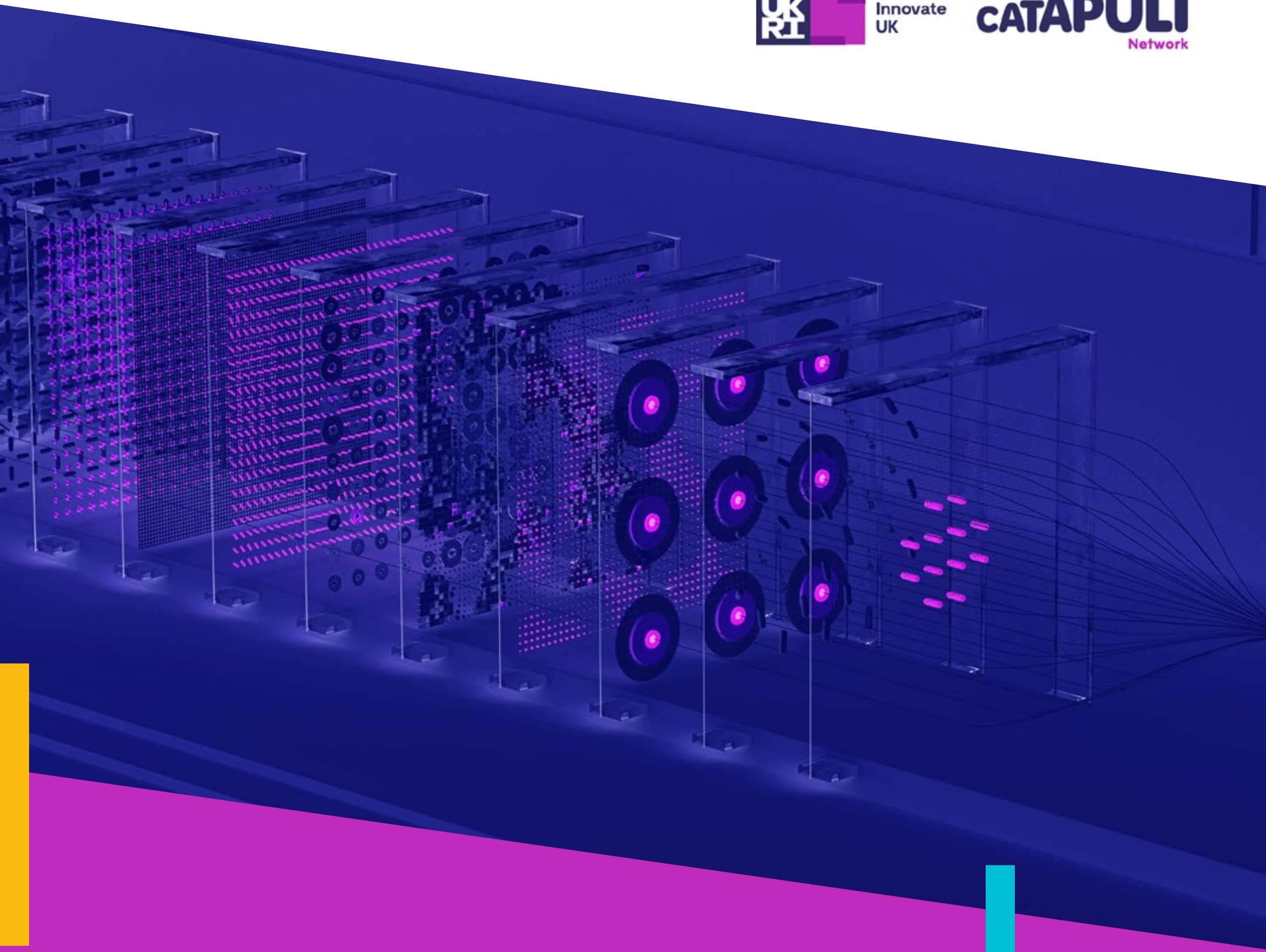


CARBON ACCOUNTING: RECOMMENDATIONS FOR DATA MANAGEMENT IN UK MANUFACTURING

AI Data Team, Digital Catapult

March 2025

CONTENTS

Document User Guide

Start of section

Previous page

Contents page

Next page

CONTENTS

SUMMARY OF RECOMMENDATIONS

Digital Catapult CCCA data framework recommendations		
Data challenges	Recommendations through the data management framework	Actions to be taken through process framework and regulations
No standardised data format for sharing carbon accounting information	Having specified solutions and formats so that data can be shared seamlessly. Solutions should encourage open data standards, having APIs for easier interaction and integration, and transparent methodologies. This would enable different carbon reporting platforms to share and verify data, promoting interoperability and data portability. It would also prevent SMEs from being restricted to a single vendor.	Coordinated vendor collaboration to get agreement on the single format. This may need to be driven by the regulations or a body that defines the standards; similar practices have been seen in the banking sector.
No third-party data verification requirements	To enable reporting of emissions that are credible and have a high degree of accuracy, cooperation and coordination among organisations is a must. Challenges faced should be shared openly and effective solutions found for them, fostering trust and transparency. To support effective sharing practice there should be a set of well-defined practices set out for verification and validation of data (internally or through external verifiers) so that organisations evidence compliance and establish trust in their data.	• Standardised carbon accounting training for a recognised qualification. • There is a need for a qualification body, such as a carbon regulator.

Digital Catapult CCCA data framework recommendations		
Data challenges	Recommendations through the data management framework	Actions to be taken through process framework and regulations
No trusted data-sharing system	It is important that the data sharing system will ensure that only the required data is disclosed. Encryption technologies, secure APIs, and implementation of cybersecurity best practices should be implemented in an organisation before data is shared.	• A 'disclose once' approach, with clear definition of disclosure requirements and formats. • Role-based access to be defined so that only authorised individuals can access the data.
No standardised emissions factor datasets, and significant gaps in scope 3 emissions factor availability	Having standardised emission factor datasets and filling in gaps as much as possible so that organisations have clarity. Also addition and approval of any new emission factors should follow a process to ensure that additions are made available to all. Refer to Energy Catapult work on carbon regulators.	• Top-level spend datasets • Additional datasets for sectors that have more granularity that would need approval • A common or central portal that defines the emission factors that can be used by all (the US LCA Commons portal is an approach that could be followed). A government body would need to regulate such a system.
Insufficient data integration within an organisation, due to data scattered across disparate systems	Implementation of data governance best practices within the organisation to promote its data integration, reducing silos where possible. Automation of data collection and processing processes in a common data storage so that the process becomes seamless in the future. This may be applicable in scenarios where an organisation has an advanced level of carbon reporting.	

Digital Catapult CCCA data framework recommendations		
Data challenges	Recommendations through the data management framework	Actions to be taken through process framework and regulations
No standardised tools	Refer to High Value Manufacturing Catapult work on tools horizon scan (work pack 4 and work pack 2b). Increased market collaboration so that organisations can move away from competition and focus on alignment and collaborations. Implementation of cross-industry or market-automated solutions that can be used by all to streamline footprinting, automate data collection, and generate value through data sharing.	• All tools should follow a set output and conversion factor for datasets, and allow the management of both primary and spend-based data. • This will need coordination between tool providers, and alignment with regulations.
Lack of clarity on data to be collected and reported for scope 3 upstream	Refer to the data collection and acquisition steps as part of the data framework recommendations. Clear guidelines on verifying and reporting of emissions covering as many scenarios as possible so that organisations can focus on the data that needs to be collected and reported. Clear guidelines on establishing boundaries when an organisation needs to report data. Mandatory reporting for all organisations would have everyone collecting the necessary data and reduce the effect of double counting.	Defining of regulations and resource burden minimised.
Using spend-based data or average data over primary data	Organisations may not initially have all the primary data available. They can start with using spend-based data to determine their highest emission contributing area and report emission and move towards collection of primary data by implementing data collection and quality practices through the framework.	Guidelines on best case data usage, varying by business type and supply chain position.

SECTION 1

This section provides context for the recommendations in Section 2, including an introduction to carbon data accounting and the data management framework

INTRODUCTION AND BACKGROUND

> A robust carbon data management system is essential for all organisations aiming to effectively manage carbon emissions within their supply chains to mitigate environmental impact. Accurately collecting carbon data and managing its quality and integrity is fundamental to achieving the objectives set by UK Government.

As the importance of environmental awareness and regulatory scrutiny increases, organisations are expected to reduce their carbon footprint and demonstrate their commitment to sustainability. The UK has aligned with international climate agreements, and our domestic climate targets place an emphasis on reducing carbon emissions, making carbon accounting a critical discipline.

To enable the UK manufacturing industry and its supply chains to manage carbon emissions data effectively and efficiently, a comprehensive carbon accounting framework needs to be designed, built and implemented.

This will enable organisations to:

- Monitor, report and verify carbon data
- Account for their carbon footprint accurately
- Make informed decisions that support reductions in greenhouse gas emissions

Technology solutions will help to automate repetitive processes that are currently being completed manually, and will help to streamline the data and make it fit for use. This will reduce the time and effort involved, resulting in a more efficient and effective carbon emissions data management process.

Implementing the recommendations in this report will only deliver effective carbon accounting when the management and standardisation of data are aligned to effective regulatory frameworks, and supported by a carbon regulator.

The Cross-Catapult Carbon Accounting programme

The Cross-Catapult Carbon Accounting programme (CCCA) began in 2022. Funded by Innovate UK, it is led by High-Value Manufacturing Catapult, collaborating with Connected Places Catapult, Digital Catapult, Energy Catapult and Satellite Applications Catapult.

It focuses on the need for a carbon accounting data management system for manufacturing and supply chains that will:

- Accurately calculate the scope 1, 2 and 3 upstream emissions that comprise most of the carbon footprints of manufacturing and supplier businesses.
- Support the creation of a comprehensive UK framework for carbon emissions, with agreed standards and tools that can be used for accounting, tracking and reporting carbon emissions throughout entire supply chains. This will inform decision-making, enabling businesses to take timely and appropriate actions to accelerate their decarbonisation.
- Unlock investment, creating an environment where manufacturers can excel on the global stage, making the UK stand out as a destination for low-carbon manufacturing.

As part of the CCCA, Digital Catapult provides recommendations for carbon data management based on the various aspects of the data landscape, including data collection, quality, privacy, interoperability, transparency and sharing. This year three report covers the outcomes of our work, provides context for the data challenges faced by the manufacturing industry, and establishes why an ideal data framework is needed.

In year one, we focused on a feasibility study using the ecometer, a tool we created to systematically capture, monitor and analyse carbon emissions.

In year two, we explored the challenges associated with managing carbon accounting data across a supply chain and best practices for achieving data consistency. These are covered in the report Carbon Accounting: a review of the existing data management landscape.

Year three has focused on recommendations for a data framework that will help organisations to manage their carbon data effectively and efficiently, enabling them to make informed decisions to accelerate decarbonisation. The outcome is this report, which provides recommendations to help organisations to effectively and efficiently manage their carbon accounting data, internally and externally across the supply chains, and move towards the carbon reduction goals.

> While this report aims to be as comprehensive as possible, it's important to remember that the landscape is constantly changing, as carbon accounting, data and technology are evolving on a continuous basis.

This work supports the outputs from other Catapults:

Promoting the consistent use of emissions data.

Managing the data through the user journey.

Steps and tools required for the carbon accounting process.

Verifying data to ensure accuracy and consistency.

DATA MANAGEMENT IN CARBON ACCOUNTING

> Accurate and well-managed data is the foundation of effective carbon accounting, enabling organisations to measure, manage and mitigate their carbon footprint in alignment with environment goals and stakeholder expectations.

The value of data in carbon accounting

Manufacturing plays a crucial role in economies around the world, but it is often linked to high carbon emissions.

- Currently, global production accounts for about 20% of carbon emissions.
- Emissions from the average company's supply chain are more than 11 times greater than those from its own operations.

In the journey towards sustainability, mitigation of emissions through efficient and effective carbon accounting is of great importance, as accurate calculation will provide a holistic view of the emissions that occur from each organisation, process or product.

To be able to account for carbon emissions accurately it is necessary to understand the origin of the data used in the calculation as well as identify the important data points. For example, use of data analytics has revolutionised the way in which organisations collect and process huge amounts of data, extracting and applying previously unseen insights to improve efficiency in operations, tracking and lowering carbon emissions, and reducing the waste generated by day-to-day processes. The Greenhouse Gas Protocol (GHG Protocol) is the global framework for measuring and managing greenhouse gas emissions from private and public sector operations and their value chains.

Data types in carbon accounting

An organisation will have two types of data: primary emissions data, and secondary (or spend-based data).

- **Primary data** is the data that an organisation collects internally from its operations and externally from its value chain suppliers or partners for specific activities conducted in that chain.
- **Secondary data** should only be used when primary data is not available or is not of the required standard or quality. Secondary or industry average data comes from such as published databases, industry associations, literature studies, generic data, and proxy data. If primary/ emissions data is not yet available, consider using spend-based metrics such as energy consumption or material expenditure, based on invoices.

Why the carbon accounting process is important

Simply having the right data is not enough to help an organisation move towards achieving sustainability. There needs to be a carbon accounting process in place that guides the organisation to measure, manage and report the organisation's carbon footprint, and empower the adoption of business strategies that reduce its impact on the environment.

In the manufacturing sector, where complexities and emissions are deeply entwined with intricate supply chains, having precise carbon accounting serves as an organisation's essential guide for moving towards more sustainable practices.

Why a framework is needed

Achieving and maintaining compliance with evolving environmental regulations poses a significant challenge for organisations. They need to be agile enough to adapt quickly to changes as they happen, a task which is made more difficult by the lack of robust data management and accurate reporting systems. They need to be able to track and report carbon emissions, and ensure consistency, transparency and reliability in alignment with regulatory standards. Working with supply chains outside their direct control adds a further layer of complexity.

The data management framework includes the tools, processes and technologies used to collect, store, analyse and report carbon emissions data. Storing and managing carbon inventory data to account for carbon emissions helps an organisation to understand its carbon footprint and take corrective action as needed. But with data comes many challenges around managing, storing, processing and exchanging it (as articulated in Greenhouse Gas Data Management by World Bank in 2016).

How this report can help manufacturing organisations

The GHG Protocol covers many areas in depth, making it heavy going for organisations that are starting out on their carbon accounting journey. Many don't have the necessary skills or human resources to be able to study the protocol documents in the depth needed.

The practical recommendations in this report deconstruct the GHG Protocol document, making the task of getting started less burdensome.

GHG PROTOCOL

Building on a 20-year partnership between World Resources Institute (WRI) and the World Business Council for Sustainable Development (WBCSD), the GHG Protocol works with governments, industry associations, NGOs, businesses and other organisations.

The GHG Protocol classifies a company's emissions into three scopes:

Scope 1

Direct emissions from owned or controlled sources.

Scope 2

Indirect emissions from the generation of purchased energy.

Scope 3

All indirect emissions not included in scope 2 that occur in the value chain of the reporting company.

SECTION 2

Recommendations for an ideal carbon data management framework

This section provides recommendations that address the carbon data management challenges already identified, as well as insights into managing carbon data. The aim is to help organisations to use their carbon data to make informed decisions that help them manage their GHG emissions more effectively and advance their sustainability journey.

THE DATA MANAGEMENT FRAMEWORK

The carbon data management framework covers the following aspects of data management, all of which are equally important:

- Data governance
- Data architecture and metadata management
- Data collection and acquisition
- Data quality management
- Data verification and validation
- Data interoperability and sharing
- Data privacy and security

Any framework that manages data should be trustworthy, and the four pillars of a trustworthy carbon data management system or framework should be:

- Data availability
- Data quality
- Data compatibility or interoperability
- Data privacy and security

Figure 1 visualises the trustworthy carbon data framework, including an example of the technologies that can be applied to it in the construction sector.

The recommendations in **Figure 2** capture the various actions required to ensure trustworthy data within this framework. This report covers these recommendations in detail, and provides examples of tools and technologies that can be used for each section.

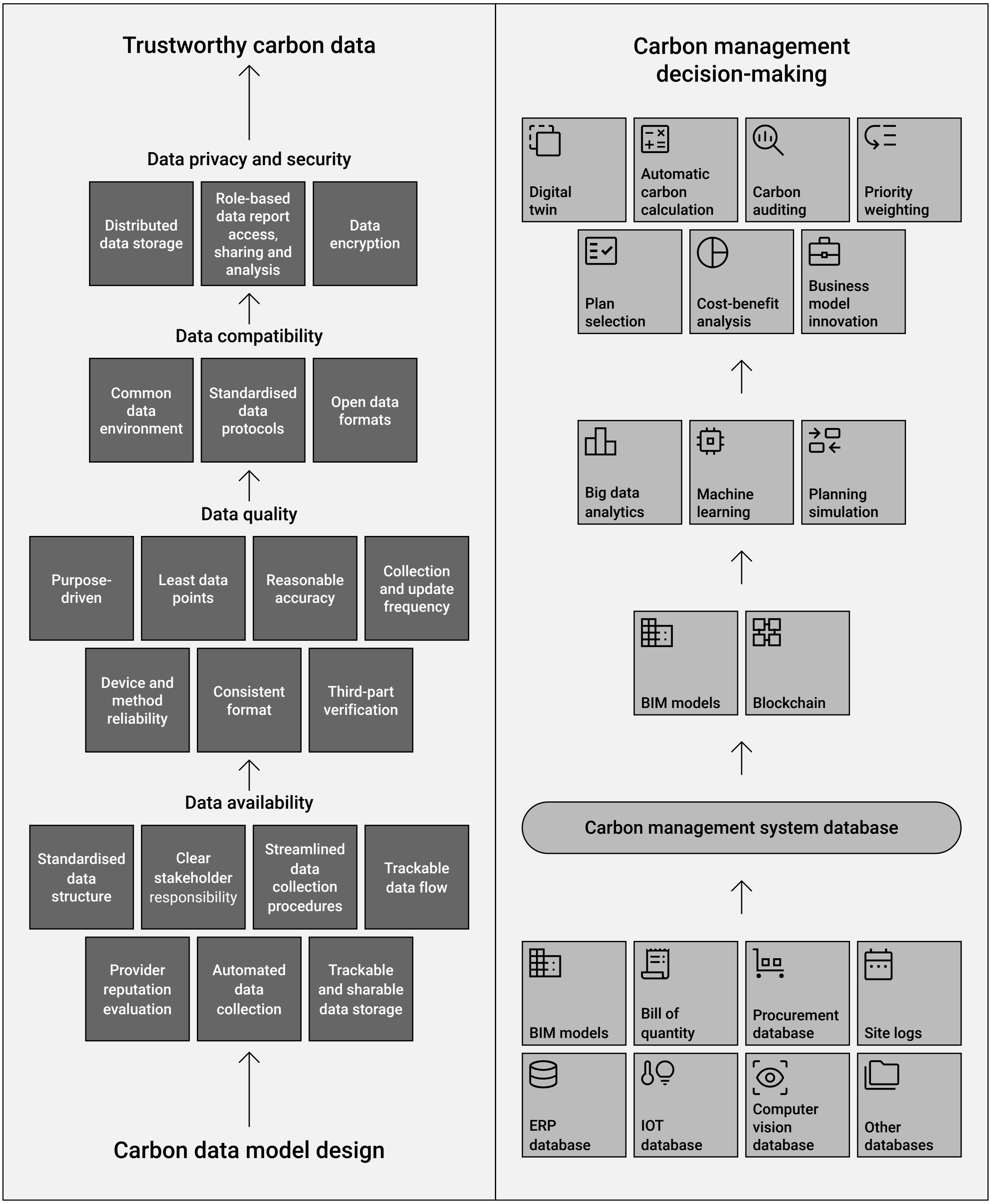


Figure 1: Trustworthy carbon data management framework and supporting technologies

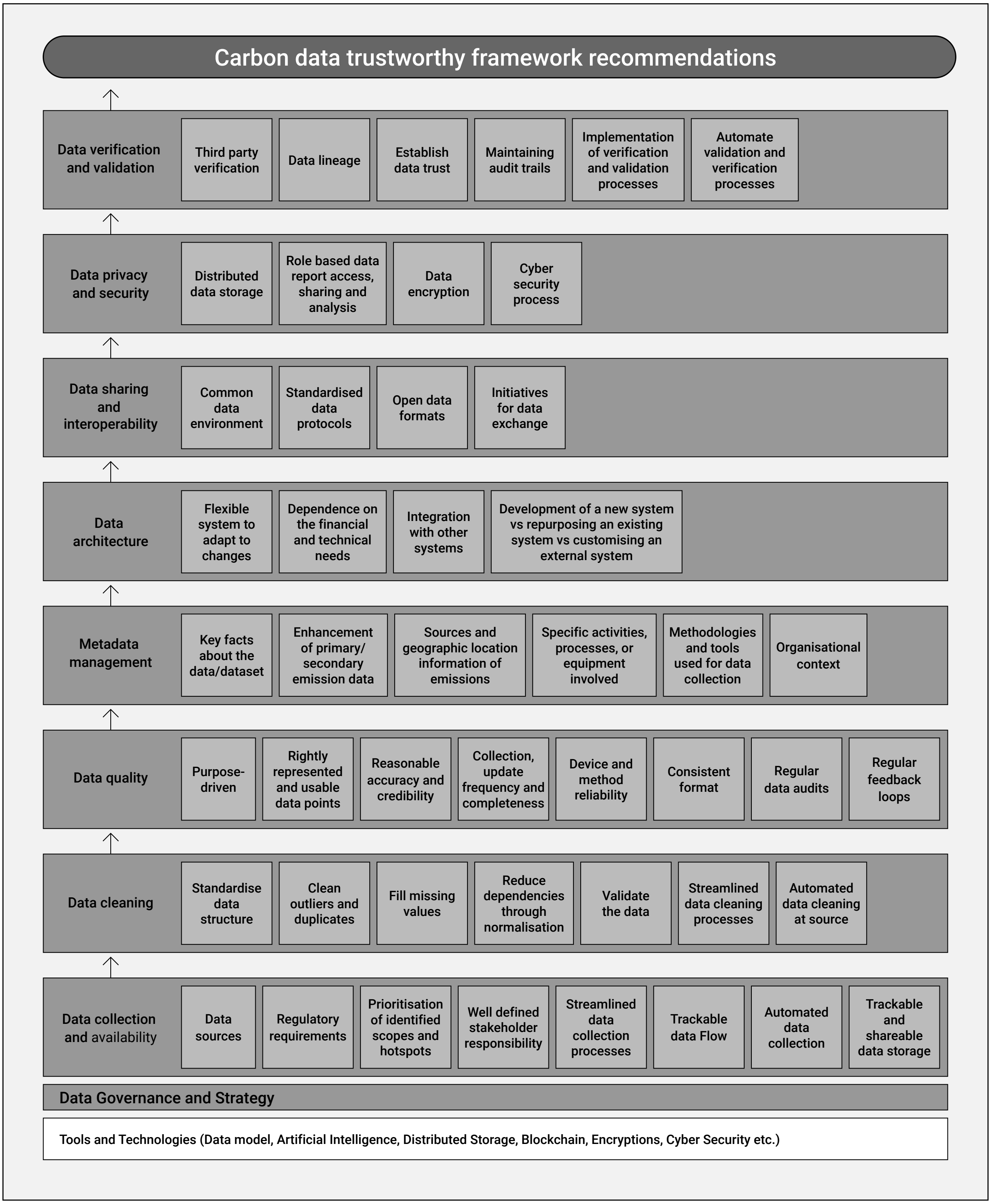


Figure 2: Recommendations for a trustworthy carbon data management system

Observing regulatory requirements

The regulatory requirements and applicable laws for carbon reporting may differ by location, geography, sector or company size. Therefore the organisation needs to familiarise themselves with the laws and regulations concerning carbon emissions in the regions that they operate in, and which typically lay out any mandatory reporting requirements that need to be followed.

Aligning with relevant scopes and categories

To understand what type of data needs to be collected and what categories their day-to-day operations fall into, an organisation needs to use the identified scopes and categories from the GHG Protocol to establish reporting boundaries. The choice of relevant carbon accounting framework can be found from those already established, such as the GHG Protocol. Selecting a standard is essential, as this will determine what data can be collected for the different categories and scope identified.

Evaluating accessibility of data

All the data needed to calculate emissions may or may not be available from the current internal or external data sources that the organisation may have. Data could come from utility bills, industry databases, logs of transportation or waste records. Organisations should prioritise data sources that may be readily available now or that are feasible to collect data from, and have a plan in place to fill data collection gaps in the future — as well improve the quality of data when collecting it.

Understanding the data needs of all stakeholders

Engaging with stakeholders both internally and externally will ensure a comprehensive approach when it comes to collecting data. Consult with sustainability teams, internal departments such as finance and procurement, and other key stakeholders in order to understand their data needs, and any perspectives they may have when it comes to data collection and use.

Identifying hotspots

To identify carbon emission hotspots, available historical data can be examined to identify any trends and patterns that may have occurred or are occurring. Rank the emissions from areas having the highest carbon emissions to the lowest. Conducting a materiality assessment that involves the relevant stakeholders to pinpoint the most significant emission sources and activities will help prioritise data collection efforts, focusing on areas with the greatest impact on environmental, social and governance issues for the organisation.

Reviewing and improving the process over a period of time

The data collection plan should always be kept in review and continuously improved. Any gaps should be identified and filled and improvements should be made on the basis of past experiences, changes in data collection methods and technologies and changes in organisational needs. This will further help in streamlining data collection processes.

Organisations should ensure that data collected is accurate, complete, consistent and representative of their activities and emission sources. Data lineage and provenance of the data should be captured when emissions are recorded. Carbon accounting data should be collected and updated regularly, ensuring that it remains up to date and reflective of the organisation's current operations.

Automate data collection

Explore any emerging trends or technologies that can automate the data collection process. Examples include the use of Internet of Things Devices or carbon accounting software solutions that can provide insight in real time on carbon emissions etc.

DATA ARCHITECTURE

> Data architecture is the structured design and organisation of data systems and processes that enable the efficient collection, storage, management, analysis, and reporting of data — in this case, relating to GHG emissions and other environmental factors.

A well-designed data architecture is crucial for accurate and reliable carbon accounting. As well as meeting current needs, it should be flexible enough to accept future requirements and any regulatory or policy-related changes.

Once the system's functional requirements, and the financial, technical and resources required to build it have been identified, there are three possible development routes to follow:

- Development of a new system internally, using external resources.
- Re-purposing an existing system — for example, when a large organisation does not want to alter their existing data structure.
- Customising an external system.

Approaches to mitigating the costs of developing a carbon data management system will depend on the scope of requirements and which of three routes is taken. In all cases, it is essential that the system is able to integrate the data from other systems and available datasets, and facilitate secure data exchange.

Data strategy and governance

Best practice dictates that an organisation should share and integrate data on a peer-to-peer basis. This makes it important to establish processes for implementing and operating data governance and data strategies over the long term.

The first step in a data strategy is to streamline and centralise the process of gathering and managing carbon-related data. This will close as many data gaps as possible, while also ensuring everyone who needs it has access to the data and also help eliminate any double reporting that may exist.

The next step is to design the data architecture in order to enrich the data, manage data from disparate sources and validate data quality. Third party support and verification is a good way to ensure that data is accurate and reliable and doing so can avoid any risks associated with bad carbon accounting data.

Data governance in carbon accounting refers to the set of practices and policies that ensure the quality, integrity, security, and responsible use of GHG and environmental data. This involves:

- Clearly defining roles and responsibilities for data ownership within the organisation, with designated individuals or teams responsible for collecting, managing, and maintaining emissions data.
- Putting in place data policies and procedures that outline data collection, storage, and reporting standards, all aligned to international reporting standards, such as the GHG Protocol.
- Implementing data security measures, including access controls, encryption, and data masking, to protect sensitive emissions data from unauthorised access or breaches.
- Establishing processes for data quality control, including validation and verification, to ensure that emissions data is accurate and reliable.

It is also important to regularly review and refine carbon accounting best practices and update data governance in response to new data practices and evolving standards.

Metadata management

Metadata generally contains the key facts about any data or dataset. It should include basic details, such as the name, generation date, version, a description of what the data is, the type of data collected (structured, unstructured, time series etc.), how it is going to be used, the data source, and if there are any restrictions on its usage.

In carbon accounting, metadata refers to the contextual and additional information that supports and enhances the primary emissions data and data specificity. This should include details about the sources of emissions (such as the activities, processes, or equipment involved), and the methodologies and tools used for data collection (direct measurement, estimation, or calculations). Metadata should also specify the time periods covered, dates of data collection, geographic information about the locations of emissions, and organisational context (which divisions or subsidiaries are included in the emissions inventory).

Using conventions such as XBRL to tag metadata can support the use of carbon data with other data, such as financial information.

Metadata supports data quality and reliability by providing a clear picture of the data's accuracy, precision, and any uncertainties or assumptions it may contain. It also ensures compliance with regulatory and reporting standards like the Greenhouse Gas Protocol and ISO 14064.

Additionally, metadata records data processing and transformation details, audit trails, and version control information, tracking changes over time to ensure transparency and enable verification. Comprehensive metadata ensures transparent, reproducible, and compliant carbon accounting, facilitating accurate reporting and informed decision-making.

Most tools or platforms that provide holistic carbon accounting solutions include metadata management. Examples include Persefoni, Sustain. Life, Greenly, IBM Environmental Intelligence Suite, Normative and many more. The Climate Action Data Trust (CAD) is an open-source metadata platform that uses blockchain technology to create a decentralised record carbon market activity in order to avoid double counting, and increase trust and build confidence through increased transparency.

DATA TYPES AND REQUIREMENTS

> Accurate and well-managed data is the foundation of effective carbon accounting, enabling organisations to measure, manage and mitigate their carbon footprint in alignment with environment goals and stakeholder expectations.

Primary data

This is the data that an organisation collects internally from its operations and externally from its value chain suppliers or partners for specific activities conducted in that chain.

Examples of primary data and its sources include:

- Kilowatt-hours of electricity consumed
- Supplier-specific data
- Transportation records such as fuel consumption, distance travelled
- Kilograms of material consumed
- Hours of time operated
- Square metres of area occupied
- Waste records, such as kilograms of waste generated
- Emission factor data
- Meter readings, purchase records, utility bills, engineering models, direct monitoring
- Product life cycle assessments

Secondary data

This is also known as industry average data, and comes from published databases, industry associations, literature studies, generic and proxy sources, and so on. It should be used only when primary data is not available, or not of the required standard or quality.

Examples of secondary data include:

- Bills for any purchased goods or services
- Financial data related to organisational spend
- Industry averages as coefficients
- Utility bills
- Emission factors from industry databases

Primary data for scope 1 and scope 2 is usually available more readily, as it comes directly from company operations and sources. Scope 3 data is harder to come by, as it depends on external sources.

Advantages and disadvantages of using primary and secondary data

	Advantages	Disadvantages
Primary data	• Emissions are represented more accurately, allowing better performance tracking. • Enables supplier tracking and benchmarking of those in the same sector; this can be done by tracking their operational changes and actions taken to reduce emissions. • Can be used to calculate emissions for major emissions-contributing areas.	• Difficult to verify the source or quality of data provided by value chain partners. • Resource-intensive, time-consuming, and costly process.
Secondary data	• Can be used when primary data is insufficient or of poor quality. • Can be useful for identifying hotspots, or calculating rough estimates of GHG emissions, to prioritise data collection, supplier engagement and reduction efforts. • Cost-effective and easier to collect than primary data. • Can be used for minor or less significant emissions-generating activities.	• May not be fully representative of all activities. • May not consider all aspects that apply when calculating emissions. • May limit the ability to track progress towards GHG reductions. • Could be difficult to quantity the operational decisions and actions taken by value chain partners.

While the GHG Protocol recommends using primary data over secondary data for any kind of carbon emissions reporting, this is not always possible. An organisation can try to collect as much primary data as possible for scope 1 and 2, and assess what is currently available for scope 3.

As an initial step, secondary data or industry average data can be used to calculate scope 3 and identify hotspots. Collection of primary data for the hotspots identified can then be initiated. For scope 3 emissions that may not be significant or are less significant, secondary data may be used if primary data is not available. Over time, the organisation can move towards collecting primary data that is accurate and of good quality for all of scope 1, scope 2 and scope 3 emissions.

Figure 3 shows an example of primary and secondary data, as compiled in the GHG Protocol, relating to scope 3 upstream emissions.

Upstream scope 3 emissions		
Category	**Examples of primary data**	**Examples of secondary data**
1. Purchased goods and services	• Product-level cradle-to-gate GHG data from suppliers calculated using site-specific data. • Site-specific energy use or emissions data from suppliers.	• Industry average emission factors per material consumed from life cycle inventory databases.
2. Capital goods	• Product-level cradle-to-gate GHG data from suppliers calculated using site-specific data. • Site-specific energy use or emissions data from capital goods suppliers.	• Industry average emission factors per material consumed from life cycle inventory databases.
3. Fuel- and energy-related activities (not included in scope 1 or scope 2)	• Company-specific data on upstream emissions (e.g. extraction of fuels). • Grid-specific T&D loss rate. • Company-specific power purchase data and generator-specific emission rate for purchased power.	• National average data on upstream emissions (e.g. from life cycle inventory database). • National average T&D loss rate. • National average power purchase data.
4. Upstream transportation and distribution	• Activity-specific energy use or emissions data from third-party transportation and distribution suppliers. • Actual distance travelled. • Carrier-specific emission factors.	• Estimated distance travelled by mode based on industry-average data.
5. Waste generated in operations	• Site-specific emissions data from waste management companies. • Company-specific metric tons of waste generated. • Company-specific emission factors.	• Estimated metric tons of waste generated based on industry-average data. • Industry average emission factors.
6. Business travel	• Activity-specific data from transportation suppliers (e.g., airlines). • Carrier-specific emission factors. • Can be used when primary data is insufficient or of poor quality.	• Estimated distance travelled based on industry-average data.
7. Employee commuting	• Specific distance travelled and mode of transport collected from employees.	• Estimated distance travelled based on industry-average data. • May not be fully representative of all activities.
8. Upstream leased assets	• Site-specific energy use data collected by utility bills or meters.	• Estimated emissions based on industry-average data (e.g. energy use per floor space by building type).

Figure 3: Examples of primary and secondary data for scope 3 upstream emissions

Emission factors

Emission factors are the GHG emissions that occur per unit of activity. These factors are multiplied by the activity data to calculate carbon emissions. Emission factors may cover different gases and are calculated in terms of units of their carbon dioxide equivalent emissions or they may cover another type of gas, such as methane per litre of fuel.

Examples of emission factor sources may include life cycle databases (LCA databases), industry associations, company-developed factors, peer-reviewed literature, published product inventory reports and government agencies.

Examples of emission factors include:

- kg CO2 equivalent emitted per litre of fuel consumed
- kg CO2 emitted per kWh of electricity consumed
- kg CO2 emitted per kilometre travelled
- kg CO2 emitted per unit of currency spent

Most such measures are of CO2 equivalent. It is less common to have pure CO2 numbers that don't include other pollutants.

The type of emissions factor needed will depend on the data that is going to be used for carbon calculations. For example, if companies collect primary activity data for certain activities then they would need emission factors relating to those activities. An organisation may also collect activity data on the basis of the emission factors they have.

For example, when calculating emissions from electricity, organisations should typically use emission factors from regionally relevant electricity sources. If an electricity supplier can provide their specific emission factor, then that supplier's data should be used as the primary source — as long as the supplier emissions are excluded from the region-specific emission factor. If that is not the case, then the average regional emission factor for electricity should be used, in order to avoid double counting.

Environmentally extended input output (EEIO) data

If the organisation uses financial spend data or secondary data for, say, a material input for a process, they will need to use EEIO emission factors to calculate upstream emissions.

EEIO models estimate the GHG emissions and/or the energy consumption that result from activities in production and the upstream supply chain for specific products and sectors. These models are derived when national GHG emissions are allocated to groups of finished products, based on the economic flows between industry sectors.

The typical output of EEIO models is the amount of GHG emitted per unit of revenue. EEIO emission factors that result from this can be used to estimate cradle-to-gate GHG emissions for a given product category or industry. When prioritising efforts for data collection, EEIO data sources are useful for identifying hotspots.

Process-based data

Process-based data is obtained by evaluating all the known energy and environmental inputs of a process and calculating the direct emissions associated with its outputs. This approach is especially useful for analysing unique processes and individual products.

The advantages of this type of data are:

- A detailed analysis of the process and any unique insights can be uncovered.
- There is a high level of specificity and focus achieved through this data.
- The concept is straightforward.

The disadvantages of this type of data are:

- Data collection may be time-consuming, costly and resource-intensive.
- The system boundary and data is usually chosen by the organisation, which may prevent data compatibility.
- It may be impractical for large-scale, multi-product analysis, or at business level.

Combining EEIO data and process-based data

Depending on the approach an organisation wants to take, they can use either the top-down EEIO method, or bottom-up process method. They can also use a combination of the two approaches. For example, EEIO could be used to calculate the upstream emissions of purchased goods or services, while a process-based approach could be used to calculate the end-of-life and downstream emissions.

Proxy data

Organisations should assess the quality of their data. If the data of the required quality is not available, proxy data can be used to fill the gaps. Proxy data is data from an activity that is similar to the one being considered as a stand-in activity. This data can be made more customised as per the requirements and representation of the said activity or can be extrapolated/forecasted or scaled up as needed.

For example an organisation may have ten manufacturing facilities out of which it has data for eight and needs to fill in the gaps for the other two. It may use proxy data for the remaining two facilities by grouping facilities having similar characteristics such as by location or type of facility and then calculate the intensity ratio for the group of facilities that data is available for which could be the total quantity of emissions per unit of production that has been produced and apply it to the two facilities. Also if a large amount of data is needed for a scope 3 category and it is challenging to collect data, an organisation could use sampling techniques to predict or extrapolate the data from a sample of activities that are a representation of the activity.

Based on the GHG Protocol Product LifeCycle Accounting and Standard, Figure 4a provides an example of types of data that can be collected for a process that uses diesel fuel.

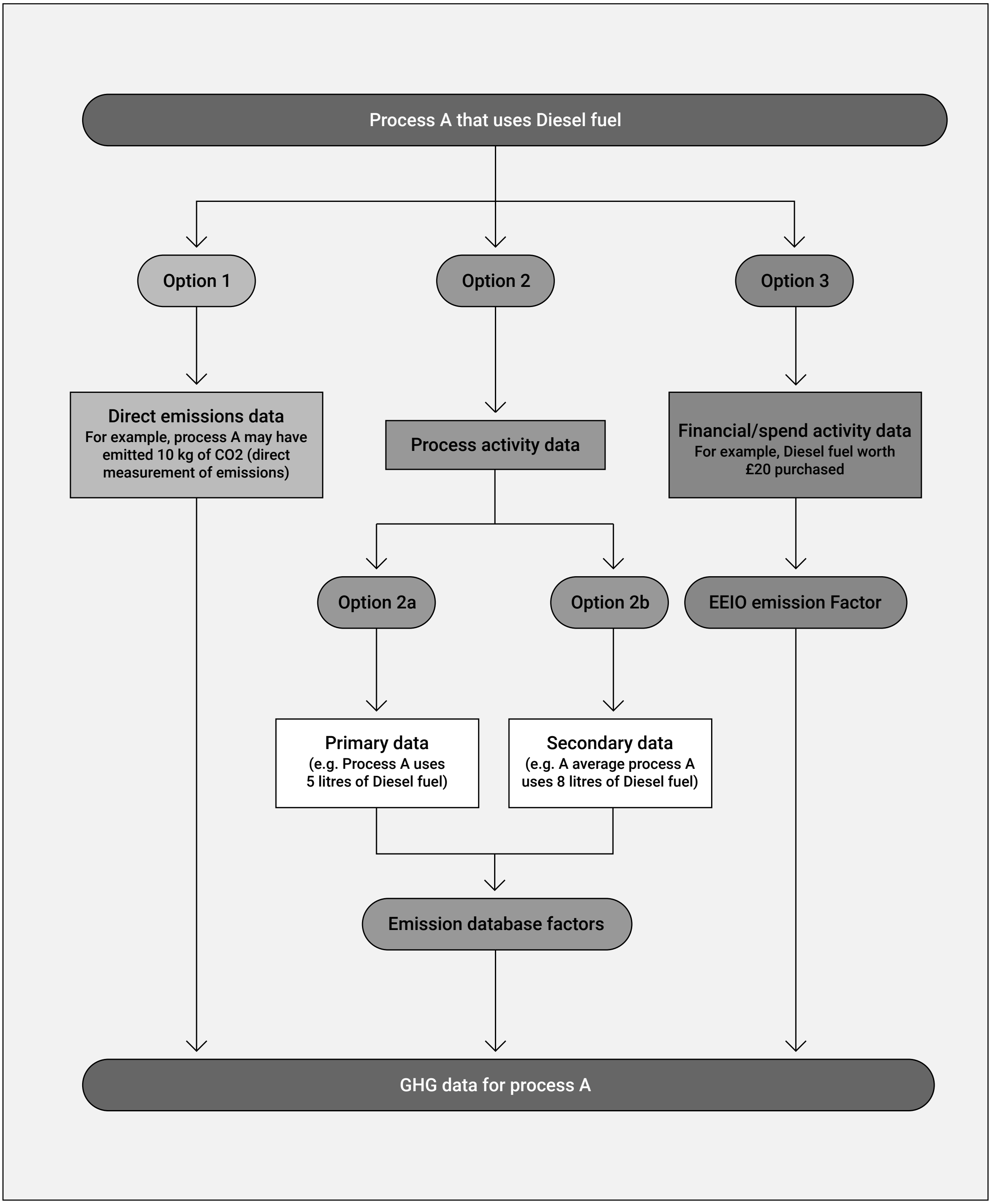


Figure 4a: Options to calculate the GHG data for a process

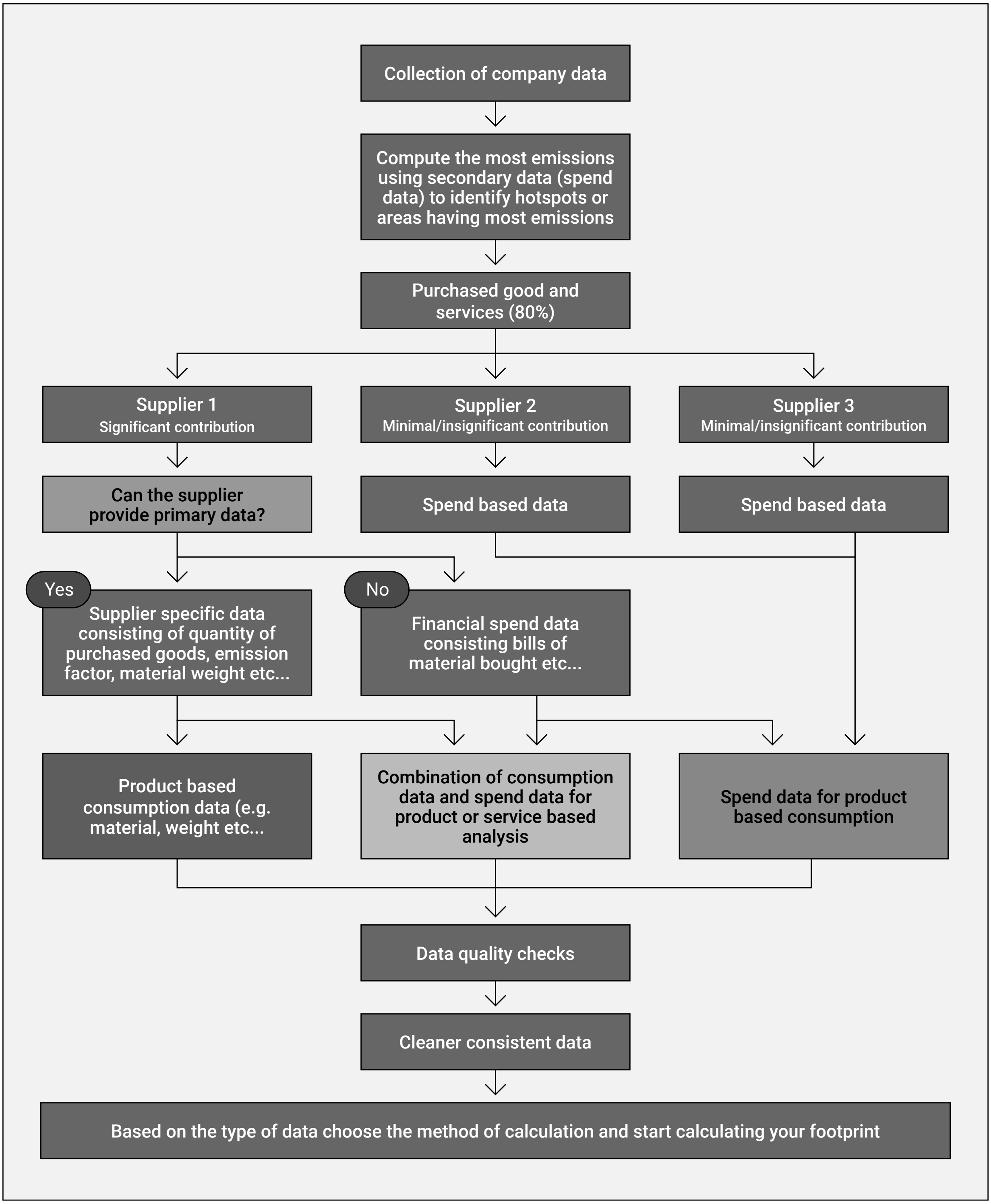


Figure 4b: Example of how primary and secondary data can be used, based on availability of scope 1 data

DATA SOURCES

> Identifying the appropriate carbon accounting data sources for an organisation requires a methodical approach, taking into account factors such as the industry, business operations, and sustainability objectives. Figure 5 shows the steps that an organisation can begin with in order to start with relevant carbon data, as set out in the GHG Protocol.

Identify contributors

The first step is to identify the major contributors to the organisation's carbon emissions. To do this, they first need to identify all the data sources, external or internal for scope 1, 2 and 3 emissions contributing to the organisation's total carbon footprint. This can be done by mapping the organisation's business structure, operations, any products or services that the organisation provides and the supply chains.

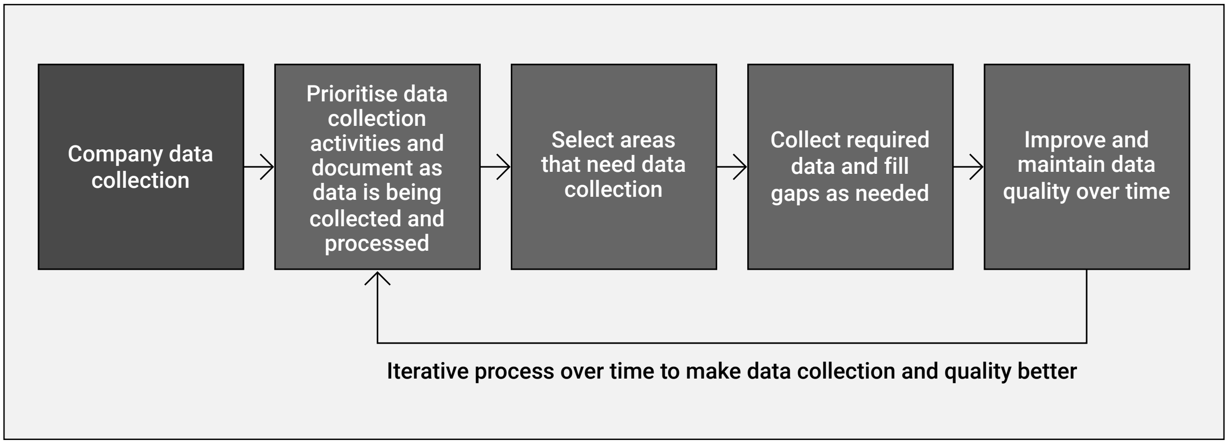


Figure 5: Steps in carbon data collection, as per GHG Protocol

When considering the supply chain, the entire value chain needs to be taken into account, upstream and downstream. The initial focus can be on the first tier of direct suppliers. It is important to consider factors such as the sectors the organisation operates in, the geographical locations involved, the types of energy being used, the modes of transport being used (including employee commutes, business travel and transportation of goods), and any waste generated from their operations.

Figure 6 shows a representative example of data that can be collected, as per the GHG Protocol categories, for an organisation dealing with fleets. This list is not exhaustive.

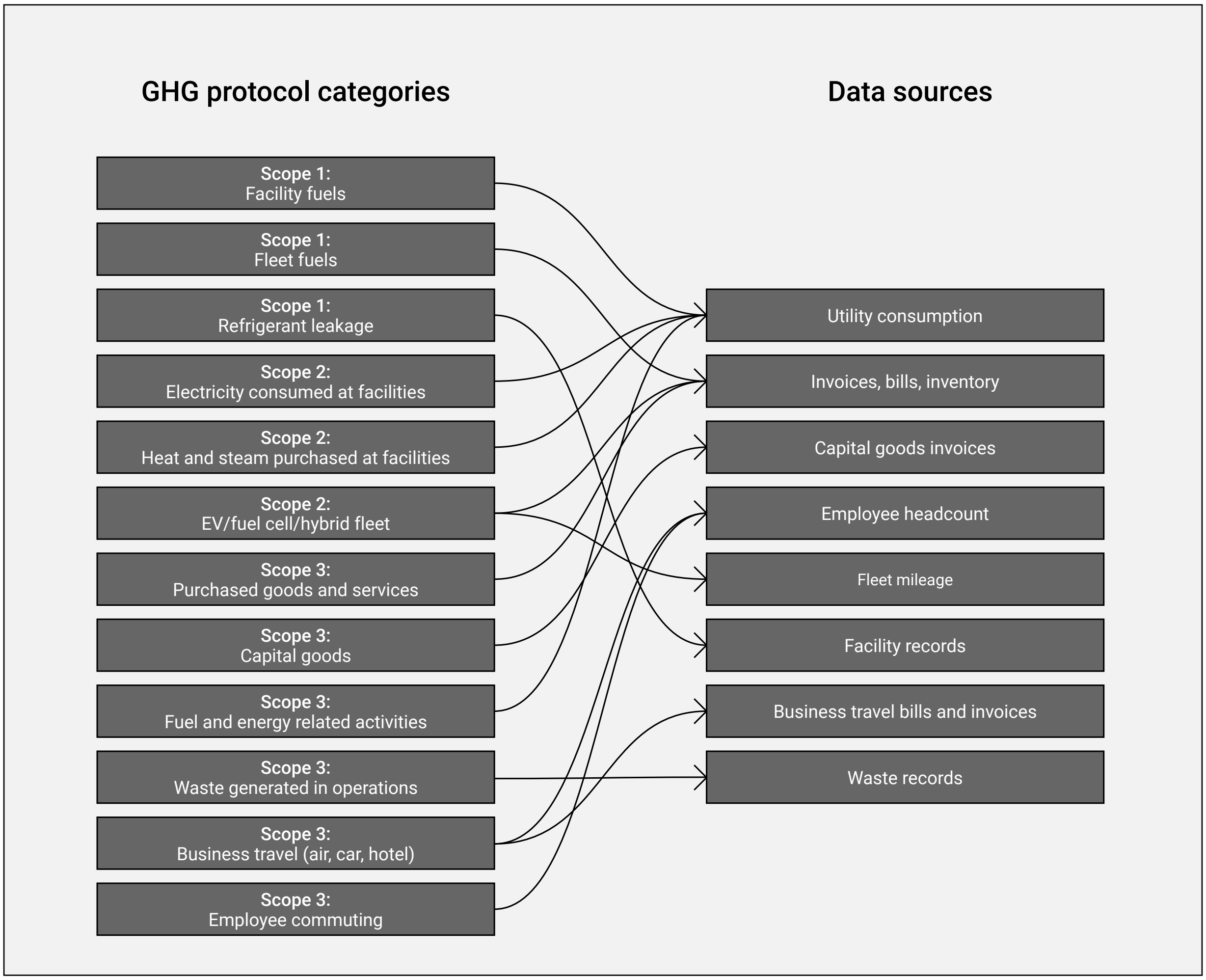


Figure 6: Example of data collected for GHG Protocol categories

Scope 1, 2 and 3 sources

Data for scope 1, scope 2 and scope 3 comes from different sources. Figure 7 depicts the different scopes for [data that needs to be collected for carbon accounting](#).

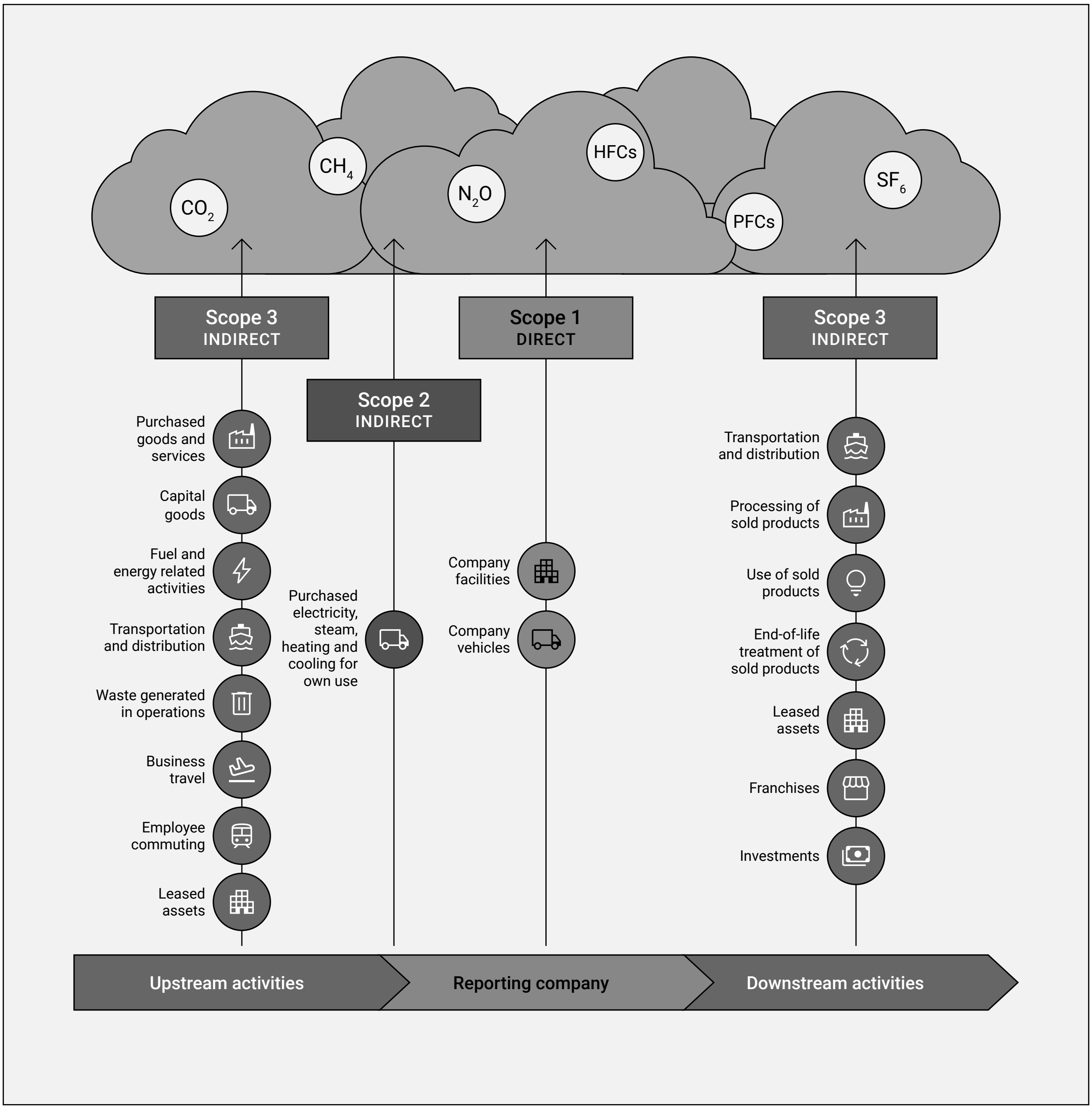


Figure 7: Scope 1, scope 2 and scope 3 emissions

Data sources for scope 1 emissions

Scope 1 carbon emissions are those from sources owned and controlled by the organisation.

These are some of the sources that scope 1 data can be collected from:

- **Records of fuel consumption:** for on-site equipment or machinery, and for vehicles that are powered by all fossil fuels, including petrol, diesel and natural gas.
- **Energy bills:** utility bill and energy consumption data, including information on the use of fuel or natural gas for heating purposes.
- **Data on heating and cooling systems being used:** the use of fossil fuels for heating systems, such as boilers and furnaces, and for cooling systems that use refrigerants.
- **Data on equipment and machinery used:** the types and numbers of equipment and machinery emitting GHGs during operation.
- **Fugitive emissions data:** unintentional GHG emissions, such as leaks from systems that handle fossil fuels or coolants.
- **Records on maintenance and repair:** data on maintenance activities that may have affected equipment efficiency and therefore emissions.
- **Chemical reactions or process emissions data:** details of chemical processes that release GHGs as byproducts (often associated with industrial operations), and data on biomass combustion, if applicable, as this can also be a source of direct emissions.

An example of scope 1 emissions and the associated type and accuracy of data is shown in Figure 8.

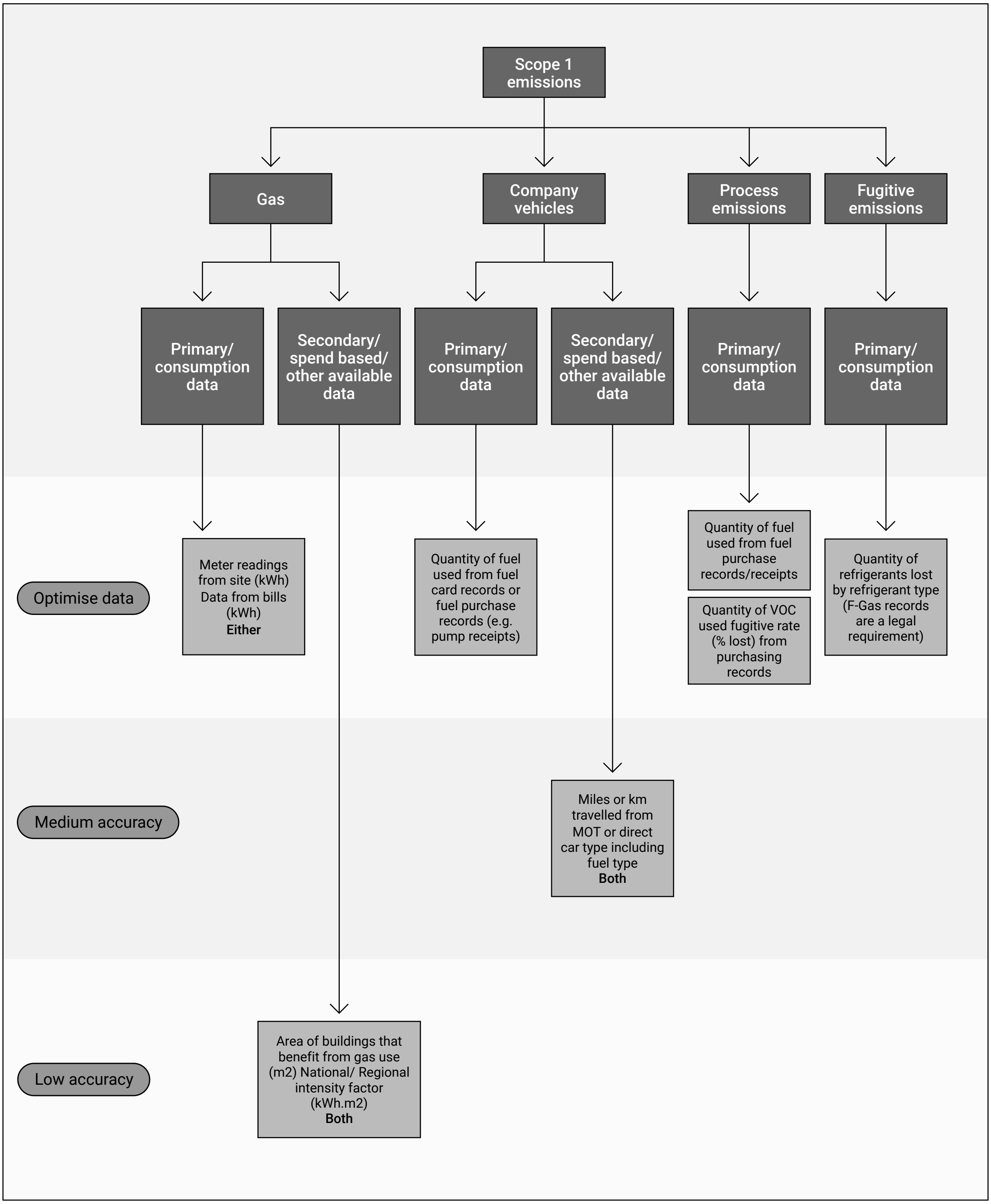


Figure 8: Example of scope 1 emissions data sources with associated levels of granularity

Data sources for scope 2 emissions

Scope 2 emissions are indirect emissions from the use of purchased steam or electricity.

Some of the sources that scope 2 data can be collected from include:

- **Utility bills and invoices:** records detailing the consumption and costs of purchased electricity, heat, steam, or cooling provided by external suppliers.
- **Renewable Energy Guarantee of Origin (REGO):** REGO digital certificates verify the purchase and use of renewable energy, and can be used with transparency through the green accreditation scheme.
- **On-site renewable energy sources:** data on energy generated by on-site renewable sources, such as solar panels or wind turbines.
- **Power purchase agreements (PPAs):** contracts that specify the terms for buying electricity, particularly relevant for organisations that source energy from renewable providers.

Scope 2 emissions can be location-based or market-based, and carbon intensity will vary depending on the source. If they are location-based or region-specific, they will most likely reflect the average carbon intensity of the grid through which consumption occurs, and will usually be calculated using grid average emission factor data. If they are market-based, emissions are calculated from electricity that an organisation may have chosen to purchase.

Using region-specific data on the basis of the carbon intensity of electricity being used can aid in accurate assessment of emissions occurring through energy consumption. Figure 9 shows an example of scope 2 emissions and the associated type and accuracy of data.

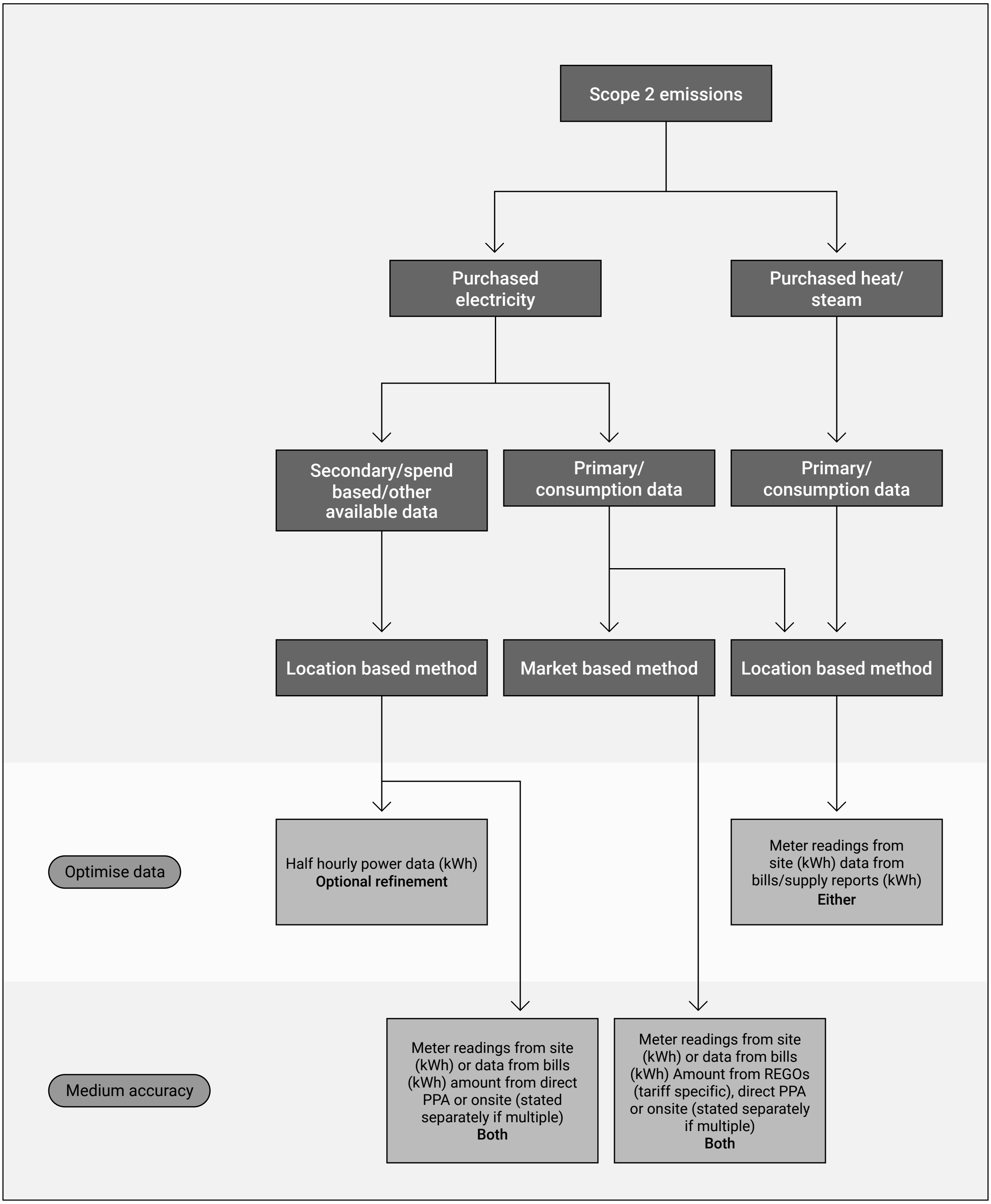


Figure 9: Example of scope 2 emissions data sources with associated levels of granularity

Data sources for scope 3 emissions

Scope 3 emissions are indirect emissions that come from an organisation's value chain or supply chain. These are generally outside the organisation's direct operations (scope 1) and purchased energy (scope 2). Data collection for scope 3 is often complex, due to the wide and complex range of activities occurring along the supply chain.

- **Business travel information:** data on employee travel activities, such as means of transport used, distance covered, travel duration, and fuel usage when available.
- **Employee commuting information:** such as the distances they travel, types of transport they use, and the frequency of their commutes.
- **Waste management information:** records of the waste produced by the organisation, including types of waste, methods of disposal, and treatment procedures.
- **Water usage and wastewater information:** data on water consumption and wastewater treatment, including the amount of water used and energy required for its treatment.
- **Supplier emissions data:** environmental impact data from suppliers, including their greenhouse gas emissions related to producing and transporting goods or services to the organisation.
- **Leased assets emissions data:** information on emissions from the use and upkeep of leased items (such as vehicles, equipment, and buildings).
- **Value chain partner emissions data:** emissions data from franchisees, other partners within the organisation's value chain or any other joint ventures.
- **Investment emissions data:** emissions data linked to the organisation's investments, including indirect emissions from funded projects.

The GHG Protocol specifies eight categories of scope 3 upstream emissions. An example of one of the categories is shown in Figure 10, detailing the associated data type and accuracy, the different calculation methods.

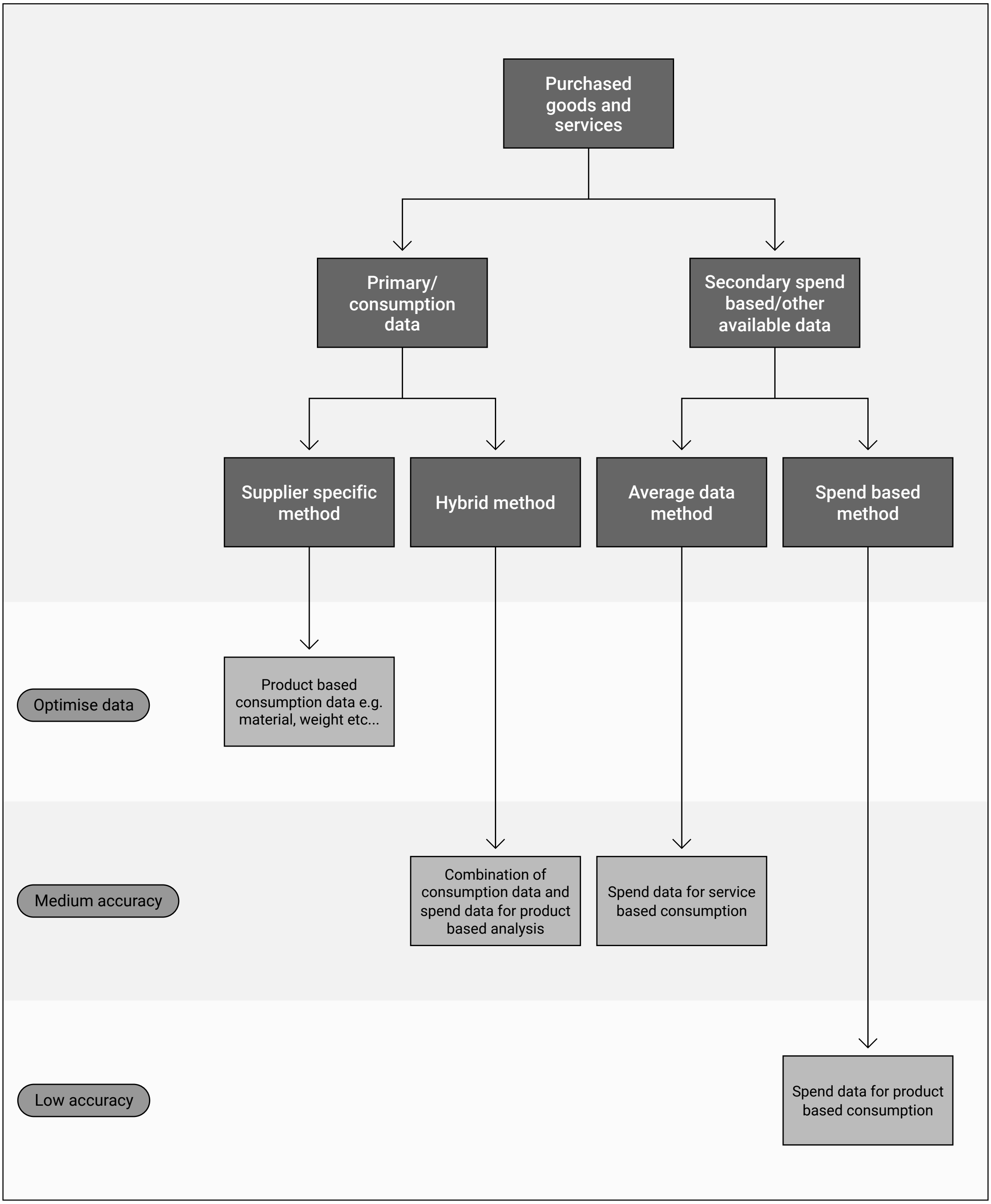


Figure 10: Example of scope 3 emissions data sources with associated levels of granularity

The remaining scope 3 categories have been mapped to the types of data presented in the HVMC report. These maps, and the link to the full versions, are included in the Appendix to this report. The categories are mapped to the diagram in Figure 11 shows the process that can be used to begin collecting data for scope 3.

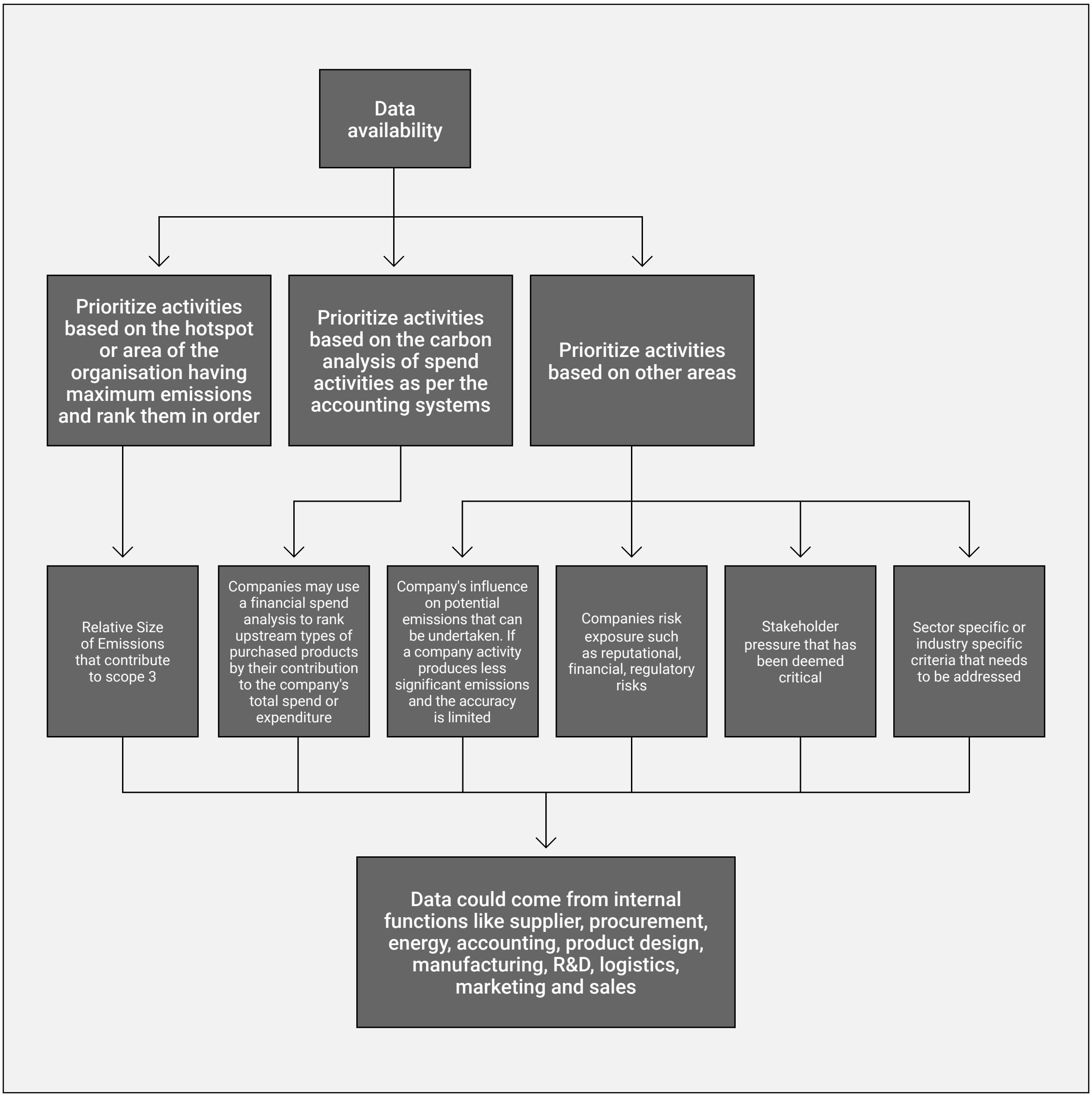


Figure 11: Example of data collection process for scope 3

Product carbon footprint data (PCF)

Data for a product's carbon footprint should consider emissions for the entire life cycle of from material acquisition, material preprocessing, production, distribution, and use through to its disposal at end of life. Figure 12 shows the process that can be used to collect data for products.

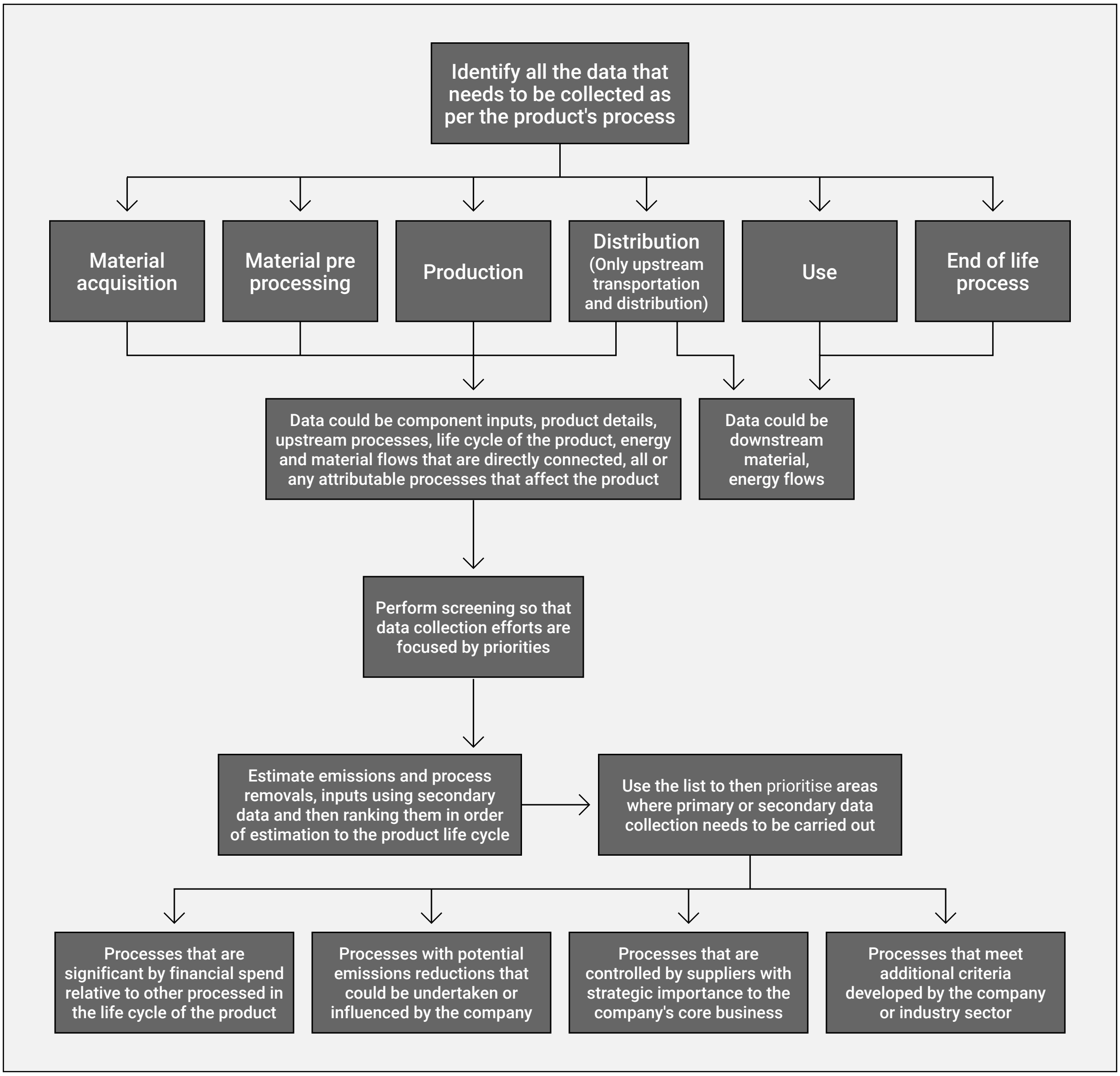


Figure 12: Example of data collection process for product footprint data

To calculate a product carbon footprint these data sources may be used:

- **Extraction of raw materials:** data on the types and quantities of raw materials being used in the production process, along with the consumption of energy during extraction and processing.
- **Emissions on material pre-processing:** information on emissions released by machinery and infrastructure during material pre-processing, creation and disposal of capital goods.
- **Production and manufacturing data:** data about the energy usage, type of materials involved, weights, and the chemical processes used during production and manufacturing processes.
- **Transport and distribution:** information on the transportation and distribution of raw materials, components and final products, covering vehicle types used, distance travelled, and fuel consumed.
- **Product usage:** information on energy consumed and other emissions during the product's use, which should also factor in energy efficiency and data on user habits (often based on surveys or industry averages).
- **End-of-life handling:** data on how products are disposed of in the organisation, including recycling rates, standard waste treatment methods, and the potential for reuse or recycling.
- **Packaging materials:** information about the packaging materials used across the product's life cycle.

Life cycle assessment and carbon accounting

Life cycle assessment (LCA) is mostly used to calculate emissions for the entire life cycle of a product or service, and unlike a product's carbon footprint, it accounts for all emissions from all industries and activities involved.

Carbon accounting is used to account for and report total carbon emissions by an organisation on an annual basis.

The table below highlights the differences between the two, as outlined by the [US Environment Protection Agency](#).

	Life cycle assessment	Carbon accounting
Characteristics	• Evaluates GHG emissions for the entire life cycle of a product or service, wherever they occur. • Provides a system perspective, considering every aspect — from raw material extraction through to end-of-life disposal.	• Comprehensive emissions accounting at organisational, local, regional, national or global level. • Evaluates carbon emissions using categories or sectors (such as energy, waste, purchased goods, manufacturing). • Carbon emissions are estimated for a calendar year.
Uses	• Identification of hotspots or large emitters within the life cycle. • Establish the GHG footprints needed, track progress, and measure the effect of reduction efforts. • Evaluate any changes there may be to policies and decisions relating to product or service's life cycle. • Comparison of impacts between alternative methods, processes, materials and practices.	• Identification of hotspots or large emissions within an organisation, locality, region or country. • Establishment of goals, baselines and targets to reduce carbon emissions. • Tracking of carbon emissions to understand trends. • Measurement of annual reductions and meeting reporting requirements as required. • Comparisons of impacts across different entities.

The data sources used for both will be similar when calculating emissions. However, there may be differences in the use of emission factors and their boundary conditions, as LCA deals with the life cycle of the product or service while carbon accounting is a more comprehensive picture of the emissions at organisational level.

When performing life cycle assessments, the LCA database needs to be referred to for emission factors, while carbon accounting may use emission factors from EEIO/MRIO databases, industry average databases and others.

DATA COLLECTION

Responsibility for carbon data collection

An organisation should establish clear accountabilities and ownership for collecting, storing and managing the data required for carbon emissions. Ascertaining and defining who will be responsible for what early on will help the organisation align to their carbon reduction goals and make their carbon accounting accurate and efficient in the long term.

Depending on the size of the organisation, the roles, responsibilities and teams assigned to carbon accounting will vary. A small organisation will have a different approach to a larger organisation with more employees.

The following departments are likely to be involved in the data collection process:

- **Leadership:** to lead the organisation with buy-in and support for support carbon reduction goals, and provide the resources and investment necessary to effectively implementing carbon initiatives.
- **Sustainability/ESG teams:** usually found in large organisations, and may often solely handle the carbon accounting and reporting process, with responsibilities set within and outside the organisation for collecting and managing data for carbon accounting; in smaller organisations there may be an individual or a group of individuals collaborating and collecting data from different entities.
- **Operations and logistics:** usually provides details or data around day-to-day operations that contribute to emissions, such as energy consumption, waste management, transport and water usage.
- **Finance and accounting:** all financial data required for spend-based calculations from this department, which also allocates and manages budget for activities relevant to emissions calculation.
- **Product design and management:** data on the entire life cycle of the product will come from this department.
- **Procurement and supply chain:** collaborates with suppliers and

supply chains to collect supplier and procurement data; they may also be able to influence suppliers to carry out data collection activities that eventually lead to their measuring and reducing their own carbon footprints.

- **Research and development:** provides data on relevant details, activities and software use that contributes to carbon emissions during the R&D phases of a project.
- **Human resources:** can provide details on employee commutes, hours spent in the office and at home, and business travel; they can also influence employees to record data and help drive employee engagement with sustainability efforts.
- **Marketing:** provides data on marketing activities that may contribute to carbon emissions.
- **IT and data management:** provides data on their own emissions-related data, and implements data collection systems to accurately manage, track and report carbon emissions.

Assigning responsibilities and roles is important, whatever the structure of the organisation. This requires collaboration and informed decision-making, and obtaining buy-in and engagement with all teams involved will help to ensure that data collected is of a high quality.

Technology and data collection

From our stakeholder interviews, it was evident that most organisations — whatever their size — gathered data manually and used spreadsheets. Scope 3 data from suppliers was usually collected through invoices — many now come with an environmental declaration of the product that is being used by the issuing company.

It was also noticeable that company practices differ significantly, depending on their size, and how advanced they are in their carbon accounting journey. Small companies and beginners use spreadsheets for tracking expenses for later carbon accounting. More sophisticated and larger companies have invested in the people and software to do this.

However data is collected, consistent and constant data collection is essential in order to track annual progress towards carbon reduction. As companies collect data and report emissions year after year, managing vast quantities of accumulating data can become a challenging and cumbersome process that runs the eventual risks of inaccuracy, inefficiency and even failure. Manual or semi-manual processes are prone to errors and lack standardisation.

Many organisations also fail to track their quarterly and yearly data progress towards carbon reduction goals. Although carbon reporting is a yearly assessment, tracking the organisation's carbon reduction goals should be undertaken on an ongoing basis. While this can be done relatively easily for scope 1 and 2 emissions, the granularity of data and time-consuming process involved in gathering scope 3 emissions information make it more challenging.

Software solutions

Using technology can streamline the carbon accounting process, enhance and automate the data collection process, and increase efficiency and accuracy of data collection — all of which support compliance with industry standards.

Software functionality usually enables organisations to:

- Automate data collection and calculations to reduce human error.
- Access precise real-time and historical energy data aligned to international GHG emissions standards, and that can be easily audited for compliance.
- Continuously monitor emissions data to track progress and adjust sustainability strategies as needed.
- Effectively manage data from various business operations, emission sources, and emissions factors.
- Scale seamlessly, with cloud-based or on-premises solutions that handle diverse data inputs, ensuring accurate and consistent calculations.

- Ensure compliance by integrating carbon accounting data with corporate financial reporting, effectively maintaining consistency and ensuring transparency and trust for stakeholders.

Whichever technology solution an organisation chooses to use, it should offer robust data processing functionality. The software should be able to support diverse carbon accounting methodologies, integrate and collect data from sources (for example, energy meters), track carbon reduction goals, perform audits, and run data checking for consistency and quality.

It needs to be able to perform calculations based on the standards and methods used industry-wide, and would ideally perform the following functions:

- Facilitate automated data capture and hierarchy management tools for GHG inventory boundaries.
- Provide access to nationally recognised carbon emissions factor datasets.
- Enable target tracking, with the ability to set and adjust baselines.
- Ensure global compatibility with multi-metric and multi-currency reporting.
- Calculate market-based and location-based scope 2 emissions according to internationally recognised protocols, such as the GHG Protocol.

Automating data collection

Technologies such as databases, IOT sensors, and artificial intelligence (AI) can make the complex task of carbon data collection much easier.

Organisations may use existing carbon accounting software and tools that can integrate with their existing systems and data thereby facilitating their data collection and in some cases automating the process, adhering to reporting requirements and automating calculation of emissions either through predefined emission factors or through real time retrieval of emission factors from a data source.

Using databases or data storage systems

Data collected should ideally be stored in databases or data storage systems, and made accessible to the required stakeholders. This can generally be facilitated using enterprise resource planning (ERP) systems, environmental management information systems (EMIS), distributed systems, or specialised software such as Carbon Disclosure Project (CDP) or GHG Protocol Corporate Standard tools. Other tools, like SAP Sustainability Control Tower or Enablon Sustainability, can help aggregate data across an organisation. Software like Procurify or Coupa can be used to track and analyse company spending, which can be correlated to carbon emissions.

There are several free resources available.

- The UK Department for Energy Security and Net Zero provides annual greenhouse gas conversion factors essential for reporting emissions.
- The National Atmospheric Emissions Inventory (NAEI) offers comprehensive emissions data for the UK.
- The Local Government Association's Greenhouse Gas Accounting Tool, free for councils in England, integrates with UK Government conversion factors to streamline reporting.
- Data.gov.uk offers a wide range of open datasets from UK public bodies, useful for environmental monitoring and carbon emissions analysis.
- Icebreaker One provides a comprehensive list of solutions for data collection (manual and automated).

Technology tools and software make it possible to integrate different data sources and types and centralise data collection. However, they can be complex and expensive to implement, and may involve a steep learning curve. When first collecting and storing data, an organisation may need to begin by storing, retrieving and managing data using a database or easily accessible and usable data storage system, and then later move on to data aggregation and centralisation using more complex solutions.

Example data models

Figure 13 shows an example of the data that needs to be collected during the construction process life cycle. It is a high-level conceptual diagram that can be the basis for the design of a data model or database schema.

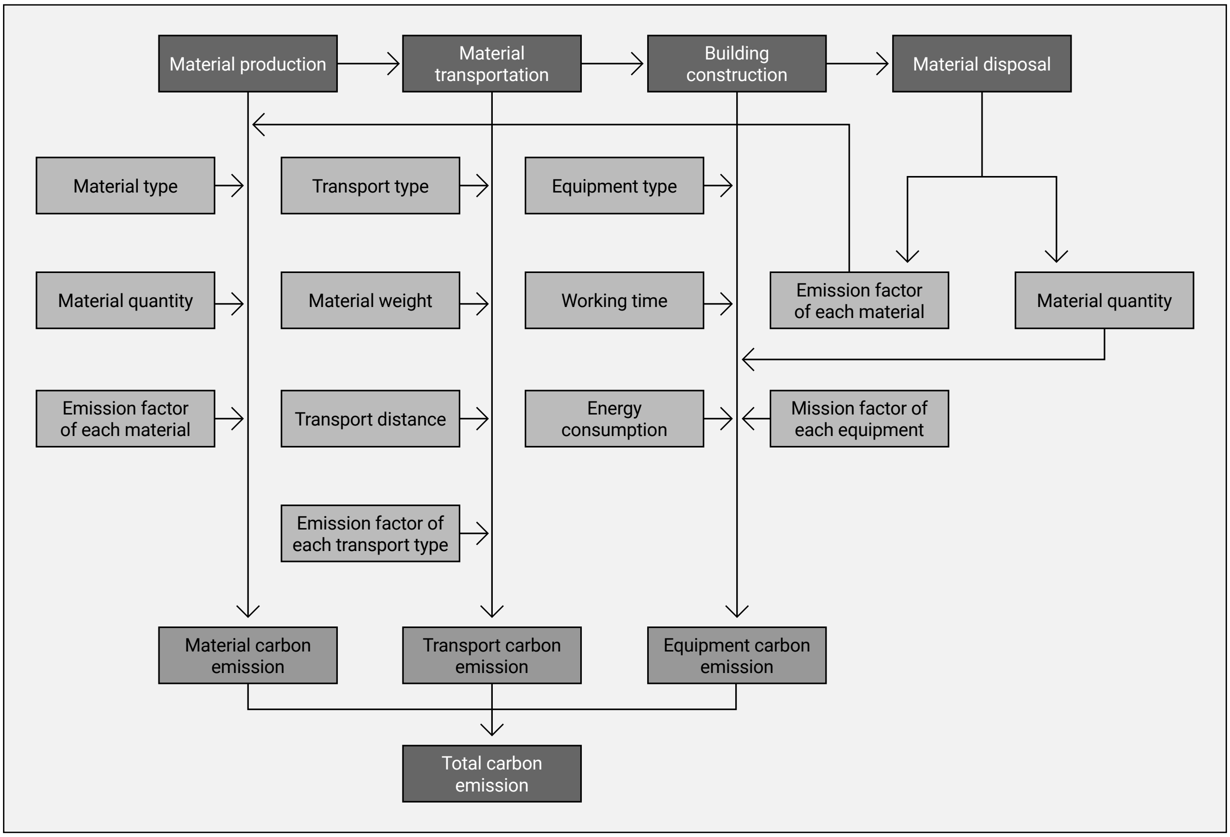


Figure 13: High-level conceptual diagram for carbon emissions in a construction process

The following two data model examples can also be used as reference sources by organisations starting to undertake carbon accounting or already on the journey. They can be used to design the carbon emissions database, and to understand the different types of data that need to be collected.

Data model 1

Net Zero Cloud Overview

Data model 2

Together for Sustainability data model to enable data exchange for the product carbon footprint.

Scope 3 data collection

Obtaining reliable and timely data from suppliers is one of the challenges in obtaining data for scope 3 emissions. To streamline and automate the process, APIs can connect an organisation's ERP systems or data centres with supplier systems. This streamlines data collection, removing the need for manual entry and consequent follow-ups, and ensures it happens consistently and in real time. This also streamlines and synchronises the process of gathering data from different geographical locations and across established organisational boundaries.

Using other emerging technologies

The advent of advanced digital technologies such as AI, blockchain, and the IOT has helped to automate data collection processes. These technologies can make data collection simpler, more efficient and consistent — they are automated and more accurate than manual data collection.

IOT sensors for real time data collection

Traditional methods as stated above mostly relied on utility bills or end of the month surveys to collect data. Using IOT sensors or AI-driven IOT sensors can help collect data in real time. For example a manufacturing company can deploy sensors in their factories or vehicles to collect the required data and store them directly into database systems. A waste management company can use sensors in their garbage collection vehicles and collect data around the amount and type of waste collected from different locations and understand the household behaviours from the waste generated, eating habits etc. The data collected can further be analysed to derive valuable insights and make decisions based on data.

Using AI with IOT sensors (smart sensors), can help with providing timely and accurate insights for decision making and also help with increasing the speed and frequency of data collection.

Using AI for intelligent carbon data collection

AI can transform carbon data collection for efficient emissions management. Through AI machine learning algorithms that can perform natural language processing (NLP) tasks, collected data can also be allocated or categorised and integrated with existing data, saving time.

Data analysis and carbon footprint calculation

To analyse the collected data against specific standards and to calculate the carbon footprint, carbon accounting software like Intelex, Carbon Trust Footprint Manager, or specialised tools that comply with various reporting standards (software that is compliant with ISO 14064, GHG Protocol, Task Force on Climate-Related Financial Disclosures (TCFD), Global Reporting Initiative (GRI) standards etc.) can be used.

DATA CLEANING

> Carbon emission data, whether internally sourced or received from the value chain, may not be immediately reusable.

The data collected may have missing values, incorrect values or duplicates, and may have to undergo a pre-processing or cleaning process in order for it to be of the quality required for effective use. Therefore data cleaning usually consists of normalising the data (standardising the format, and checking and correcting missing values and duplicates) and verifying the data. This process should be automated rather than manual.

Value of data cleaning

Having unclean (noisy) or incorrect data could skew calculations, resulting in incorrect conclusions and actions. Cleaning also helps to establish the level of accuracy of the dataset, which will help determine the percentage accuracy of any KPIs or studies that have been carried out.

When data is cleaned:

- There is more trust and confidence in the dataset.
- Identified errors can be corrected for the future.
- Data collection areas needing improvement can be identified.
- Decisions can be made what to include or exclude, and if the dataset suits the purpose.
- Better cost savings can be achieved.

Data normalisation

If using a database it is good practice to normalise the data so that there are less dependencies than otherwise needed. Databases should be structured in such a way that one attribute is defined at a time so that if unwanted data needs to be deleted it may be done so without having to delete data that it may have a dependency on.

Standardising the data

Data usually comes from different sources and presented in different ways, so it needs to be consistently formatted so that it can be treated as one dataset.

Establishing standardised naming conventions for data elements, categories, and emission sources helps prevent confusion and ensures that data is easily understood by all stakeholders. Standardised data reduces the likelihood of errors, discrepancies, and double counting, which can compromise the accuracy of emissions calculations and reporting. Standardisation ensures that data can be seamlessly integrated into modelling tools, software, and reporting systems, facilitating efficient carbon accounting processes.

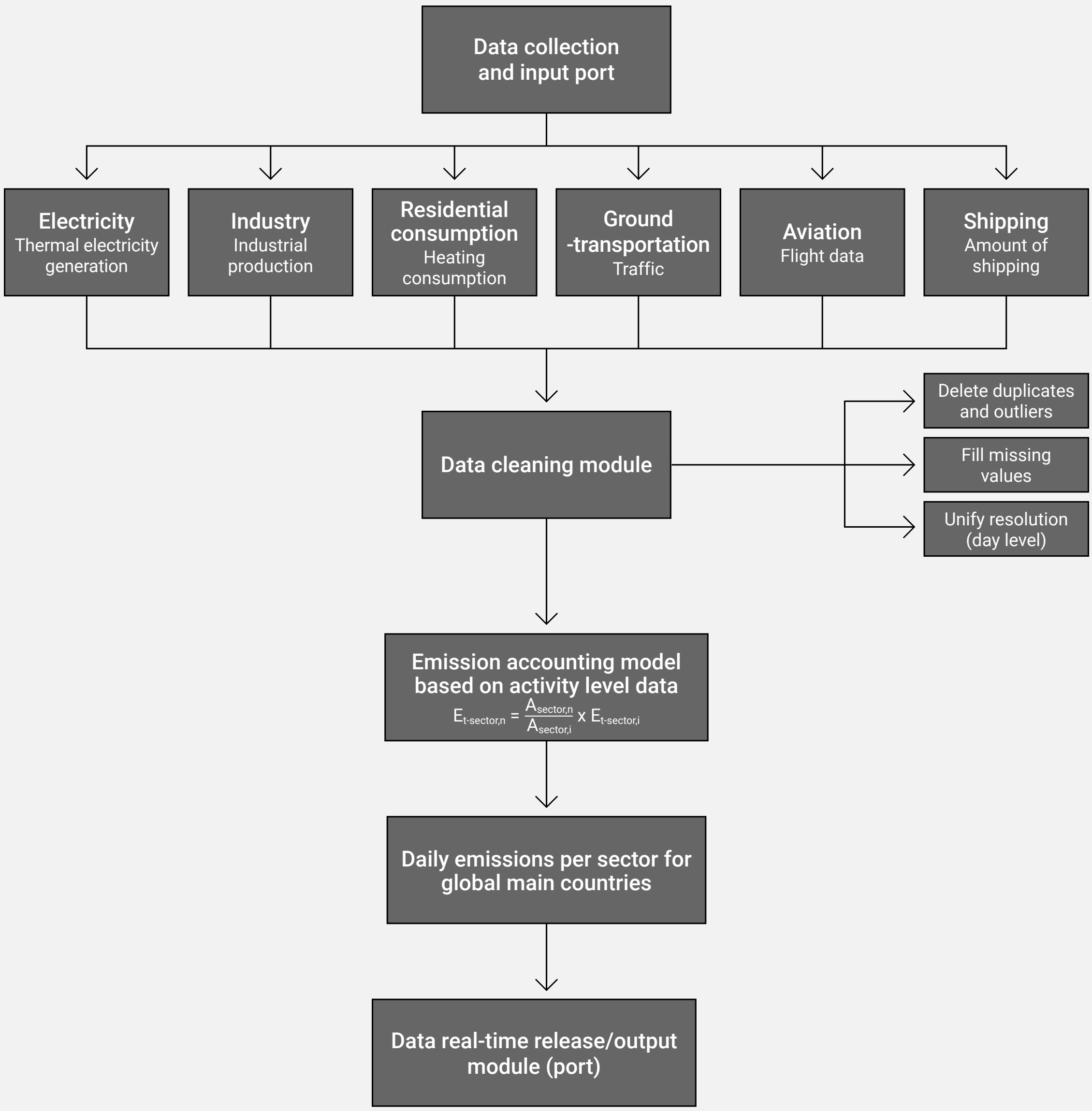
Deleting outliers and duplicates, adding missing values

After data is merged or integrated there could be duplicated records that need to be deleted. Missing values need to be updated to the correct format or values found elsewhere in the dataset — any remaining incomplete can be deleted. Any outliers or false values need to be deleted if they represent a scenario that can't have occurred.

Data validation

Once data normalisation is complete, it needs to be validated for other issues or inconsistencies to ensure it is within acceptable values or boundaries. If any more issues are found, they need to be analysed and fixed until the data reaches the required quality and state.

Figure 14 shows an example of the data cleaning process applied to a near real-time carbon emissions system framework.



Near-real-time carbon emission system framework. $E_{t\text{-}sector,n}$, $A_{sector,n}$, $A_{sector,i}$, $E_{t\text{-}sector,i}$ represent total emissions for one sector for the nth day, activity data for one sector for the nth day, activity data for one sector for the ith day, and total emissions for one sector for the ith day, respectively.

Figure 14: Example of a data cleaning module, from the paper by Liu et al, 2022.

Technology and data cleaning

The use of ERP systems or similar can help to automate the data collection process, but will not improve or standardise data quality. There are separate software solutions that can be used to clean the data and support integration. Tools such as SAS Data Quality, Informatica, Oracle Enterprise Data Quality, Melissa Clean suite, Alteryx, dplyr, pandas, and IBM Infosphere Datastage can be used to automate this process.

AI and machine learning technologies can identify anomalies and patterns to help automate the data cleaning process. They can also be used to detect other errors (such as spelling mistakes), suggest missing values, and identify and resolve data duplicates through the use of clustering algorithms.

DATA QUALITY

> For carbon emissions management mitigation actions to be effective and efficient, data should be accurate, reliable, consistent and updated in a timely manner for the analysis of trends and patterns.

The quality and representation of correct data matters more than the quantity of data, and it is important to have a clear purpose when designing the carbon data model. Attention should be given to the reliability of the data collection methods — devices used should be thoroughly examined. Only the most valuable data should be collected with accuracy and frequency.

Figure 15 provides an overview of the aspects of data quality that an organisation needs to consider when collecting, managing and storing data, as outlined in An Advanced Big Data Quality Framework Based on Weighted Metrics, 2022.

A single source of truth (SSoT) refers to a common, authoritative repository or database where all environmental and emissions data is stored, managed and maintained. SSoT promotes data consistency by ensuring that all emissions data is standardised to follow uniform data formats, units, and naming conventions. It helps maintain data integrity by minimising the risk of duplicate or conflicting data entries.

Quality aspects	Dimensions	Definition
Reliability	**Accuracy**	The degree to which data is reliable and describes real-world values.
	Uniqueness	Ensures that there are no duplicated records.
	Validity	Assures that data conform a specific format and complies with the defined business rules.
Availability	**Accessibility**	The Extent to which data is available and easily accessible.
	Security	Ensures that access to information is appropriately restricted.
Usability	**Ease of Manipulation**	The degree to which data could be used and manipulated for its intended use.
	Completeness	Assures that there are no missing values, and all the expected attributes have values.
	Readability	Refers to the ease of understanding of information.
Relevancy	**Freshness**	Refers to how recent and up-to-date the data is.
	Consistency	The extent to which data are coherent and does not contain contradictions.
	Credibility	Refers to how much data is credible and can be trusted.

Figure 15: Aspects of data quality

Data quality should adhere to the five characteristics or indicators specified by the GHG Protocol: **technological, temporal, geographical representativeness, completeness and reliability.**

Figures 16 and 17 show these five indicators and the scoring for each.

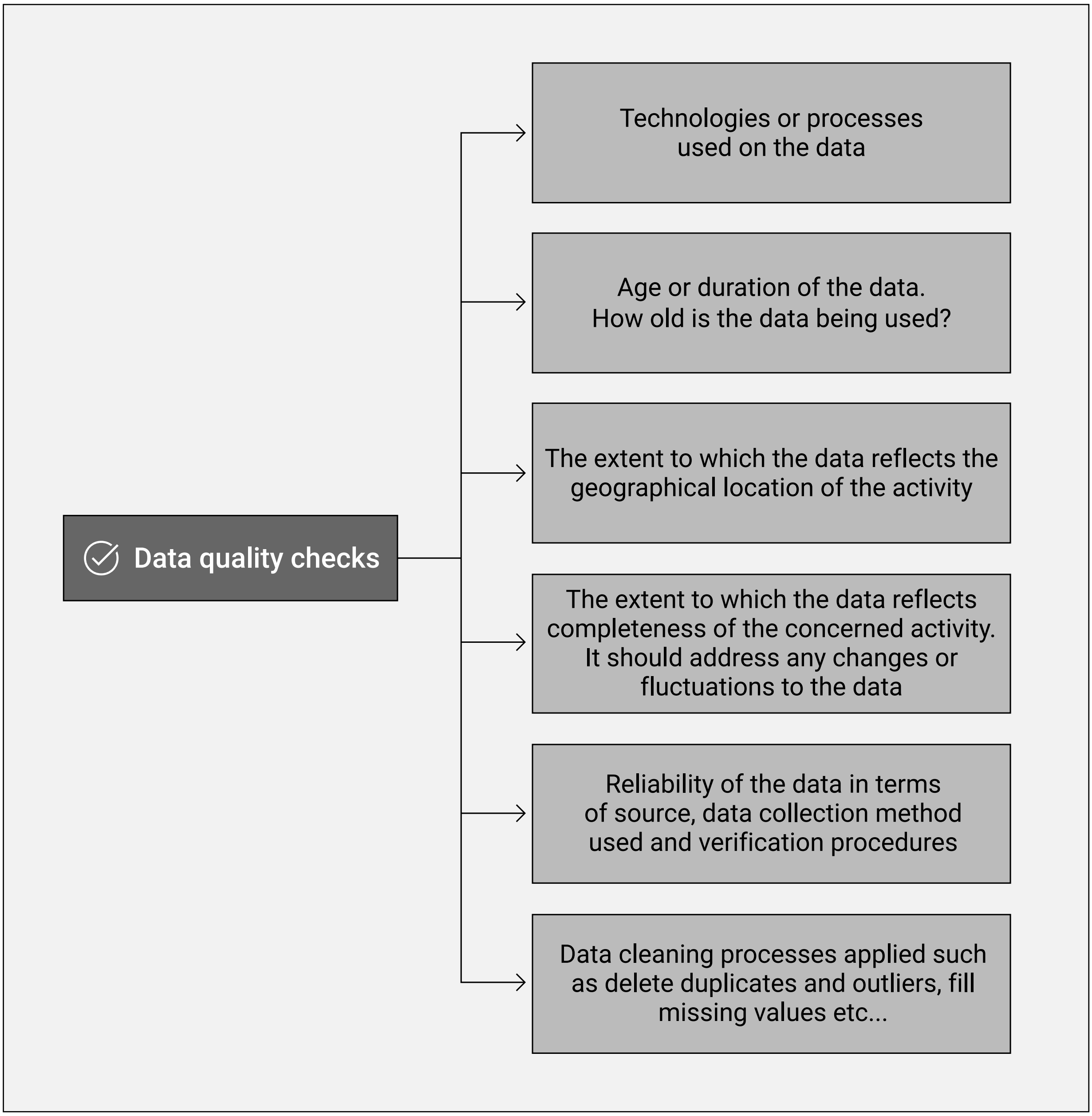


Figure 16: Data quality checks

Score	Representativeness to the activity in terms of:				
	Technology	Time	Geography	Completeness	Reliability
Very good	Data generated using the same technology.	Data from the same area.	Data from all relevant sites over an adequate time period to even out normal fluctuations.	Data from all relevant sites over an adequate time period to even out normal fluctuations.	Verified data based on measurements.
Good	Data generated using a similar but different technology.	Data with less than 6 years of difference.	Data from a similar area.	Data from more than 50 percent of sites for an adequate time period to even out normal fluctuations.	Verified data partly based on assumptions or non-verified data based on measurements.
Fair	Data generated using a different technology.	Data with less than 10 years of difference.	Data from a different area.	Data from less than 50 percent of sites for an adequate time period to even out normal fluctuations or more than 50 percent of sites but for a shorter time period.	Non-verified data partly based on assumptions, or a qualified estimate (e.g. by a sector expert).
Poor	Data where technology is unknown.	Data with more than 10 years of difference or the age of the data are unknown.	Data from an area that is unknown.	Data from less than 50 percent of sites for shorter time period or representativeness is unknown.	Non-qualified estimate.

Figure 17: Data quality scoring

Improving data quality

Improving data quality gradually over a period of time is important, as this will help the organisation to ensure accountability, transparency and alignment to their carbon accounting goals.

The following steps can improve data quality, as outlined in this Guide from xTonnes: Exploring data sources for carbon accounting.

Employee education

Everyone involved in carbon accounting should be educated on data collection, quality and reporting processes and the importance of timely and accurate data, so they know how their actions contribute to the organisation's sustainability goals.

Regulate data collection processes

To establish standards and procedures across the organisation, data collection responsibilities, methods and sources should be clearly defined. Policies and processes need to be documented and followed, so that transparency and consistency is ensured and comparisons can be made if needed.

Data validation and verification

Validate and verify the accuracy and completeness of data by carrying out regular audits, identifying any discrepancies by comparing data between time periods and reliable sources, and rectifying them as needed. To verify emissions, use the help of independent third-party auditors to get credible certifications.

Integrate feedback loops

Establish an environment of fostering data transparency and open communication. A process of regular feedback loops ensures that continuous improvements in data collection and quality can be made, and employees should be encouraged to report any discrepancies.

Use of technology

Using available carbon accounting software to track and manage data accuracy and quality also automates the process of managing data, reducing manual errors and saving time and costs. Such tools can also ensure that the latest emission factors and calculations are used to calculate carbon emissions.

Technology and data quality

Machine learning algorithms enhance carbon data quality by detecting irregularities in energy consumption or emissions. Natural language processing tools analyse supplier reports, internal documents and regulatory requirements to extract necessary data. Integrating IoT sensors with artificial intelligence and edge computing allows real-time processing of high-quality data at the source, ensuring accuracy and negating the need for subsequent verification. An example of such integration can be seen in an example of machine learning pipelines using recurrent patterns of IOT data to detect sensor drift and repair any erroneous or corrupt data.

Off-the-shelf data quality tools can help in automating quality control and remediation through features like profiling, metadata management, matching and monitoring. Examples include SAP, Informatica, IBM Infosphere, Talend, and SAS.

DATA VALIDATION AND VERIFICATION

> All existing and proposed emission reduction programmes require some kind of verification or certification of self-reported data, and there will be differences in how these verification processes are implemented.

Data validation and verification are essential to ensure the accuracy and reliability of greenhouse gas (GHG) emissions and environmental data, and involve checking data for accuracy, completeness, and consistency. This helps identify errors, anomalies, and discrepancies in the data, ensuring that only high-quality data is used in emissions calculations, and helps establish data trustworthiness. The report by Satellite Applications Catapult covers the application of remote sensing technologies in carbon accounting, so that data can be validated and verified and trust can be built.

Developing carbon accounting data verification processes and structures that have credibility and are efficient is one of the most important challenges that all countries currently face. Therefore when developing such a system it is important to consider the principles of truthful disclosure, consistency, transparency, ethical conduct independence, and due professional care. The Role of Third-Party Verification in Emissions Trading Systems highlights these principles and provides the verification best practices followed by governmental bodies such as the US EPA and EU-ETS in order to conduct and monitor their processes.

Data validation processes

- Maintaining detailed audit trails that document data validation and verification activities is good practice to ensure data traceability and trust in the quality of the data. Such trails provide a record of how data was validated, who conducted the validation, and results of the verification process.
- It's important that emission factors and methodologies used for emissions calculations are verified against recognised standards and best practices, ensuring they are up to date and applicable.
- Cross-checking data from different sources helps to ensure consistency (for example, comparing energy consumption data from utility bills with data from internal meters).
- Periodic internal audits of data management processes and emissions calculations and the type of data used will help to identify areas for improvement and maintain data quality over time.

The findings from data validation and verification processes should be used to continuously improve data collection, reporting methodologies, and emissions reduction strategies.

Third-party verification

Independent verification by a third party, such as an external auditor or expert, is a crucial step, although it is not a legal requirement. An independent review and confirmation of the accuracy of emissions data and calculations adds credibility to emissions reports. However, independent third-party verification can be costly, and as there are multiple different methods currently in use, it's important that the chosen party should use credible and recognised methods that will help to establish trust. Successful validation and verification processes usually end with the third-party auditor or verifier providing a limited or reasonable statement of assurance. Limited assurance means that the nature and extent to which verification activities were conducted fall within a reduced assurance level. Reasonable assurance means that the nature and extent to which verification activities were conducted fall within an assurance level that is high but may not be absolute.

Data lineage

Data in an organisation may come from — and be manually shared between — different departments. Data lineage is the process of tracking and visualising how data moves through an organisation, from its source to its destination, and if there is no clear process in place, then the organisation will not only be not able to validate and verify the data. This means that it will not be able to establish trust in the data and its quality.

Tracking and recording changes made to the data over time, and the business reasons behind those changes, is a good practice for understanding data lineage.

There are many solutions for tracking the lineage of data. A data catalogue can be used to track internal datasets and the relationships between different data systems. However, these may not be ideal for ad hoc data, or data from external sources such as value chains. This is where a blockchain system can be used for trustless transactions — those that do not require trust in a third party. Once the data is entered into the system, it cannot be tampered or modified with. The concept of trustless data can be used to reduce and eventually eliminate the reporting of incorrect data.

Data lineage tools like Manta, Atlan, and Talend focus on visualising and tracking data flow and transformations across complex systems. Version control systems (such as Git and DVC) track changes to datasets over time, promoting transparency and collaboration. Master data management tools (like SAP Master Data Governance and Profisee) ensure the consistency and accuracy of key data assets, providing a single source of truth with lineage tracking.

Data verification/validation and technology

Usually, organisations manually cross-reference new emission data against existing industry standards or databases. Technology solutions such as AI can automate this verification process, validating the accuracy of new data entering the system against existing industry standards, such as the Global Reporting Initiative (GRI) or the EPA's own emission control system and factors. This saves time and minimises the risk of errors.

It is important to generate auditable footprint reports that ensure the quality of data. Each state of data management, from ingestion to reporting, needs to be put through rigorous validation processes. This also needs to be reviewed by the organisation's stakeholders to validate and confirm the alignment of the outputs to business goals. Visualisation can help to achieve this, for example, by displaying a trend analysis over a monthly or quarterly cycle.

Data quality tools

In addition to checks that an organisation may perform on data sources, they should also be using data quality tools to detect anomalies at the pre-processing stage, ensure that the quality of activity data is not affected by data transformations, metadata values are intact, and that no records are either duplicated or deleted during the validation process. These can all be tracked via a dashboard system or alerts system, using business intelligence and reporting tools such as Tableau or Power BI, or sustainability reporting software that can handle complex data sets and generate comprehensive reports and dashboards.

Other audit management solutions that specialise in environmental compliance are available, such as ISO 14001 audit tools, or sustainability assurance services from third-party providers. Using standardised templates, datasets and transfer methods can also make verification easier.

Double counting can also be addressed through the use of better, cheaper and simpler auditing techniques and capabilities, or, as this article by the Carbon Data Trust explains, by using blockchain technology.

e-liability

E-liability (that is, environmental liability) is an accounting system that addresses issues of transparency and double accounting in the supply chain by using the principles of cost accounting, allocating carbon emissions across different stages in production.

Any inputs of raw materials or intermediary goods that make up a product carry an e-liability. This means an organisation can obtain a comprehensive picture of its emissions when it adds these inherited e-liabilities to its own scope 1 emissions and the carbon cost associated with production of the product.

Traditional MRV (monitoring, reporting and verification) systems such as the EU-ETS, Vera, and Gold Standard are currently characterised by manual data collection, soft copy reporting, and long or annual reporting cycles. They are also heavily centralised (resulting in a lack of transparency), are more prone to errors, time-consuming, and may become costly over time.

To overcome these challenges, digital monitoring, reporting and verification (D-MRV) systems were introduced: second generation MRV systems that are envisioned to be integrated, fully automated carbon accounting systems that are built using service-oriented sensing technologies, smart contracts for automating data collection, and monitoring and distributed ledger technologies (such as blockchain) for security and transparency for emission registrations. Verification processes can be automated using AI for pre-conditioning. While such a system may initially involve costs initially, they are usually more cost-effective and efficient in the long term.

An example of EU-ETS systems that may be built using distributed ledger technology. Towards greener trade and global supply chain environmental accounting includes further examples of such systems, in the context of both DMRV systems and carbon tracing in supply chains. The article Emerging digital MRV tools shape a new climate information ecosystem includes links to some notable workshops and solutions.

However, D-MRV systems do not solve the verification problem entirely. Verifiable Carbon Accounting in Supply Chains proposes an example of a blockchain-based verifiable carbon accounting system as an extension to the digital monitoring, reporting and verification system, in order to improve PCF verifiability and apply D-MRV principles and systems to emissions calculations as well as data sources.

DATA SHARING AND INTEROPERABILITY

> Data interoperability is the ability of different systems or components to communicate and operate together seamlessly, even if they are from different vendors or use different technologies.

It ensures that emissions data can be integrated from various sources, such as energy consumption data from utility companies, transportation data from logistic partners and operational data from internal systems.

Data sharing refers to the ease with which environmental data can be moved or shared between different databases or platforms. It allows organisations to collect data using various tools, store it in different databases, and share it with partners or regulatory authorities without compatibility issues.

For effective scope 3 emissions data sharing and fusion between the different stakeholders in a value chain, it is important that carbon data is compatible or interoperable, consistent and comparable over a period of time. Requirements for a carbon data model should therefore be interoperability across different data platforms, sharing or transferability between different stakeholders and compatibility across different devices and versions. A desirable or ideal data environment would be common, work with open data formats, and follow standardised data protocols (as described in A carbon data trustworthiness framework for the construction sector).

Having a common data environment enables different organisations and stakeholders to share or exchange data, increasing collaboration and coordination. Using open data formats makes data easier to access by a range of users, increasing transparency and accountability, and helping to build and establish trust in carbon data.

A comprehensive and reliable picture of carbon data management can be created by having a set of standardised data protocols that would ensure easy data integration from different data sources.

The consistency and quality of data can be improved by standardising data protocols and the formats being used. This will also reduce the risk of errors occurring and make the data more reliable. Eventually, the usage of open data formats and standardised data protocols will not only help to reduce costs associated with carbon emission data collection, analysis, storage and management but will also help to reduce costs when integrating data from different sources.

To achieve data sharing and interoperability, environmental data can be structured using standardised formats such as XBRL, XML, and JSON.

Financial reporting and XBRL

Eventually, reporting carbon emissions will require alignment with financial reporting, as the sustainability efforts of an organisation will have financial implications, and vice versa. The underlying principles of financial practice ensure consistency checks and direct comparability for any two sets of accounts since both have been created using the same standard principles and methodology.

Some of the underlying principles that govern the financial domain are as follows:

- Financial reporting is essential for every organisation and is embedded through mandates and regulations — adherence to standards and regulations is required of every entity.
- Income and expenses are measured consistently in all organisations, irrespective of when and how they occur, usually involving the use of standard tools.
- Data can be transferred or shared between one entity and another using recognised and available standard formats.
- Reporting mechanisms are all standardised, and all audits follow standard rules.
- Skilled resources are widely available, with a single governing body — the Financial Reporting Council (FRC) — overseeing recognised accounting qualifications; this means that everyone works the same way and the skills used are the same in all organisations, aligned to the requirements of the sector's reporting systems.

XBRL (eXtensible Business Reporting Language) is a standardised markup language primarily used for financial reporting and the exchange of financial data between different software applications and systems. A method similar to that used in financial accounting to approach carbon accounting.

The 2016 paper by Seele on digitally unified reporting reviews the use of XBRL format for reporting financial and business information. This can be done using a common data repository from where the integrated XBRL report data can be obtained in real time. common repository. Using XBRL can help to fill current sustainability gaps in transparency, accountability, data exchange and interoperability.

An example of automated data sharing and sustainability reporting via a trust framework for scope 2 emissions in banking has been developed by IceBreakerOne. Another example is the open banking programme which standardised the sharing of data between banks to increase transparency.

Energy and the digital spine

The energy sector is one of the significant creators of carbon emissions, and their proposed technology developments include many improvements that should be considered as recommendations for the carbon accounting data framework. Energy Systems Catapult has recently outlined the concept for a digital spine for the energy industry. A digital spine is a thin interoperable layer that forms a connection or one that establishes communication between different organisational systems, hardware and software forming a grid in real time.

Existing systems work in fragmented environments and in silos that limit information sharing. Digital spine solves this problem by providing an open access and cohesive infrastructure that can be jointly governed as well as operated for the public good. Implementing a digital spine can address the issues raised around data interoperability and sharing.

A digital spine supports interoperability and standardised data sharing in three key ways, which are subject to technical and governance considerations:

- **Data preparation** involving a containerised solution that formats data to a given standard, creating an interoperable data-sharing infrastructure.
- **Ecosystem trust**, created by a governance or trust framework for data sharing, with defined roles, responsibilities and security controls.
- **Data sharing** between two entities, standardised according to technical capabilities and associated governance.

Product carbon data exchange initiatives

Credible emissions data is a key enabler to support policymakers, investors, and industry to make low carbon decisions and track progress towards net zero. An example data sharing system would be a distributed approach to data management and sharing of information between data generators and data consumers that can support different carbon accounting requirements.

Guidance and initiatives for product carbon footprint (PCF) data exchange have been provided by TFS and WBCSD. Through the PACT framework, WBCSD highlights the data model as well as technical specifications for data sharing.

More examples of data exchange and interoperability for product scope 3 upstream emissions and the value chain can be found in Supply-chain data sharing for scope 3 emissions, which references initiatives such as OS-Climate, Catena-X Automotive Network EV and Smart Freight Center.

It is important that data sharing initiatives remain open to further collaboration, so that eventually there can be one nationally harmonised standard across industries for easier exchange of emission data and consistency in data measurement.

From the Energy Systems Catapult report, Carbon accounting and standards in industry: a framework for innovation and growth, Figure 18 depicts an example of data generators and consumers and how they can work together in a distributed environment to share data.

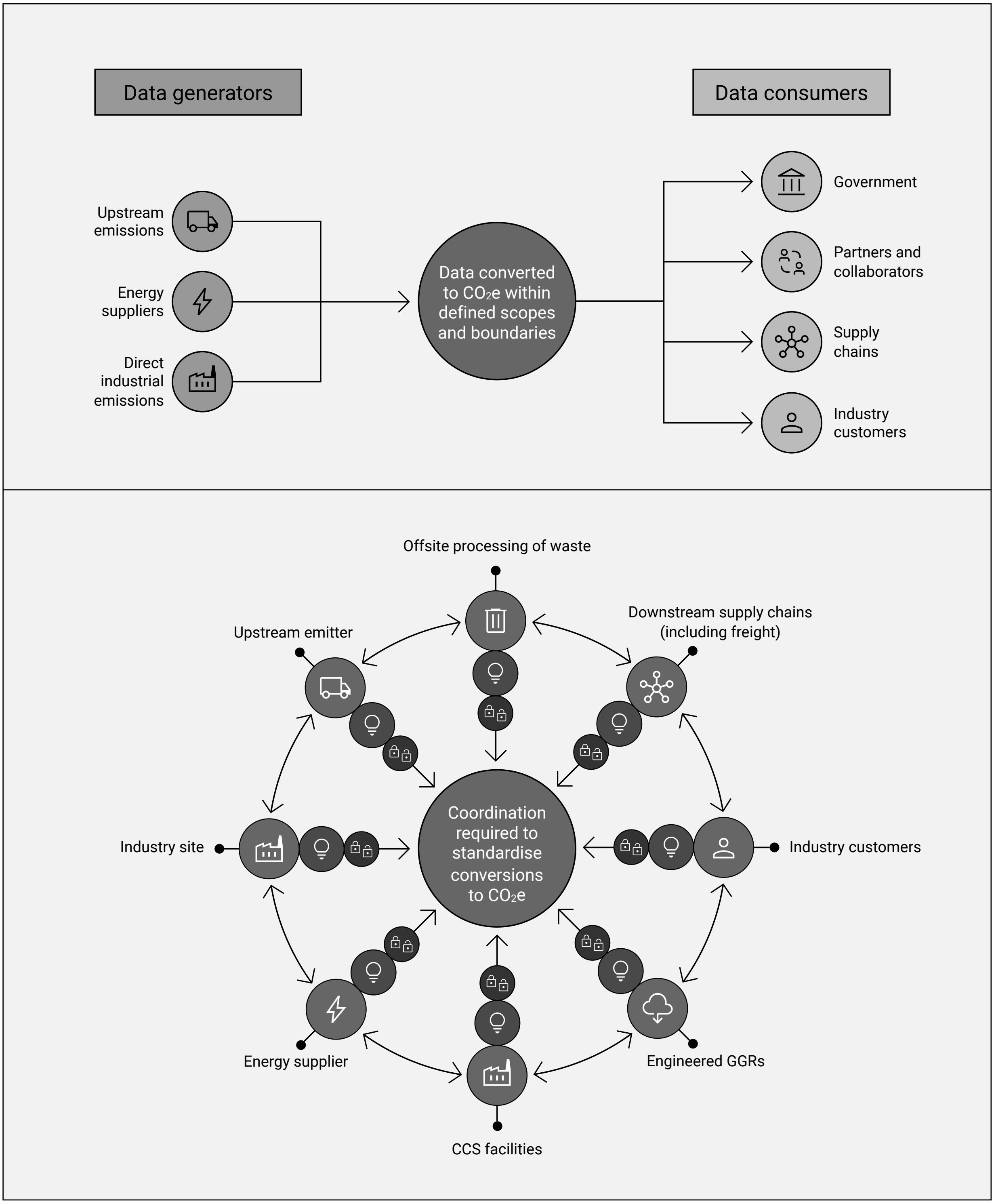


Figure 18: Distributed data approach for data sharing between generators and users

DATA CALCULATION

> There are three different methodologies that can be used to calculate carbon emissions, depending on the availability of data: activity data-based calculations, spend-based calculations and average data-based calculations.

Activity data-based calculations

This calculation method focuses mainly on the use of primary or activity data to calculate carbon dioxide emissions. It measures those activities that directly contribute to the organisation's emissions. This approach would generally require the organisation to carry out a detailed analysis or examination of the activities such as transportation records, waste generation records, energy consumption etc. The emission factors used are more generally more specific to the organisation's activities and therefore can be used to make more precise calculations.

For example, to understand or model the emissions from the organisation's use of vehicles, data on the type of vehicle, distance travelled, fuel consumption and vehicle weight will be needed to calculate the emissions. Another example would be on the goods purchased from the supplier wherein the organisation would provide information on the emissions specific to the supplier (which should be usually provided by the supplier if possible) and the quantity of goods purchased.

The method should be the one generally used for carbon calculations due to the accuracy it may provide as the data comes from the organisation's specific activities and emission factors. It would also help drive better decision making and open up a whole set of actions that the organisation can take in the future. But to use this methodology is also challenging as primary data is not always available and the process of data collection is a difficult, time consuming and resource intensive process.

Spend data-based calculations

The spend based method of calculation calculates the emissions using data from the organisation's financial spend (for example the amount spent on fuel during transportation can be used to calculate emissions). There is a correlation between the amount of money spent to the emissions generated, and the spend is multiplied by relevant emission factors to derive the final calculation. This method results in secondary data.

This is one of the easiest methods of calculating emissions, as spend data is readily available and calculations can be done quickly. It is less accurate than activity data-based calculation, but more accurate than using average data. For organisations having to deal with complex supply chains, obtaining accurate data may not be easy, and spend-based calculations may be useful for measuring emissions. They are also useful as a first step for identifying emission hotspots in an organisation starting out on their carbon accounting journey.

Average data-based calculations

This calculation method also uses secondary data to calculate emissions. Owing to the lack of primary or activity-specific data, secondary data (such as estimates, sector emissions or industry averages) are used when calculating emissions for a service, activity or product. For example, vehicle emissions may be calculated using the available industry average emission factor, or employee commutes calculated using an estimate based on average distance travelled, typical mode of transport, and average number of commuting days per week.

Using average data makes the calculation generic, and the least accurate approach, but can be useful when a rough estimate is needed, the required information is not available, or is a challenge to obtain.

Comparison of data calculation methods			
Feature	**Activity data-based**	**Spend data-based**	**Average data-based**
Data required	Primary data or detailed data specific to the activities internal and external to the organisation.	Secondary data or financial spend-based data that can be found in bills, invoices etc.	Estimated or secondary data.
Emission factors	Specific to the activity or operation.	Averaged industry or sector-specific values.	Averaged industry or sector-specific values.
Complexity	Determining an emission factor for a particular activity or product may not be easy when multiple variables and varying factors are used.	Financial data is usually readily available (such as purchases made in service industry organisations).	Data is easily accessed, allowing rapid estimation of carbon emissions to identify hotspots.
Accuracy	Data is accurate and precise due to the granular and detailed nature, allowing detailed analysis as each activity is linked to its respective emission factor.	A direct relation is assumed between financial spend and emission factors, but this may be more complex — accurately attributing an emission to a spending category may not be possible.	Using average or estimated data may not accurately reflect emissions; this relies on assumptions and generalisations.
Applicability	Identifying emission hotspots to target (areas where emission reductions are possible).	Useful when organisations have to make estimations for their supply chains.	Average emission factors can be applied to a wide range of situations; useful for organisations without direct control of production or supply chains.
Relevance to business operations	Delivers actionable insights, identifying specific activities and processes that are contributing largely to emissions.	Reflecting on the financial expenditure and purchasing decisions that affect carbon decision making.	Spending is usually fully tracked, making it easy to monitor emissions via expenditure over time.
Regulatory compliance	May be required by regulations or industry-specific standards.	May be required by regulations or industry-specific standards.	May be required by regulations or industry-specific standards.
Reporting	Greater accuracy when reporting emissions.	Less accurate than activity-based.	Less accurate than spend- based.

All three methods can be used individually or in hybrid combination, according to data availability. Depending on the type of method used, the sources of data may vary — for example, organisations may use secondary data methods to gather a rough estimate of their emissions first, then move on to using primary data methods, depending on their time and resource constraints. More on the different types of methods used for emissions calculation for different scope 3 categories can be found in the [GHG Protocol Scope 3 Calculation Guidance].

For example, for the scope 3 category of 'purchased goods and services' there are four methods that can be used: supplier-specific, spend-based, hybrid, and the average data method. Figure 19 shows an example of a decision tree for selecting the calculation method.

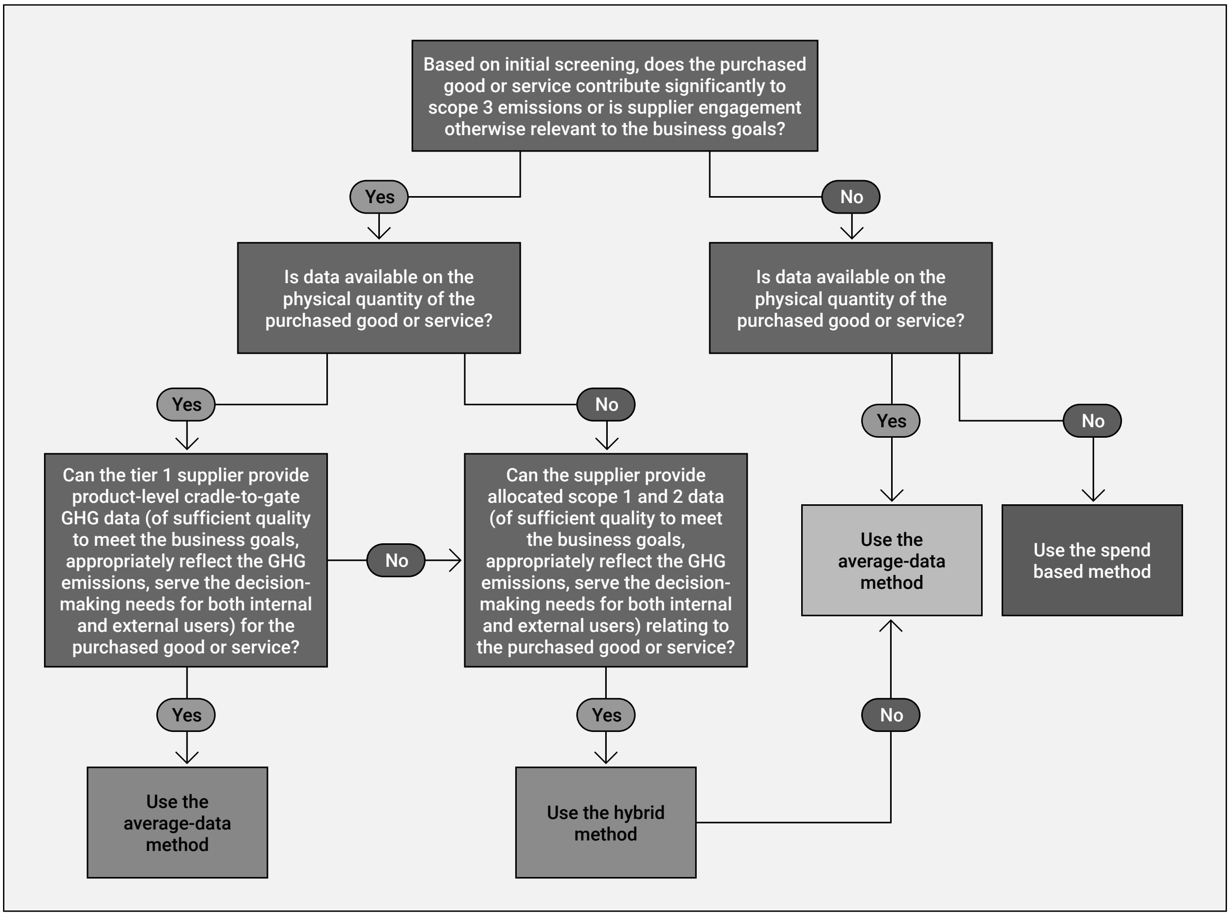


Figure 19: Decision tree for selecting the calculation method for emissions from purchased goods and services

Data accuracy and specificity

Data used for supplier-specific and hybrid methods may more accurately represent an individual supplier than spend-based or average data methods. But the results produced may still not be an accurate representation or reflection of the contribution made by the product to the company that is reporting the scope 3 emissions. At times, data collected from a supplier can be even less accurate than industry average data, and a considerable degree of uncertainty can be involved when allocating a supplier's emissions to a product purchase, depending on the type of allocation method used.

Data accuracy is generally derived from the granularity of emissions data, the reliability of the data sources (in this case, the supplier) and any allocation techniques used.

DATA PRIVACY AND SECURITY

> One of the most important aspects of managing data is its privacy and security.

Carbon data may include data relating to people, company strategies, technologies being used, costs and finance, all of which could be considered as confidential to a business. Whether primary or secondary spend-based data, it should still be considered as sensitive and will need privacy and security when sharing, so that only the required information is gathered and managed, and access is restricted to the relevant parties.

Data privacy and security technology

Many technologies can be used to ensure data security and privacy, including distributed data storage techniques, such as distributed ledgers and smart contracts. Distributed ledger technologies (DLT) such as blockchain are now being used to ensure data security and privacy protection. They also avoid the risk of a single point of failure associated with centralised data stores by using cloud-based platforms to store copies of data in multiple secure locations.

DLT enables and establishes data trust when sharing, helping an organisation to calculate its carbon footprint accurately and also to estimate the carbon sensitivity and intensity. If needed, external verification is also made possible when data is shared securely.

Data privacy can be achieved using encryption technologies. Encryption increases the willingness of organisations to share data, as it gives them control over access, reducing risk.

Recent innovations in cryptography include homomorphic encryption (HE), a set of algorithms that allows users to perform mathematical operations on encrypted data without needing to decrypt it first. HE is made possible due to increasing compute power, and examples include zero-knowledge proofs and secure multi-party computing.

- Using **zero-knowledge proofs**, an organisation can prove that they meet or satisfy given criteria without revealing confidential information — this could include proving that the carbon footprint of a product meets or is below a determined threshold.
- **Secure multi-party computing** allows computation or analysis to be carried out on combined data without either of the parties having to reveal their private input.

Data privacy can also be achieved by putting secure role-based data access in place, so that sharing, reporting and analysis rights are given to designated parties only.

Data trusts

Another way of sharing data securely is by appointing a data trust. Data trusts are independent and neutral parties that facilitate data sharing between parties while protecting privacy. Data trusts support trustworthy data processing, and provide ethical, governance and architectural support within the boundaries of the law. They also need to provide the technical tools, such as encryption technologies, to secure and protect the data they hold. Examples of data trusts include Catena-X, SINE Foundation, and OS-Climate which were all established as non-profit organisations to enable sharing and interoperability of carbon data.

Cyberthreats

Cyber-attacks present a security risk to all data, and a solution proposed by PACT is for organisations to decentralise their data and enable peer-to-peer sharing, which could decrease risk due to hackers and cyber-attacks. Using data trusts, encryption technologies, decentralised data storage, and cybersecurity technologies, processes and practices can enable organisations to have better privacy and security thereby mitigating their data privacy concerns when sharing data.

Data privacy and security are included in most of the holistic carbon accounting platforms or tools. Some specific tools with a greater focus on privacy and security of data are the Microsoft Sustainability Cloud, which leverages Microsoft's AI and data capabilities to enhance data visibility and security, and the OneTrust ESG Module, which provides interactive dashboards for measuring logistics carbon footprints and analysing data across various transport modes. The OneTrust ESG Module ensures data privacy through compliance with frameworks like the GLEC Framework and DIN EN 16258 European Standard.

CONCLUSION

The different pillars of the data recommendation framework and their respective principles can help in the generation of carbon management data that is of good quality and trustworthy, and can serve as a guide when planning and designing carbon data models in an organisation.

Any industry should start with defining guidelines, principles, roles and responsibilities and procedures for carbon data management at sector level and then extend these to other related sectors so as to increase compatibility between sectors.

Leveraging advanced digital technologies

Within this report, tools and technologies for managing carbon data have been mentioned, and wherever possible, the most recent innovations should be used to automate data collection processes, such as data models, remote sensing, sensing devices, IoT, computer vision and AI.

Advanced analytics techniques, such as big data analytics, AI and ML, and simulation technologies for activities such as automated carbon calculations, forecasting and projections, carbon auditing, and cost benefit analysis can be undertaken using trustworthy carbon data to support an organisation's carbon management decision-making.

While the cost and resource burden associated with investing in and implementing tools is something that not all organisations can easily, effective carbon accounting data management will be a requirement in the long-term as data accumulates. Organisations unable to adopt new tools straight away should start with manual data processes, and slowly transition from manual to automated processes as the size and complexity of both the data and the business grow.

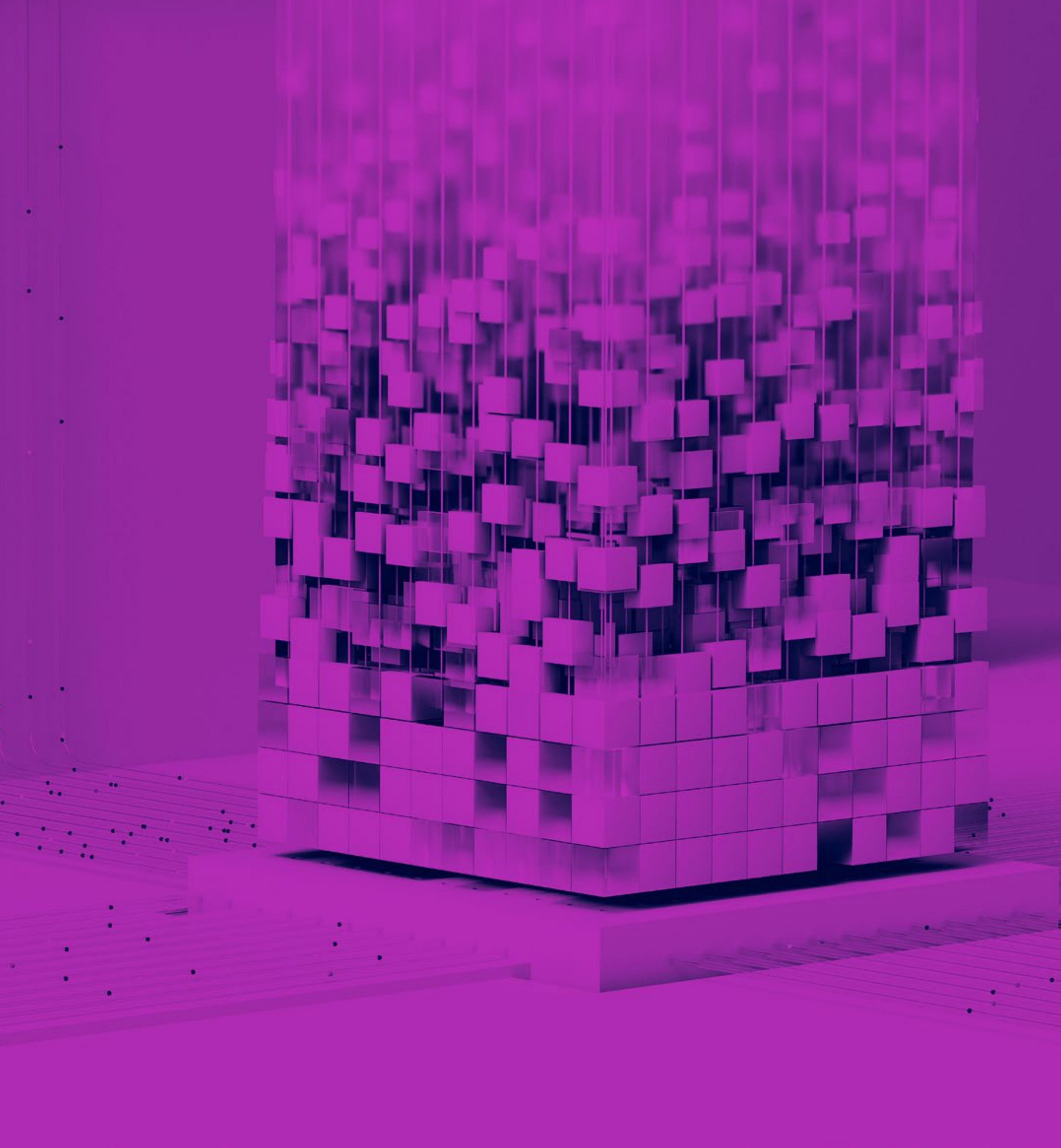
Next steps and further work

With the help of the recommendations provided in this document and the outputs of the CCCA programme as a whole, organisations can improve their data management processes and practices, particularly in terms of carbon data. This would collectively improve their ability to monitor, track and report carbon emissions, while promoting transparency across industry when it comes to sharing data and establishing trust.

Organisations can then make better-informed decisions to help them achieve their sustainability goals, contributing to efforts being made in tackling climate change.

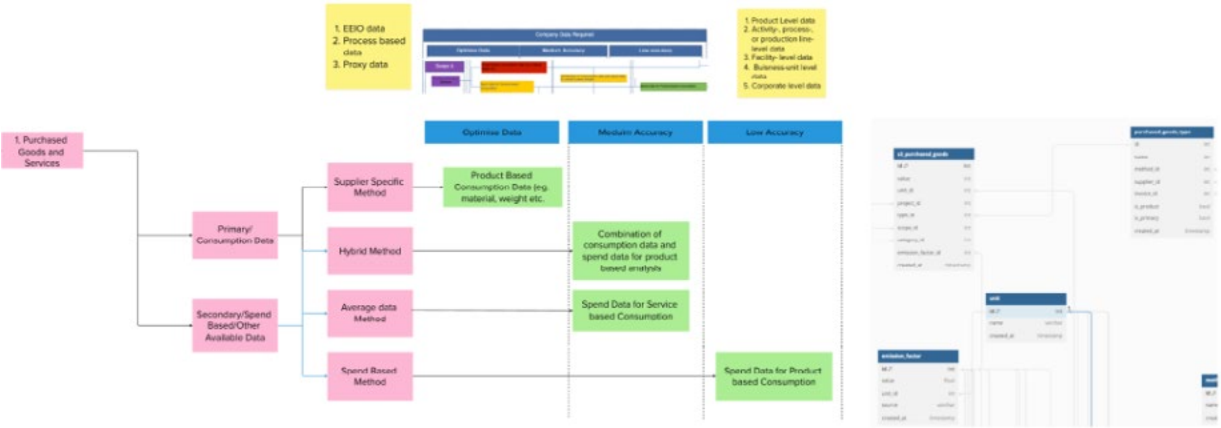
For carbon accounting data management recommendations to be effectively implemented, robust policies and processes need to be put in place. This can be made possible with the help of carbon regulators, sector and industry-specific guidance and processes, and guidance for how all the policies, regulations and processes will tie up at the national level. Streamlining policies and process frameworks will make it easier to implement and strengthen data management processes.

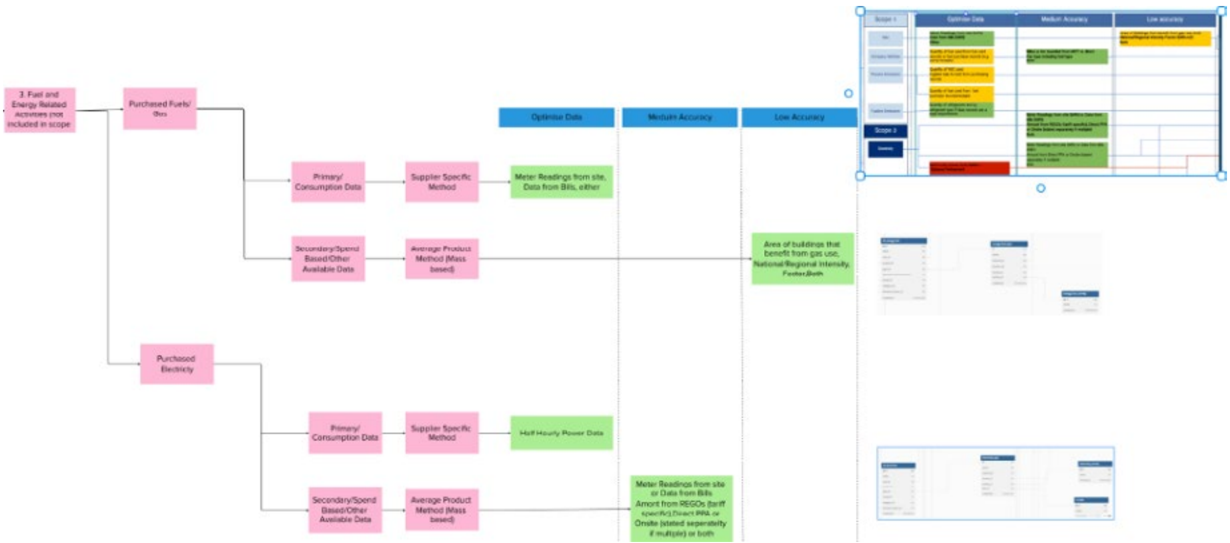
APPENDIX



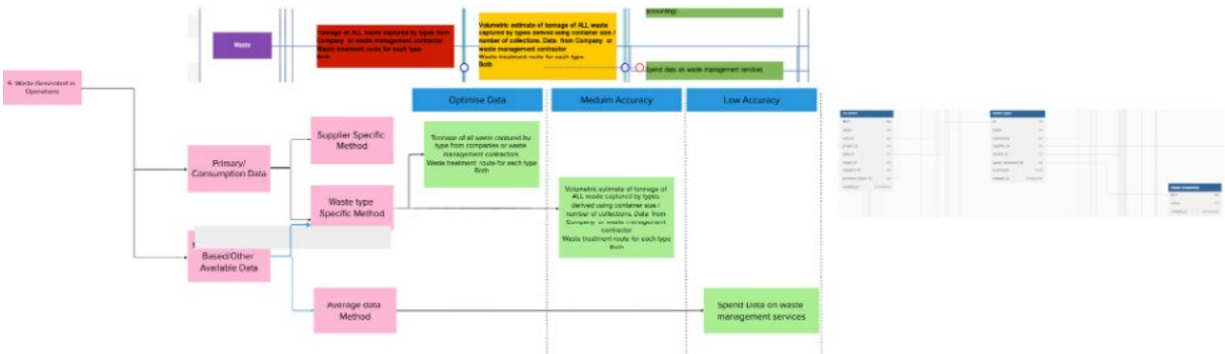
Supporting workflows

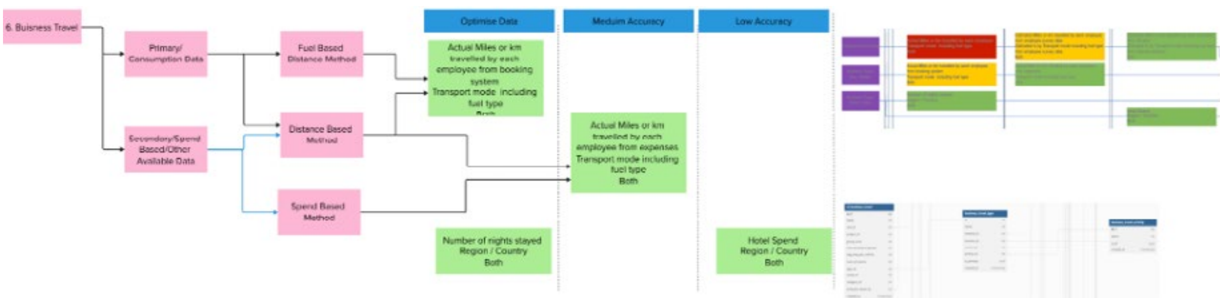
The following images are stored on Mural, and can be viewed at full size via this link.

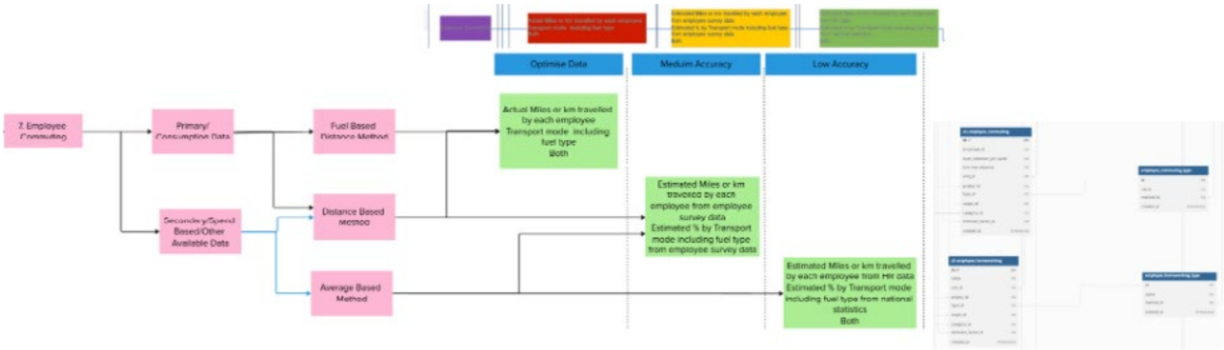


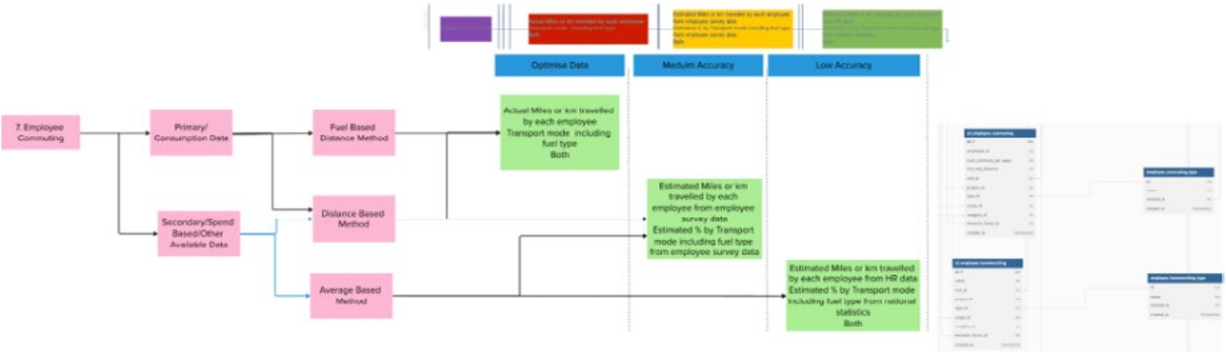


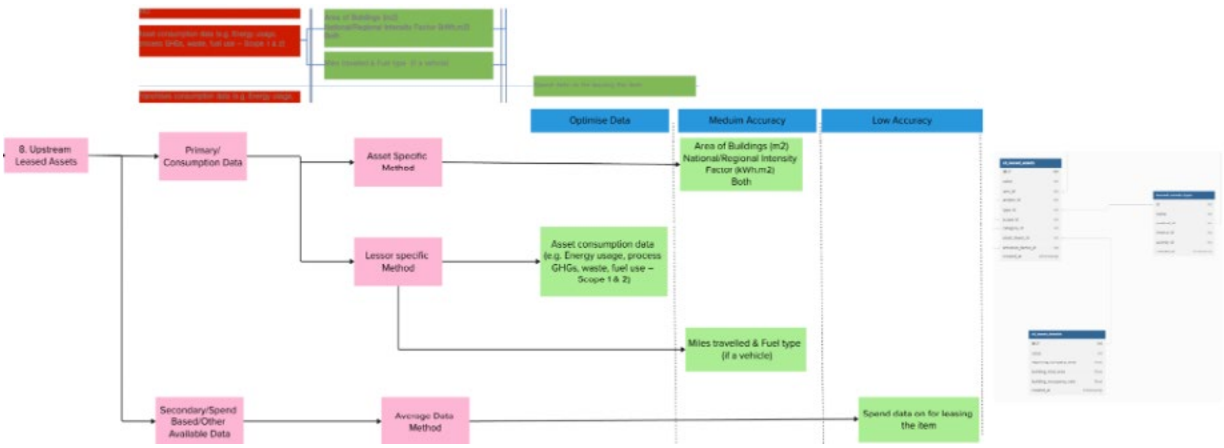

Upstream Transportation
Optimise Data
Medium Accuracy
Low Accuracy
Primary/ Consumption Data
Fuel Based Distance Method
Quantity of Fuel used from the supplier
Secondary/Spend Based/Other Available Data
Distance Based Method
Miles or Km travelled from supplier Car type including fuel type Both
Spend Based Method
Transport specific spend based invoices (availability probably dependent on industry)
Transport specific derived from what brought where and included in purchasing aboveRISK of double accounting)
Distribution
Primary/ Consumption Data
Site Specific Method
Secondary/Spend Based/Other Available Data
Average Product Method (Mass based)


Waste
Optimise Data
Meduim Accuracy
Low Accuracy
Supplier Specific Method
Primary/ Consumption Data
Waste type Specific Method
Based/Other Available Data
Average data Method
Spend Data on waste management services


6. Buisness Travel
Optimise Data
Meduim Accuracy
Low Accuracy
Primary/ Consumption Data
Fuel Based Distance Method
Actual Miles or km travelled by each employee from booking system Transport mode including fuel type
Distance Based Method
Secondary/Spend Based/Other Available Data
Actual Miles or km travelled by each employee from expenses Transport mode including fuel type Both
Spend Based Method
Number of nights stayed Region / Country Both
Hotel Spend Region / Country Both


Optimise Data
Meduim Accuracy
Low Accuracy
7. Employee Commuting
Primary/ Consumption Data
Secondary/Spend Based/Other Available Data
Fuel Based Distance Method
Distance Based Method
Average Based Method
Actual Miles or km travelled by each employee Transport mode including fuel type Both
Estimated Miles or km travelled by each employee from employee survey data Estimated % by Transport mode including fuel type from employee survey data
Estimated Miles or km travelled by each employee from HR data Estimated % by Transport mode including fuel type from national statistics Both


Optimise Data
Meduim Accuracy
Low Accuracy
7. Employee Commuting
Primary/ Consumption Data
Secondary/Spend Based/Other Available Data
Fuel Based Distance Method
Distance Based Method
Average Based Method
Actual Miles or km travelled by each employee Transport mode including fuel type Both
Estimated Miles or km travelled by each employee from employee survey data Estimated % by Transport mode including fuel type from employee survey data
Estimated Miles or km travelled by each employee from HR data Estimated % by Transport mode including fuel type from national statistics Both


Optimise Data
Meduim Accuracy
Low Accuracy
8. Upstream Leased Assets
Primary/ Consumption Data
Secondary/Spend Based/Other Available Data
Asset Specific Method
Lessor specific Method
Average Data Method
Area of Buildings (m2) National/Regional Intensity Factor (kWh.m2) Both
Asset consumption data (e.g. Energy usage, process GHGs, waste, fuel use – Scope 1 & 2)
Miles travelled & Fuel type (if a vehicle)
Spend data on for leasing the item

UKRI Innovate UK

CATAPULT Network