes adopt SIC as a standard activity classification such as DUKES, ESOS and ECUK. This may reflect priorities in emissions versus energy data collection however.
	Certain sectors are under-represented, with SMEs and non-energy-intensive industry being prime examples.	• Need to limit cost is a dominant factor. • Data collection is undertaken by separate bodies, and is not coordinated.	• The size of "Other" sectors in, for example DUKES, (and growth over the years). • From an approximate assessment of which sectoral disaggregations are most used by the datasets, it is clear that there is a weighting towards larger scale industries and facilities. Around a third of the datasets considered all sectors, with a slightly smaller proportion interested only in the main industrial sub-sectors. Another third considered a specific single sector, and this was largely dominated by energy industry statistics. From this it is also clear that there is a bias towards data on energy supply rather than energy demand. • The Climate Change Agreement Sector Performance Data perhaps shows the most comprehensive level of sectoral data, with 53 sectors represented. No dataset clearly addressed energy demand in SMEs, although this is a consideration in the consultation on the SECR framework. This considered broadening the threshold for energy and carbon reporting to companies using a minimum of 40,000 kWh per annum, rather than the current criteria of company size.[10]
Poor continuity	Available data are old, or only cover a short time period. Reluctance to provide clear transparent updates to data.	• Data collection is linked to specific policies, so stops when policy ends. • Data collection changes frequently and is not integrated into a single, consistent and sustained process.	• The Usable Energy Database (Box 2) was produced as a snapshot for a specific research project. Despite no recent updates it is believed to be widely used due to a lack of alternatives.
Lacking detail	Data do not have the required granularity for policy analysis or tracking progress with actions. This includes a lack of geographic information.	• Need to limit cost. • Uncertainty in sectoral allocation. • There is concern not to burden industry, which constrains the amount of data which could be collected in connection with support packages.	• No known geographical disaggregation of industrial energy data. A geographical disaggregation at some level can be calculated by processing site level data given in the CCA and EU ETS datasets, but this is limited in sector coverage, fuel disaggregation and end use information. • Within Energy Consumption in the UK (ECUK) energy demand data are available disaggregated into 259 industrial sub-sectors. However, this is only available for 2007, with no more recent update.[11]
	Technological detail is lacking and relies on old datasets (for example the UKERC Usable Energy Database).	• Technology is not the focus of the larger statistical surveys.	• Energy use is annually reported as an aggregate for a sub-sector. There is no known annual release that covers the use of different technologies. Even the volume of currently implemented technologies is uncertain with unknown data held by trade associations.

Box 2: Usable Energy Database (UED): Building an industrial energy database

The UED was constructed by researchers at the University of Bath to improve the representation of the industrial sector within the UK TIMES model (a whole systems cost optimisation model). The database involved two aspects: baseline data on energy use, emissions and output levels for the year 2010 and information on technological options to reduce emissions likely to be available until 2050.

The database was split into a number of subsectors: iron and steel, cement, chemicals, food and drink, paper, steam systems and motors. Subsector specific data sources were used in the UED construction (for both baseline and technology data), including information from industry trade associations for some subsectors. This indicates the lack of detail available with currently available industry wide datasets. The database was funded by the UK Energy Research Centre (UKERC) and is freely available through the UKERC Energy Data Centre.[12]

Due to the nature of the funding the UED represents a snapshot of UK industry and has not been updated since its initial publication. However, attendees at our workshop indicated the value of the UED and the desire for it to be updated.

2.3 Key issues and the need for a data strategy

As outlined in **Table 3**, there are a variety of interlinked driving factors that tend to lead to data which has limited long-term value and that is difficult to use to inform policy. The following outlines the key issues:

- **Differing purposes**. While some data collection and dissemination occur as part of regular 'general purpose' statistical surveys or census, other data collection is often one-off, tailored to more specific purposes or individual policies; for instance when evaluating the uptake of a grant scheme, or informing a modelling exercise. When considering each dataset in isolation, it makes sense to tailor the scope, granularity and classifications used to the specific purpose of the exercise in order to maximise cost-effectiveness and utility. However, this makes it more difficult to make the data more widely useful and/or to maintain a sustainable data flow.
- **Diverse responsible bodies**. Data collection for a specific purpose will often be undertaken by bodies relatively isolated from one another, such as policy teams in different departments. This can lead to duplicated efforts, and inconsistency in the scope, granularity and classifications used.
- **Cost/burden**. Ultimately, data collection and dissemination activities are costly in time and money, both to the bodies collecting the data and to the subjects of the data collection (e.g. industrial facility operators). For this reason, data collection tends to be limited in sample size, scope, granularity and classifications required to achieve the goal. There is no analysis of the longer term added value of a dataset beyond its immediate policy purpose.

The end result is that decision-making systems rely too heavily on individual experts or proprietary datasets and the slow and piecemeal paper publishing cycles of the academic community. It is rare that the data collected behind policies and/or research projects are openly available for other uses and/or users. In many cases, transparent, regular data collection can expose poor decision-making and areas where strategies are not working. It is often easier not to collect data and updates that might expose the failings. Finally, data means power. It therefore makes sense that organisations and sometimes even individuals are not so keen to share it.

Figure 2 outlines the key components within a 'data flow', shown here in the case of industrial energy data. The 'drivers' in a data flow are the external pressures which require improvements to data quality and possibly quantity, including the availability of new funding which demands a better evidence base. 'Inputs' to a data flow may be existing policies and strategies which define what data is collected and how, and what the strategic priorities of a given organisation are. 'Processes' include data collection and compilation, expert elicitation, and ultimately, decision-making. The 'outputs' of a successful data flow can be seen as sources of new evidence, which in turn inform advice and reports, which can then feed into the new strategy and policy of governments or organisations.

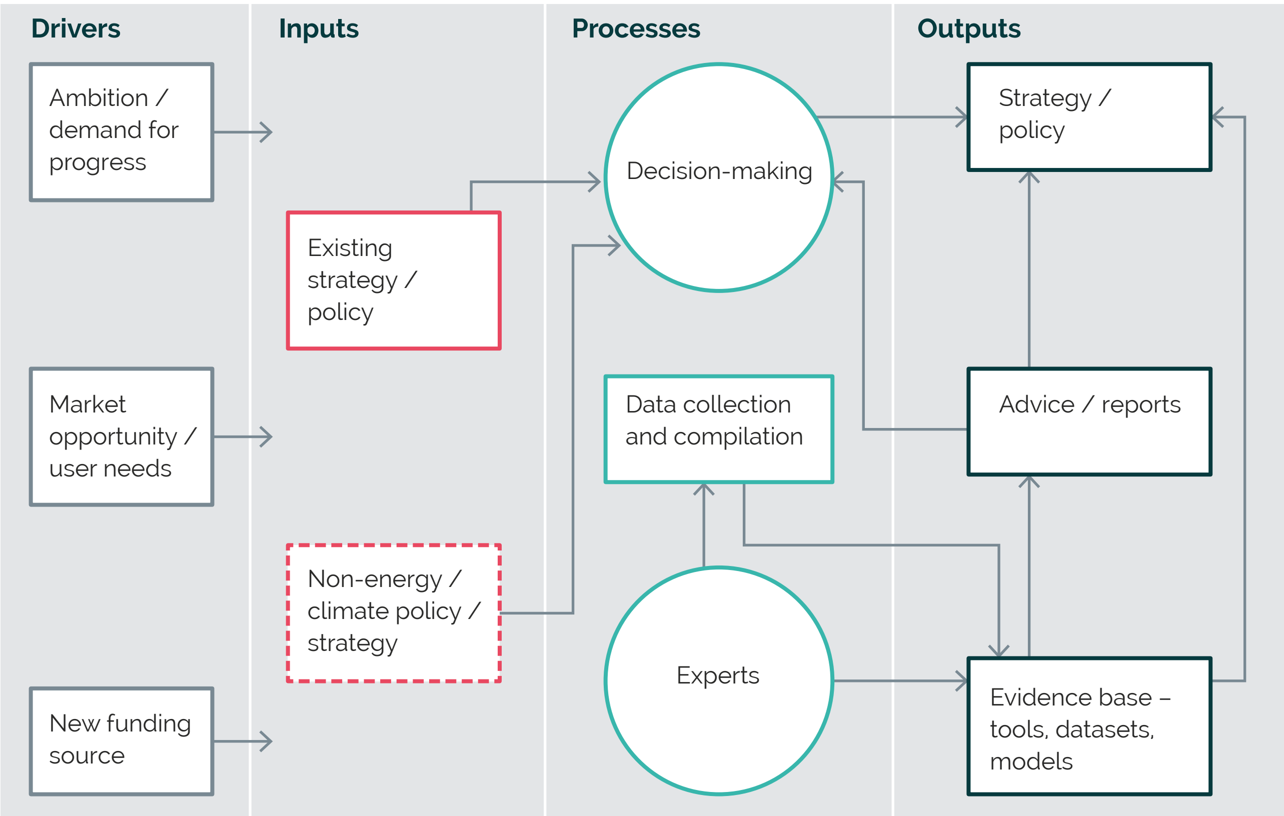

Figure 2: Indicative components in a data flow for industrial energy.

3. Recommendations to maximise data value

The following section lists specific recommendations based on the work undertaken looking at industry energy demand. The work and recommendations also have the potential for wider application to data that can help to navigate a successful path for growth and environmental protection across other sectors of the economy including transport, waste, agriculture and domestic consumption. Many of the recommendations draw on the principles and practices in the Code of Practice for Statistics,[13] published by the UK Statistics Authority for producers of official statistics.

Figure 3 summarises the ten recommendations, split into four different areas, which aim to maximise the value of data collected. Valuable data can be considered to have the following key attributes:

- **Regular and updatable** with well managed changes to methods, data sources and assumptions.
- **Transparent** with well-defined nomenclature, categorisation and taxonomies, clearly documented methods, data sources and assumptions and understanding of uncertainties and limitations.
- **Flexible** and useful in a range of different applications to produce indicators, reports and briefings for multiple audiences across different geographies and timeframes and to provide general and/or specific messages.
- **Accessible** through a central platform alongside other datasets of a similar type and quality, as part of a searchable and easily navigable service linked to dataset taxonomies.

Recommendations for an industrial energy data strategy

Data users and needs

1 Identify stakeholders, their roles and responsibilities, and their data needs

2 Develop a list of strategic data-related questions and indicators

3 Establish a list of required datasets and indicators

Data management

4 Consider appointing a body (such as ONS) to manage open-access datasets

5 Maximise data collection opportunities and consider the long-term value of data beyond a single project

Data quality

6 Ensure consistency of concepts, terminology and classifications between datasets

7 Standardise the presentation of data, documentation on assumptions and uncertainty, and ensure transparency

8 Improve data granularity in the industrial sector, in terms of sub-sector and spatial disaggregations, and technology information

Communicating data

9 Disseminate datasets in a consistent, searchable format

10 Ensure all datasets collected using public funding become open access

Figure 3: Recommendations for an industrial energy data strategy, categorised by theme.

3.1 Data users and needs

Understanding data users and their needs, as undertaken in our analysis, form the first three recommendations.

1. **Develop a list of stakeholders, their roles and responsibilities, and their data needs.**
2. **Develop a list of strategic data related questions and indicators.**
3. **Establish a list of required datasets and indicators.**

3.2 Data management

4. **Consider appointing a central public body (such as ONS) to manage open access datasets**. This would include responsibility for indicator generation and applying appropriate Quality Assurance and Quality Control (QA/QC) rules and dissemination approaches. Relevant public bodies would act as curator for data flows. A non-political public body is preferable. The public body would need to be relatively long-lived, independent, trusted, and experienced in managing data.

A central body controlling the data (and possibly its collection) would facilitate implementing many of the recommendations below by: improving coordination and consistency of data collection and dissemination; being independent of varying political priorities; and avoiding duplication to minimise respondent burden and maximise cost-effectiveness of data collection efforts. ONS produce a variety of statistical publications, disseminated in multiple formats. Developing new multi-format statistical publications focussing on specific areas where data are needed to solve problems (e.g. industrial energy demand and environmental impacts) is one potential option for dissemination, for instance in the form of a new 'annual statistical release'.

5. **Maximise data collection opportunities and consider the long-term value of data beyond a single project**. Data collection is often undertaken for a single dataset linked to a specific policy. Taking a broader view to data collection opportunities can significantly improve its value, summed up by the principle of "collect once, use many times". There may be cases where collecting data at a finer level of granularity than is required for the current application places marginal extra burden on the process but could significantly increase its value. There are currently many missed opportunities to collect data from industry due to a fear of overburdening industry. However, there are many schemes that are awarding funding to industry, and it would be reasonable to require that industries which benefit from these schemes to report more.

   Setting up a regularly updated dataset offers greater value than a single snapshot. The cost of setting up data collection is generally far higher than the cost of maintaining collection, especially if data will be needed again periodically. **When planning new data collection stakeholders should consider its value in tracking progress on policies or informing future policies, and thus the case for ongoing collection**. The UK's National Atmospheric Emissions Inventory (NAEI) (Box 3) is an example of environmental data that is collected and used for multiple reporting activities as well as tracking government policy, supporting business assessments of environmental impact, and informing the public. To ensure ongoing collection, the data collection could be incorporated into another related data flow, which a public body would be well suited to carry out. An example of this process is with the UK's official carbon footprint statistics, produced by the University of Leeds and supplied annually to Defra. Yearly releases incorporate methodological updates and improvements. The University of Leeds has been contracted to produce this, establishing a consistent and publicly available data flow.[14] Once a data flow is established it is harder to discontinue, especially if it becomes an annual statistical publication.

   The value of collected data may also be increased by considering the applications of its indicators beyond the planned one. Indicators which highlight the benefits of a particular course of action for the economy, health, transport, or other environmental good, will help shape cross-cutting engagement. Examples of cross-cutting policies where energy consumption data could contribute include:

- National Strategies: the Industrial Strategy aims to establish the world's first net-zero carbon industrial cluster by 2040[2]. Industrial energy efficiency could be an important factor in realising this.
- UK Sustainable Development Goals (SDGs):[15] goal 9 (to build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation) is set using an indicator of CO2 emissions per unit of value added. Improved energy efficiency could contribute to reducing CO2 emissions.

Box 3: UK National Atmospheric Emissions Inventory (NAEI): An example of good data flow

The UK NAEI has compiled an annually updated time series of statistics on the emissions to and removals from the atmosphere of greenhouse gases and air pollutants. This well-established data flow uses a range of national statistics, reported data from regulated industry, international emission coefficients and expert judgement to estimate a complete picture of emissions and removals from 1990 to the latest possible year (N-2).

This data flow is used in national statistics, informs the calculation of some industrial release estimates, informs on progress with meeting emissions reduction targets and engages a range of scientific, data providing and decision-making stakeholders around action and progress. It also forms a strong foundation and starting point for research into policy impacts and projection scenarios.

The datasets are compiled according to agreed standard methodologies and good practice rules. The data are compiled into internationally agreed nomenclatures but can be converted to and presented in other nomenclatures as well. International reporting requirements ensure that the data are transparent and provide sufficient numerical and descriptive detail on the uncertainties, methods, data sources and assumptions used to compile them. The data undergoes a range of internal audits and a peer review process and is subject to a programme of continuous improvement. This ensures that the usefulness of the data for decision-making can be understood.

The data flow for the NAEI is designed to serve a range of different purposes and is designed to offer a multipurpose tool to support policy development and provide advice. It also is able to help meet some of the UKs specific reporting requirements on air pollution and climate change.

3.3 Data quality

6. **Ensure consistency of concepts, terminology and classifications between datasets**. It is important for data to be collected under a harmonised system of classification, in terms of the definitions and terminology used, geographic codes, and categorisations for variables such as economic sectors, goods, activities, energy and technology. This would improve consistency across departments and allow for datasets to be more easily combined. The standard format would also have to be applied to external contractors. A good example of how the standard classifications can be harmonised across many data providers is provided by Eurostat (Box 4). There are a number of standard classifications (e.g. Standard Industrial Classification; SIC) that can be used. However, it would also be necessary to introduce some new classifications and taxonomies to help define fields such as equipment types, energy efficiency measures, energy sources and technologies. An example classification system developed for industrial energy demand is shown in Figure 4. The classification system would also have to be held by one body which would be in charge of its creation with input from other relevant departments, as well as updating the classifications with new values. Implementing a classification across departments would have challenges since fields within the classification system would have to be agreed. This could be difficult as fields required by different departments may not necessarily align. In addition, for data that need to be reported to larger organisations, such as the NAEI and European Pollutant Release and Transfer Register (E-PRTR), the classification system would have to align enough to be mapped to external classifications.

Equipment	Measures	Energy source	Industry
Energy consuming equipment activities: **Cooling** • Freezing • Refrigeration • Space **Heating** • Drying • Space • Water & liquids **Lighting** **Physical machinery** • Motors • Pumps • Other	Energy efficiency measures: **LEDs** **Temperature** • Insulation • More efficient heating • More efficient cooling	**Electricity** **Fossil fuel** • Coal • Gas • Oil • Other **Other** **Renewable** • Biomass • Hydro • Other • Solar • Wind	**Energy production** • CHP • Electricity • Other **Mining & quarrying** • Metals • Minerals **Production** • Animal &vegetable products • Ceramics • Chemicals • Metals • Other • Textiles • Wood & paper

Figure 4. Example classification developed for this project.

7. Standardise the presentation of data, documentation on assumptions and uncertainty and ensure transparency. Adopting a standard format could make data and QA/QC more transparent. Establishing standards for data and data fields would be advantageous. This could follow national and international standards such as the INSPIRE (European Commission) and ONS national statistical data standards. Documentation should also be standardised to ensure transparency in the assumptions and uncertainties of datasets. Greater transparency will allow data to be questioned and improved more efficiently. Documentation standards would also have to be adopted by external contractors, with new contracts for data collection requiring a clear outline of where there are uncertainties in the data to be provided. An example of a dataset with documentation on the QA/QC process is the UK Environmental Accounts (Box 5).

Box 4: EUROSTAT: Harmonisation across multiple datasets and national bodies

Eurostat – the Statistical Office of the European Commission (EC) – is the central focal point of the European Statistical System (ESS). In this system, national statistical institutes (e.g. ONS in the UK) and other national authorities (such as other government departments) collect and process data across a multitude of topics, for a variety of purposes.

Eurostat works with national bodies to enable them to supply data in a harmonised format, in order to publish consistent European-level datasets. Among other activities, Eurostat's role is to maximise the harmonisation across Member States and themes by:

- Providing standard classification schemes and nomenclatures for reporting, and mapping from other international classifications to aid consistent use of these;
- Outlining specifications for data collection, along with methodological guidance on how to do so;
- Implementing QA/QC checks on the data, to ensure data meets certain standards;
- Requesting and publishing standardised metadata.

Additionally, where data supply is stipulated by an EC directive, supplying data to Eurostat provides national authorities with a mandate for data collection, if it would otherwise not be collected.

8. **Improve data granularity in the industrial sector, in terms of sub-sector and spatial disaggregations, and technology information**. Analysis of the datasets against the 'Policy Questions' and 'Indicators' showed that industrial energy data currently lacks detail to effectively inform policy. Gaps included missing technology and process information, for instance on both what is currently used and what would constitute 'best practice'. A list of technologies and processes used within each sector and sub-sector would be highly beneficial.

For comprehensive assessments of industrial energy demand, data should be able to be split by sector, technology, product, fuel and geography. For example, in the iron and steel sector it would be necessary to be able to break down data to a level where relevant interventions can be identified; this means to the individual process and technology level, for instance at the blast furnace and electric arc furnace level. In order to be effective there would need to be consistency in levels of disaggregation across datasets, something which the dataset analysis indicated does not occur. Datasets often had different geographical and sectoral splits, making comparison and combination difficult. A classification system could be used to improve consistency across datasets. For some sectors site-level data may be valuable. An example of a dataset with flexible granularity is the UK Environmental Accounts (Box 5). Data on smaller energy users, SMEs and non-energy-intensive sub-sectors is also lacking from currently available datasets. In order to decarbonise rapidly and at scale, an understanding of all energy users is necessary, and this is currently a clear gap in the data.

Box 5: UK Environmental Accounts: Flexible data granularity and transparent QA/QC documentation

The UK Environmental Accounts are a series of wide-ranging datasets produced by the ONS as 'satellite' to the UK National Accounts.[16] They are mostly released on a year minus 2 basis, and are designed to cover the intersections between economic activity and environmental impacts. The Accounts include natural asset, physical flow, and economic activity data, providing among others the following indicators:

- Natural assets: land consumption, ecosystem accounts, oil and gas reserves, natural capital flows, forestry resources;
- Physical flows: atmospheric emissions, energy use, material flow accounts, emissions intensity;
- Economic activity: environmental taxes, environmental goods and services sector, environmental protection expenditure, renewable energy sector.

Data are provided at various levels of granularity and disaggregation; for example, energy use data is available by industry reallocated to final consumers, by source, and by fuel type.

The Accounts are produced in accordance with the System of Environmental Economic Accounting (SSEA) and undergo rigorous quality control and assurance processes as part of the suite of ONS publications. These processes are also well documented and publicly available.[17] However, they are often compiled using third party data which can be proprietary, which may have implications for transparency, and entrench reliance on expert judgement.

3.4 Communicating data

9. **Disseminate datasets in a consistent, searchable format**. For the analysis of the 68 industrial energy demand data sets there were 15 datasets that are not publicly available, although some of these may possibly be acquired if requested. In order to maximise public awareness of datasets, they should be highly visible online. The file formats and data structure should be as far as possible consistent and up to date. In the ideal case, a single public body would store all related datasets in a single location, and where feasible pre-combined so the data can be accessed via a single user-interface. A good example of this is the NOMIS portal for labour market statistics in the UK (Box 6).

   There may be issues of confidentiality which require access to the data to be controlled. Depending on the degree of sensitivity there are several options. Less sensitive data can be made available through secure online portals such as the Secure Lab of the UK Data service, requiring a registration process and subject to usage restrictions. More sensitive data may need to be kept offline, with access and analysis taking place only within particular locations.

10. **Ensure all datasets collected using public funding become open access**. A key barrier to consistent dissemination of datasets by a centralised body is that often data collected by external contractors can become the property of those contractors, despite public money being used to fund the creation of that dataset. This varies significantly from case to case, but it is highly desirable that going forward all contracts for data supply specify that the government will own the data at the end. This benefits not only government, but also the general public and civil society who would have access to publicly funded data.

Box 6: NOMIS: Coherent dissemination of data from multiple surveys

NOMIS – the UK labour market statistics portal – was created by the Office for National Statistics (ONS) for viewing and downloading data collected from 14 different sources.

This includes the results of large surveys such as the Labour Force Survey and the Annual Survey of Hours and Earnings operated by ONS, as well as census information, and administrative records such as business counts and births, marriages and deaths as provided by other departments.

The portal provides the facility to view and download:

- Pre-defined statistical reports at the national, regional and local level offering a variety of different geographic categorisations;
- A query tool allowing flexible selection, viewing and download of data from individual data sources, using a standardised interface and classifications.

Registration is not necessary, so the datasets are available to all.

4. Benefits of the recommendations

There are a number of benefits from implementing these recommendations for a range of stakeholders, as summarised in Table 8.

The recommendations could provide **continuity and enhance collaborative working** within and between **government departments** in support of good decision-making. Government policy is transient, and civil service staff turnover can be high. A robust data strategy would **embed the institutional processes** and links between departments in a transparent and durable structure, mitigating against reliance on short-term policy focus, knowledge of individual staff members, and informal arrangements. Staff moving between departments would not need to learn a new approach and there would be a common understanding of the data. A shared system across departments would allow for the efficient creation and use of datasets as well as reducing the need to change approach if transferring to a new department. This would also shape a common understanding of the data. Data that are continuously collected and available would allow for more agile decision-making and flexibility to answer ministerial questions. There would be less need for delays while data is collected after a question is raised. A more openly available data system would also contribute to transparent government processes, potentially reducing freedom of information (FOI) requests, and would amount to a statement of open government.

By maximising both data collection opportunities and the value of such data, the costs of data collection could be reduced as the process will be more efficient. It is clear that an understanding of the overlap, trade-offs and potential synergies between different data collections activities is required. For example, collecting data at a slightly finer level of sectoral resolution may seem unnecessary in the context of a specific policy. However, if this means it can also be used for another purpose in place of a second data collection exercise then this will be more efficient overall. This approach of looking more broadly at the data flows, reduces costs for both the bodies commissioning data collection and reporting into data collection activities.

More information and data available could **stimulate commercial innovation and sound business decision-making**; the data could be used to identify gaps in the market and develop new business models. Data linked to energy consumption and reducing energy demand could be used to identify key areas for reducing energy costs. Data available for specific processes could be used for the benchmarking of processes and attract capital investment in new technologies. The benefits of data availability to businesses were previously identified in the 25 Year Environment Plan. The National Industrial Symbiosis Programme is an example of data sharing within industry (Box 7). The recommended data strategy could also **streamline reporting for businesses and therefore reduce the reporting burden**. Reporting can be seen as a reasonable expectation of business benefitting from policy incentives, and could provide further feedback on how to improve the effectiveness of policy delivery. An example of this is the Streamlined Energy and Carbon Reporting (SECR)[10] framework, which is providing another data flow between stakeholders. While the analysis done shows an increase in costs to business compared to the Carbon Reduction Commitment (CRC) Energy Efficiency Scheme,[18] which it is replacing, much more information is being reported and this will be available to multiple government departments and business.

The proposed recommendations will improve the ability of **academics/civil society to extend their applications** and **improve the basis for questioning and scrutiny** of government and industrial decision-making.

Table 8: Benefits of the recommendations to each generic stakeholder category.

User	Benefit of recommendation
Government analyst (including consultants contracted by government)	• Decision-making is informed by a wider evidence base • Improved ability to respond quickly to ministerial needs for evidence • Greater interdepartmental collaboration • Economic efficiency of data collection • Continuous improvement of datasets through clear documentation of assumptions and uncertainties • Continuity and consistency within and between departments • Reduced need for learning processes with staff turnover • Potential for fewer Freedom of Information (FOI) requests with greater public trust in open government
Academic researcher	• Greater confidence in robustness of datasets • Improved ability to scrutinise government decision-making • Improved ability to provide high quality evidence to inform decision-making
Civil society	• Better availability, accessibility, and navigability of data for diverse needs • Greater trust in government decision-making through transparency and a sense of open government
Industry	• Identification of new business opportunities • Ability to improve operational efficiencies • Greater ambition through comparison to industry best practice • Ability to identify areas to reduce energy costs • Streamlined reporting burden • Access to more useful policy incentives designed around their data

Box 7: National Industrial Symbiosis Programme (NISP): Data sharing within Industry

The National Industrial Symbiosis Programme was a voluntary agreement funded by Defra from 2005 to 2014. It aimed to create resource efficiency opportunities across sectors in the UK, in identifying new material flows between businesses. By establishing multilateral networks between companies, the programme saved the use of 7.9 million tonnes of raw materials between 2005 and 2009, as well as creating £131m in savings to members of the scheme.[19]

A critical part of the scheme's delivery was in the creation of the Core Resource for Industrial Symbiosis Practitioners (CRISP), an online portal which acted as central repository for resource data. This allowed mutually beneficial connections to be made between companies. CRISP also enabled standardised data collection and collation across each participating region of the UK, and was considered important in providing a data resource for research beyond the scope of NISP.[20]

NISP members would submit data on material flows within their operations, and companies were assigned a unique identifier. Data were maintained in CRISP aligned with a set of data management guidelines. The data were spatially referenced, and could be disaggregated by the type and quantity of a given resource.[21] Among the metrics adopted in the datasets were those quantifying benefits, including: cost savings, job savings and creation, raw material savings, diversion from landfill, CO_2 reductions, elimination of hazardous waste, water savings.[22]

WRAP funding was withdrawn in 2014, and NISP is now independently managed by International Synergies.[23]

5. Barriers to the recommendations

Barriers to implementing the recommendations include:

- **Perceived additional burden**. While many agree that the current approach is insufficient for the data needs of stakeholders, some stakeholders are content with the current system and felt that it was 'fit for purpose' considering the additional burden of change. There is a perception that mainstreaming previously 'one-off' data collection activities, increasing the detail requested in surveys, or increasing the sample size of existing surveys will result in greater burden on industry. However, if implemented smartly, improved coordination within and between departments could minimise this, and could remove the need to report according to several different conventions. Indeed, there is an argument that existing data collection practices are already burdensome and that there is considerable scope for improvement. The administrative barriers of needing to enlarge and potentially restructure the national entity to accommodate this would be an important consideration, although this is already being considered in the case of SECR.[10]
- **Industry confidentiality concerns**. This was commonly cited in the SECR consultation[10] and could also be a barrier to the implementation of the recommended data system. However, with centralised management of open source datasets as well as a tiered and user-specific system for more sensitive and granular data (such as the ONS's virtual microdata laboratory), this does not represent a substantive barrier.
- **Short policy turnaround and short-term policymaking**. The short turnaround of policy design and the often short-term nature of policymaking also pose barriers. But readily available and easily navigable datasets could adapt to changing demand. Enhanced levels of confidence in dataset reliability and methodological rigor could also improve the quality of, and trust in, decision- and policy-making.
- **Resistance to transparency**. Some organisations would prefer to operate with a lack of public transparency to avoid challenge or scrutiny of their operations and/or their poorly performing actions. However, by introducing new standards of reporting and documentation, this creates a level playing field which establishes a minimum expectation of quality. In time this can be expected to improve practice through

agenda-setting, both in terms of data management and real-world operations.

- **Reluctance to relinquish ad-hoc data collection**. Organisations may be reluctant to lose the flexibility to carry out impromptu data collection. Implementing a data strategy would increase the administrative and financial overheads involved in planning data collection, and may introduce delays into the process. The implementation of a data strategy needs to be carefully thought through, to still allow policy and analytical teams to carry out crucial data collection in a timely and targeted manner. Evaluative frameworks must be in place to consider the costs and benefits of ad-hoc data collection in each particular policy case, as it sometimes may be more appropriate to carry out time-limited, specific data collection. However, as previously discussed, a well-coordinated data strategy has potential to reduce the (costly) need for ad-hoc data collection once established.
- **Economic costs to implementation**. There are both real and perceived costs associated with the recommendations, including establishing the secure and centralised infrastructures for dataset storage, as well as coordination of company reporting. However, implementation of the recommendations could be more financially efficient than existing practice, and would additionally provide scope for industry to identify cost-savings in their own practice, in turn driving productivity. By broadening the appeal of a given data collection exercise to multiple stakeholders and end-users, the costs of data collection are shared, and this diversifies potential funding sources.
- **Current lack of a body responsible for data management**. A key recommendation above is number 4: Consider appointing a central public body (such as ONS) to manage open access datasets. This facilitates the adoption of the other recommendations, without such a body it is difficult to coordinate action across different stakeholders.

6. Conclusion: Data strategy for a low carbon transition

In conclusion we find that the current availability and approach to evidence on industrial energy is highly inadequate. It relies too heavy on individual experts or proprietary datasets that are not open to scrutiny and are updated on a sporadic basis. Decisions are made on out of date analysis that were never designed for that particular purpose. The current approach is not cost effective where individual studies are commissioned to repeat similar analysis for different Government bodies and agencies. There are multiple benefits to be gained with the implementation of a data strategy, and though particularly of relevance to the industrial energy space, there is clearly potential for broader application across other UK sectors. The analysis has shown that current data management practices around industrial energy are inefficient and do not substantively respond to current and future stakeholder data needs. This in turn leads to missed opportunities for the identification of efficiency improvements which would benefit industry and hence the broader economy.

We conclude that there is an urgent need for a data strategy that gathers linked data on emissions, technologies, costs linked to wider datasets on jobs and other environmental impacts.

We highlight that there are multiple benefits to this approach ensuring that Government and industrial funds are spent most effectively, allowing wider scrutiny and ensuring the alignment of multiple government strategies managed across a number of departments.

We conclude that this dataset must be updatable, open access, independently managed ensuring consistency of concept, terminology and classifications.

The need for such a dataset is urgent as the time available to adequately respond to the climate emergency is short. Trade-off exists between having good evidence to make good policies and the cost associated with having good evidence – if you spend time and money to get the evidence right, then the money and time will be spent more effectively.

We conclude that BEIS and Defra should help establish a central public body, with the co-operation of key agencies such as ONS and the CCC, to deliver an improved evidence base for industrial energy, emissions and technologies.

Annex 1. Stakeholder consultations and engagement

A1.1 Stakeholder interviews

One-to-one phone interviews were conducted with 7 relevant stakeholders, including representatives of government agencies including BEIS, Defra, CCC, and ONS, as well as a consultant. The unstructured interview covered 3 key areas, but was directed by stakeholders to determine what the data priorities were for their organisation. The broad topic areas covered included:

- Gauging the satisfaction of stakeholders (both data users and suppliers) with the quality of current data flows;
- Determining the issues in current data use/provision, e.g. transparency, accessibility, coverage;
- Identifying what the ideal state of data would be for the current and future needs of each stakeholder.

The outcomes of these informal interviews provided direction to the planning of the stakeholder workshop, in terms of what would be valuable topics to cover, and which points could be validated against insights from other stakeholders.

A1.2 Stakeholder workshop

The project team organised a stakeholder workshop during January 2019 for 16 external stakeholders, including 11 attendees from BEIS, 2 from the CCC, 1 from Defra, 1 from ONS, 1 from CREDS (Oxford) and 1 affiliate of the University of Leeds. 6 project team members coordinated the workshop. It was decided that the workshop would be oriented to public agency stakeholders in-keeping with the focus of the analysis on data strategy, policy levers for improved data flows and what policy questions these could help answer.

A briefing note on the project was circulated to attendees prior to the event, as well as an agenda. Certain stakeholders were asked to give short presentations to establish the position of their organisation on data quality, availability, and strategy. This provided scope for the resulting discussion. Information was captured anonymously by means of two note-takers, as well as a feedback form issued at the end of the session to all participants.

Questions posed to workshop attendees included the following:

1. How do you use industrial energy data in your analysis?
2. What improvements to existing data would increase its value?
3. Where are the current gaps in data?
4. What would a best practice data strategy for UK industrial energy and resource demand look like?

Copies of both the presentations given and of the workshop findings were issued after the event.

Annex 2. Dataset assessments

A2.1 Datasets evaluated

An initial list of 68 datasets was compiled. This list also included databases that were considered to possibly exist as well as those that were known to. Each dataset was described by the following information:

General information on the dataset:

- An overview of the dataset
  - The data status (Existing data/ data not accessible)
  - File type
  - Availability/charges/location of the data
  - Confidentiality or use restrictions
- Scope and resolution
  - Any information contained about technology or processes
  - The geographical scope and resolution
  - Temporal scope and resolution
  - The activity classification used
  - The scope of energy types
- Information on how the data is gathered
  - Source of data
  - Survey type: how complete is the data collection
  - Respondent type: who the data is being gathered from
  - Body responsible for data collection
  - Data collection frequency
- Quality
  - If QA/QC is applied
  - If uncertainty is quantified
  - Quality issues

The aim was to complete the information laid out above for each dataset. However, for many datasets it was not possible to fully determine all of the required parameters especially information on the quality control of the datasets. There were also many datasets that were not publicly available for which it was not possible to find out the above information either.

Before undertaking further assessment, the list of datasets was refined to make the number more manageable. Of the 68 datasets 31 were excluded from further analysis due to lack of information or not being publicly available and not being relevant enough to energy efficiency (Table A2.1).

Table A2.1: Initial dataset list and reasons for exclusion from further analysis.

Datasets	Exclusion criteria
A2 / Part B Permit Reporting	
Aluminium Institute	
Annual Purchases Survey	Doesn't correlate to an efficiency indicator.
BEIS Energy and Emissions Projections	
BEIS energy technologies installations (subnational datasets (electricity and gas)	Insufficient information/access.
BEIS Energy Trends	
Building Energy Efficiency Survey	Insufficient information/access.
Cambridge Econometrics	Insufficient information/access.
Climate Change Agreements	
Committee on Climate Change	
Climate Change Levy	Doesn't correlate to an efficiency indicator.
CHP Quality Assurance Program	Insufficient information/access.
Chemical Industries Association	Insufficient information/access.
Coal Authority	
Coal Statistics	
CRC Energy Efficiency Scheme	
DECC Electricity Surveys	Production-based indicator.
Downstream annual gas survey	
DUKES	
Environment Agency	Doesn't correlate to an efficiency indicator.
Electrical Contractors' Association	Insufficient information/access.
Economy based IDBR	Doesn't correlate to an efficiency indicator.
ECUK	
Electricity Statistics (same as DUKES, Energy Trends)	Production-based indicator.
Energy Demand Research Project	Insufficient information/access.
ENUSIM model	Insufficient information/access.
Environmental Accounts	
E-PRTR	
ESOS	Doesn't correlate to an efficiency indicator.
EU ETS: Annual Returns	

Table A2.1: Initial dataset list and reasons for exclusion from further analysis.

Datasets	Exclusion criteria
EU ETS: UK allocations	Doesn't correlate to an efficiency indicator.
EU ETS: UK results	
EUROSTAT Annual energy balances (nrg_10a)	
EUROSTAT Monthly energy statistics (nrg_10m)	
EUROSTAT Statistics on production of manufactured goods	
Gas statistics	
Gas suppliers	Insufficient information/access.
HMRC/Treasury	Doesn't correlate to an efficiency indicator.
International Energy Agency	
International Steel Statistics Bureau	
LCP Database	
LSOA/MSOA electricity consumption	
Market transformation programme	
Mineral Statistics Yearbook	
Mineral Products Association	Insufficient information/access.
NAEI	
Non-domestic EPCs	Insufficient information/access.
Non-domestic National EE Data-Framework	Insufficient information/access.
Non-domestic Smart Meters	
Natural Resources Wales	Doesn't correlate to an efficiency indicator.
ODYSEE-MURE	
Ofgem – FITs	Doesn't correlate to an efficiency indicator.
Oil statistics	
ONS Labour Market Survey	Doesn't correlate to an efficiency indicator.
Platts Market Data	Doesn't correlate to an efficiency indicator.
Pollution Inventory	
PRODCOM (Eurostat)	
RDCO (oil purchasing)	Production-based indicator.
Renewables statistics (RESTATS)	Production-based indicator.
Non-domestic Renewable Heat Incentive (RHI)	
Subnational electricity and gas consumption	
Total final energy consumption – subnational	
Other trade associations	Insufficient information/access.
UK offshore operators association	Production-based indicator.
UK PIA, downstream oil sector	
UKERC Usable Energy Database	
Wales Flexis	Insufficient information/access.

A2.2 Data Assessment:

The next step was to assess the suitability of each dataset for answering a range of existing and potential future user needs (see Annex 3). In order to assess the suitability of the datasets against the data needs identified, firstly a list of indicators for energy demand was put together (see Annex 3). These indicators represent a description of the quantity in question (e.g. energy use, value-added, physical production), scope and breakdown of data that would be required to address particular needs or questions. Each indicator was mapped to the list of questions for which it is relevant. In a separate step, the indicators were also mapped to the datasets containing suitable data. By using the indicators as an intermediary, the datasets could then be matched to the questions that they could be used to help answer.[1]

A2.2.1 Matching Indicators to datasets

While some indicators have a high number of datasets underlying them, such as total sub-sectoral energy consumption which has 12 datasets, the majority have under 5 and significant portion are covered by only one or no datasets (table A2.2). While this does not show how well each dataset could be used to provide the indicator it does highlight that there are many gaps in the data currently available, in particular information on technology. The dataset identified as the most valuable and broad for information on technology was the UKERC Usable Energy Database and while it has fairly high sectoral coverage it is a one-off dataset and therefore out of date. It is also worth highlighting that this database was based on literature that was older than the database itself. This is especially important as a few stakeholders during the workshop indicated that it was still widely used and that a new version would be useful.

Table A2.2: The number of datasets from which an indicator can be derived. Indicators that can be derived from no or only one dataset are highlighted.

Indicator		No. of relevant datasets
Energy consumption	Total economy energy consumption	2
	Total economy energy consumption by source	4
	Total industry energy consumption (absolute/share of TFC)	6
	Total industry energy consumption by energy source	7
	Total sub-sectoral energy consumption (absolute/share of industry consumption)	12
	Total sub-sectoral energy consumption (absolute/share of industry consumption) by source	8
	Energy consumption of cross-cutting processes	0
	Energy consumption of sub-sector processes	1
	Energy consumption of sub-sector processes by source	1
	Fuel intensity	0

1 Full details of the characteristics assessed for each dataset, and of the mapping between datasets, indicators and user needs / questions is available from the authors on request.

Table A2.2: The number of datasets from which an indicator can be derived. Indicators that can be derived from no or only one dataset are highlighted.

Indicator		No. of relevant datasets
Industrial activity, value added	Total industry value added	1
	Total industry energy consumption per total industry value added (energy intensity)	1
	Sub-sectoral energy consumption per sub-sectoral value added (energy intensity)	1
	Energy consumption/GHG emissions per value added for each sub-sector process/product	0
	Industrial sub-sector value added	2
Industrial activity, physical output	Total industry physical output	0
	Total industry energy consumption per unit of physical output (energy intensity)	1
	Total industry GHG emissions per unit of physical output (emissions intensity)	0
	Total sub-sectoral industrial physical output	10
	Sub-sectoral energy consumption per unit of physical output (energy intensity)	3
	Sub-sectoral GHG emissions per unit of physical output (emissions intensity)	2
	Specific energy consumption (SEC) (per tonne produced)	1
Emissions	Total economy GHG emissions	3
	Total industry GHG emissions	5
	Total industrial sub-sector GHG emissions	11
	GHG emissions of sub-sector processes	2
	GHG emissions of sub-sector products/equipment	0
	GHG emissions by facility	3
	Direct/indirect/process emissions by industrial sub-sector	1
Technology	Energy consumption of sub-sectoral products/equipment	3
	Per unit equipment energy intensity	0
	Production volume of energy efficient equipment	1
	Sales of energy efficient industrial equipment	1
	Technology penetration rates	1
	Installation frequency of efficiency equipment	4
	Sales by product	2
	Capital cost of energy efficient equipment	1
	Efficiency improvement potential of technology	0

A2.2.2 Matching datasets to questions

Using the mappings created between indicators and datasets, and between questions and indicators, the availability of datasets to address user needs was analysed (table A2.3). Excluding the qualitative questions there is a high number of datasets that could be applied to each question, with the lowest being three for "How does industrial energy demand affect trade, employment and economic output?". However, this is likely to be because it wasn't the main focus of the project and in the initial stages the team had not set out to gather demographic and economic data.

Table A2.3: Matching of the datasets with questions		
Parent Question	**Sub-Question**	**No. of relevant datasets**
What is the potential demand reduction through energy efficiency improvement in industry?	What potential demand reduction could be achieved through energy efficiency improvement to existing industrial processes?	12
	What potential demand reduction could be achieved through existing energy efficiency technologies?	7
	What potential demand reduction could be achieved through new energy efficiency technologies?	7
	What is the demand reduction potential of industrial energy efficiency technologies by Technology Readiness Level?	7
In which industrial activities would funding to improve energy efficiency be most cost-effective?	Which industrial activities have the largest contribution to industrial energy consumption?	16
	What is the specific energy consumption (SEC – energy use per unit of physical output) for each sub-sector, and how has this changed over time?	4
	What is the most cost-effective pathway for industry GHG mitigation?	4
Which technologies have the most potential to improve energy efficiency?	What is the state of the existing technology 'fleet' by industrial sub-sector?	7
	What is the potential for different technologies to deliver GHG emission reductions in industrial sub-sectors, and by when?	7
	What data is needed to assess a capital allowance scheme for industrial energy efficiency? (Qualitative)	0
How much energy consumption can be decarbonised by industrial sub-sector?	What is the potential for fuel switching (e.g. electrification/hydrogen/biomass) in industrial sub-sectors by fuel type?	8
	What net change in industrial energy consumption can be anticipated from high electrification scenarios?	12
	How could fuel switching affect national fuel demand levels?	14
What is international best practice on industrial energy efficiency?	What is the UK's industrial energy efficiency?	19
	What energy savings could be made by switching to international best practice efficiency?	26
	What data collection practices are used to enable/monitor international energy efficiency schemes? (Qualitative)	0
	What policies underpin international best practice energy efficiencies? (Qualitative)	0
CCS applications	What would be the net change in energy demand as a result of high CCS implementation?	17
How is industrial energy demand linked to the broader national economy?	How does industrial energy demand affect employment levels?	3
	How does industrial energy demand affect economic output?	3
	How does industrial energy demand and reduction policy affect trade?	3

However, simply comparing the numbers of datasets relevant to one question versus another is potentially misleading. The relevance or completeness of individual datasets in supplying an indicator can vary (i.e. can a dataset fully or only partly supply the required indicator on its own), as well as the degree of overlap or complementarity between different datasets relevant to the same question. These are aspects which are difficult to quantify without detailed examination of the data themselves, which has not yet been undertaken.

Nevertheless, the mapping of relevant datasets undertaken so far does serve to:

1. Highlight evidence gaps, where zero or very few have been identified; and
2. Direct efforts to investigate datasets relevant to key questions in more detail.

Digging deeper using for example the first question "What potential demand reduction could be achieved through energy efficiency improvement to existing industrial processes?." There are 12 datasets that have been identified as able to contribute to answering the question. However, while some provide some sectoral breakdown (environmental accounts and EUROSTAT tables) other do not (IEA, which provides an overall UK result, and UKPIA which is for oil refining only). Also, none of the datasets are broken down to the process level. This lack of complete sectoral and process breakdown will hinder the ability the answer the question using the available datasets. It should also be noted that there is repetition between some of these datasets for example the ECUK data, Gas Statistics and Coal Statistics source their overlapping data from the same sources and the EUROSTAT tables are the same data but with different time resolution (monthly and annual).

Annex 3. Policy questions, indicators & datasets identified

Table A3.1: Policy Question list.

Hierarchy	Question
1.0	What are the most cost-effective mitigation pathways to industrial decarbonisation? (i.e. how to most efficiently allocate funding?
1.1	For each sector, how much energy consumption can be electrified / decarbonised?
1.1.1	What is the potential for fuel switching to decarbonised electricity, hydrogen etc.?
1.1.1.1	What net change in energy consumption can be anticipated from high electrification (e.g. of heat) scenarios within industry?
1.1.1.2	How could fuel switching affect national fuel demand levels?
1.2	Where is the greatest energy inefficiency in industry? I.e. where are the greatest opportunities for intervention?
1.2.1	Which industrial activities have the largest contribution to industrial energy consumption?
1.2.2	What is the specific energy consumption (SEC – energy use per unit of physical output) for each sub-sector? How has this changed over time?
1.3	How much emphasis should be on replacing existing approaches, how much for new/innovation
1.3.1	What potential demand reduction could be achieved through energy efficiency improvement to existing industrial processes?
1.3.2	Which technologies have the most potential to improve energy efficiency?
1.3.2.1	What potential demand reduction could be achieved through existing energy efficiency technologies?
1.3.2.2	What potential demand reduction could be achieved through new energy efficiency technologies?
1.3.2.3	What is the demand reduction potential of industrial energy efficiency technologies by Technology Readiness Level?
1.3.2.3	What is the energy demand by technologies within sub-sectors?
1.3.2.4	What is the potential for energy demand reduction in cross-cutting technologies within the non-energy-intensive sub-sectors of industry?

Table A3.1: Policy Question list.

Hierarchy	Question
1.3.3	What is the state of the existing technology "fleet" by sector?
1.3.4	What is the potential for different technologies to deliver GHG emission reduction in the industrial sector and by when?
1.4	Where might intervention create systemic change beyond a specific project?
1.5	What savings could be made by switching to best practice?
1.5.2	What are international best practice energy efficiencies?
1.5.1	What is the UK's industrial energy efficiency?
1.6	What is the potential increase in energy demand from the uptake of CCS?
2.0	Does the collection of different approaches to improve industrial energy efficiency equate the scale of reduction to achieve the carbon budgets?

Table A3.2: Indicator list.

Indicator ID	Indicator Group	Indicator
1.1	Energy Consumption	Total economy energy consumption
1.1.1	Energy Consumption	Total energy consumption by source
1.1.2	Energy Consumption	Total industry consumption
1.1.2.1	Energy Consumption	Total industry energy consumption by energy source
1.1.2.2	Energy Consumption	Total sub-sectoral energy consumption
1.1.2.2.1	Energy Consumption	Total sub-sectoral energy consumption by source
1.1.2.3	Energy Consumption	Total consumption of cross-cutting processes
1.1.2.2.2	Energy Consumption	Energy consumption of sub sector processes
1.1.2.2.3	Energy Consumption	Energy consumption of sub sector processes by source
1.2	Energy Consumption	Fuel intensity
2.1	Industrial Activity	Value added
2.1.1	Industrial Activity	Total industry value added
2.1.1.1	Industrial Activity	Total industry energy intensity
2.1.1.1.1	Industrial Activity	Sub-sectoral energy intensity
2.1.2	Industrial Activity	Industrial sub-sector value added
2.1.2.1	Industrial Activity	Energy consumption/GHG emissions per value added for each sub sector process/product
2.2	Industrial Activity	Physical output
2.2.1	Industrial Activity	Total industry physical output
2.2.1.1	Industrial Activity	Total industry energy intensity
2.2.1.2	Industrial Activity	Total industry emissions intensity
2.2.1.3	Industrial Activity	Total sub-sectoral industrial physical output

Table A3.2: Indicator list.

Indicator ID	Indicator Group	Indicator
2.2.1.3.1	Industrial Activity	Sub-sectoral energy intensity
2.2.1.3.2	Industrial Activity	Sub-sectoral emissions intensity
2.2.1.3.3	Industrial Activity	Specific energy consumption
3.1	Emissions	Total economy GHG emissions
3.1.1	Emissions	Total industry GHG emissions
3.1.1.1	Emissions	Total industry sub-sector GHG emissions
3.1.1.1.1	Emissions	Direct/indirect/process emissions by industrial sub-sector
3.1.1.1.2	Emissions	GHG emissions of sub-sector processes
3.1.1.1.3	Emissions	GHG emissions of sub-sector products/equipment
3.1.1.1.4	Emissions	GHG emissions by facility
4.1	Technology	Energy consumption of sub-sectoral products/ equipment
4.3.1	Technology	Sales of energy efficiency industrial equipment
4.3	Technology	Per unit equipment energy intensity
4.2	Technology	Efficiency improvement potential of technology
4.4	Technology	Technology penetration rates
4.3.2	Technology	Production volume of energy efficient equipment
4.3.3	Technology	Installation frequency of efficiency equipment
4.5	Technology	Sales by product
4.3.4	Technology	Capital Cost of energy efficient technology

Annex 4: Project team

This report was jointly produced by the Centre for research into Energy Demand Solutions (CREDS) and Aether Limited. CREDS is a UKRI funded research centre established in 2018, involving more than 80 academics across 13 UK academic institutions. The centre aims to understand the role of energy demand in achieving a low carbon transition, with 3 sector and 3 cross-cutting themes (buildings, materials and products, transport and mobility; and digital society, flexibility, and policy and governance, respectively), as well as specific challenges on the decarbonisation of heat and fuel and transport poverty. The material and products theme is led by Professor John Barrett at the University of Leeds, considering the role of industrial energy demand, energy efficiency, and the energy embodied in materials and products.

Aether Limited provides world experts in environmental data analysis and interpretation, and is at the forefront of greenhouse gas (GHG) and air quality pollutant emissions calculation and review. Most of Aether's portfolio involves working with international organisations and delivering to national level Governments, but Aether also undertakes work at regional and local scales. Aether's work is motivated by a strong drive to increase the quality, transparency and accessibility of emissions and other environmental data, which is vital to underpin climate and air quality mitigation policies and measures.

The project team included:

CREDS/University of Leeds

- **Jonathan Norman**: Jonathan is a Senior Research Fellow in the Sustainable Research Institute at the University of Leeds, and part of the CREDS Materials and Products theme. He previously worked on industrial energy demand reduction as part of CIE-MAP at the University of Bath, and has a background in Mechanical Engineering with previous experience in industry.

- **John Barrett**: John leads the CREDS Materials and Products theme and is a Professor in Energy and Climate Policy in the Sustainability Research Institute at the University of Leeds. John was the Director of the Centre for Industrial Energy, Materials and Products (CIE-MAP) and a Co-Director of the UK Energy Research Centre.
- **Alice Garvey**: Alice is a Research Assistant in energy demand at the University of Leeds as part of the CREDS 'materials and products' theme, with a broad research interest in low carbon transitions.
- **Peter Taylor**: Peter is Chair in Sustainable Energy Systems at the University of Leeds and a member of a number of high-profile RCUK research centres including CREDS, the UK Energy Research Centre, the Centre for Climate Change Economics and Policy and the Supergen Energy Networks Hub.

Aether Limited

- **Justin Goodwin**: Justin is one of the four founding Directors at Aether, with over 20 years' experience in GHG emissions inventory compilation and review, and in developing national monitoring, reporting and verification systems.
- **Mark Gibbs**: Mark is a Principal Consultant at Aether, specialising in improving industrial emissions reporting and data flows, following 12 years working for the Oklahoma Department of Environmental Quality.
- **Richard German**: Richard is a Senior Consultant at Aether, specialising in data analysis, database design and visualisation.
- **Lucy Garland:** Lucy is an Assistant Consultant at Aether, with a strong interest and expertise in emissions reporting in industry and the power sector.

References

1 House of Commons Library. 2019. Legislating for net zero – Briefing paper CBP8590. London: House of Commons Library.

2 BEIS. 2017. Industrial Strategy: Building a Britain fit for the future. London: Crown Copyright.

3 BEIS. 2017. The Clean Growth Strategy: Leading the way to a low carbon future. London: Crown Copyright.

4 Defra and EA. 2018. Our waste, our resources: A strategy for England. London: Crown Copyright.

5 Defra. 2018. A Green Future: Our 25 Year Plan to Improve the Environment. London: Crown Copyright.

6 Innovate UK. 2019. Industrial Strategy Challenge Fund – wave 3 shortlist. London: Crown Copyright.

7 Innovate UK, BEIS, and UK Research and Innovation. 2019. Industrial Strategy Challenge Fund: for research and innovation. London: Crown Copyright.

8 BEIS. 2019. Creating a Clean Steel Fund: call for evidence. London: Crown Copyright.

9 Pout, C. H., MacKenzie, F., and Bettle, R. 2002. Carbon dioxide emissions from non-domestic buildings: 2000 and beyond. Watford, UK: Buildings Research Establishment.

10 BEIS. 2018. Streamlined Energy and Carbon Reporting: Government response. London: Crown Copyright.

11 BEIS. 2019. Energy Consumption in the UK. London: Crown Copyright.

12 UK Energy Research Centre. 2019. Energy Data Centre. London: UK Energy Research Centre.

13 UK Statistics Authority. 2020. Code of practice for statistics. London: UK Statistics Authority.

14 Defra. 2019. Official Statistics: UK's Carbon Footprint. London: Crown Copyright.

15 Office for National Statistics. 2019. UK data for Sustainable Development Goal indicators. London: Crown Copyright.

16 Office for National Statistics. 2019. UK Environmental Accounts. London: Crown Copyright.

17 Office for National Statistics. 2019. Environmental accounts on energy use QMI. London: Crown Copyright.

18 BEIS. 2015. CRC Energy Efficiency Scheme. London: Crown Copyright.

19 European Commission. 2009. Waste Prevention Best Practice Factsheets: National Industrial Symbiosis Programme (UK). Brussels, Belgium: European Commission.

20 Mirata, M. 2004. Experiences from early stages of a national industrial symbiosis programme in the UK: determinants and coordination challenges. *Journal of Cleaner Production*, **12** (8–10): 967–983. doi: 10.1016/j.jclepro.2004.02.031

21 Jensen, P.D., Basson, L., Hellawell, E.E., Bailey, M.R. and Leach, M. 2011. Quantifying 'geographic proximity': Experiences from the United Kingdom's National Industrial Symbiosis Programme. *Resources, Conservation and Recycling*, **55** (7): 703–712. doi: 10.1016/j.resconrec.2011.02.003

22 Cecelja, F., Raafat, T., Trokanas, N., Innes, S., Smith, M., Yang, A., Zorgios, Y., Korkofygas, A. and Kokossis, A. 2015. e-Symbiosis: technology-enabled support for Industrial Symbiosis targeting Small and Medium Enterprises and innovation. *Journal of Cleaner Production*, **98** (July 2015): 336–352. doi: 10.1016/j.jclepro.2014.08.051

23 International Synergies. 2019. Delivering industrial symbiosis in England. Birmingham: International Synergies.

About CREDS

The Centre for Research into Energy Demand Solutions (CREDS) was established as part of the UK Research and Innovation's Energy Programme in April 2018, with funding of £19.5M over 5 years. Its mission is to make the UK a leader in understanding the changes in energy demand needed for the transition to a secure and affordable, low carbon energy system. CREDS has a team of over 100 people based at 15 UK universities.

CREDS is funded by UK Research and Innovation, Grant agreement number EP/R035288/1

About Aether

Aether was founded in 2008 by senior members of the UK's national emission inventory team. The initial team of four has now grown to more than 20, based in Oxford.

By transforming complex emissions data into robust emissions inventories Aether can deliver demonstrable environmental intelligence that can be used to inform policy decisions. Policy makers are then able to make informed, evidence based, environmental and economic decisions towards helping to reach emission reduction targets.

ISBN: 978-1-913299-02-6

CREDSadmin@ouce.ox.ac.uk

www.creds.ac.uk

@CREDS_UK

www.linkedin.com/company/credsuk/

UK Research and Innovation

Engineering and Physical Sciences Research Council

Economic and Social Research Council