GFLI branded data

Methodology & Procedures for a Data-in Project

Version 1 (October, 2023)

Table of Contents

1 GFLI Branded data methodology

The following methodology is the result of the evaluation during the GFLI Branded Data Pilot that ran from February 2022 to March 2023. Any company interested in branded data through the Global Feed LCA Institute (GFLI) shall comply with the following methodology.

1.1 Introduction

The GFLI branded data methodology is created to be used by interested companies wanting to conduct an environmental assessment of their brand-specific products/ingredients compliant to the guidelines of the Global Feed LCA Institute. The aim of the GFLI branded data is providing guidance for a company-specific LCA dataset that considers the full chain of custody[1] through reliable, quantitative and, ideally, primary data. The GFLI acknowledges the difficulties of collecting this full scope of data and therefore integrated minimal requirements into the methodology that allow consistent GFLI branded datasets with the incentive to improve the quality of the data as the market matures.

Data users will be able to identify the quality and scope of each branded dataset through the publication of information on which sources were primary data and the Data Quality Rating (DQR).

This is a living document that is intended to be updated regularly. Please note that with a significant methodological update, your branded datasets may require new data to be collected in order to remain aligned with the GFLI branded data methodology. Consult the GFLI Procedures for a Branded data-in Project for details on such processes.

1.2 Chain of Custody and tiers

The chain of custody is a standard of full supply chain responsibility, detailing all relevant partners in a supply chain during the production of an ingredient (isealalliance.org, 2016). While the branded data methodology (at this point) will not require documentation supporting such a system, it indicates the shared responsibility for the production of an ingredient and structures this in terminology the GFLI upholds. These can be categorized in multiple models with various properties, including but not limited to:

- Identity preservation
- Segregation
- Mass balance, divided into batch level, site level, and group level
- Certificate trading

Each model qualifies for a level of (physical) traceability of their ingredient.

Accommodating the chain of custody, the GFLI uses the tier system. "*A tier represents a level of methodological complexity. Usually, three tiers are provided. Tier 1 is the basic method, Tier 2 intermediate and Tier 3 the most demanding in terms of complexity and data requirements. Tiers 2 and 3 are sometimes referred to as higher tier methods and are generally considered to be more accurate on condition that adequate data are available to develop, evaluate and apply a higher tier method*" (IPCC, 2019)

In the branded data methodology, suppliers beyond the applying company's own facilities, are identified as tier 1 supplier (those supplying directly to the applying company) and tier 2 suppliers (supplying indirectly to the applying company with intermediary (processing or production) steps. In the criteria for data collection (chapter 1.8), required data (quality) from tier 1 and tier 2 suppliers are shown.

---

[1] Chain of Custody is described as '*The custodial sequence that occurs as ownership or control of the material supply is transferred from one custodian to another in the supply chain'. (Adapted from: WB, WWF Alliance for Forest Conservation and Sustainable Use, 2002)*.

1.3   Distinction in data quality

The GFLI methodology distinguishes three types of data, with the distinction of primary and secondary data coming from the terminology in the FAO-LEAP guidelines (FAO-LEAP, 2016):

- **Primary data:** refers to data sourced directly from the farming/fisheries or manufacturing process (on-farm or on-facility data), which are directly and completely under the control of the applying company.
- **Secondary data:** refers to default data coming from global sources on which the GFLI database relies for its regional datasets, e.g., FAOstat, EUROstat; or may be data collected through default models e.g. LUC, fertilizer emissions).
- **Adapted secondary data:** refers to not-primary sourced data but would be considered of higher representative value than the default secondary data (DQR ≤2.0). At minimum, the applying company must know where the material is produced and with which technology in order to select the most representative dataset. This could be presented in ways of regionalized (& publicized) databases, improving default data by updated and/or more representative data input, a market mix of processed ingredients better representing the trade data of the company, or secondary datasets supplemented by primary data.

Due to the nature of branded data, primary data should be the leading source in a full chain of custody for the ingredient. Secondary data can be relied upon when there is no data available from the suppliers or upstream processes of your ingredient, but this has a negative impact on the Data Quality Rating (DQR). Adapted secondary data presents a relative solution towards more representative data for the region or process but must be shown as a higher quality data source as compared to secondary data (which can be calculated through the DQR).

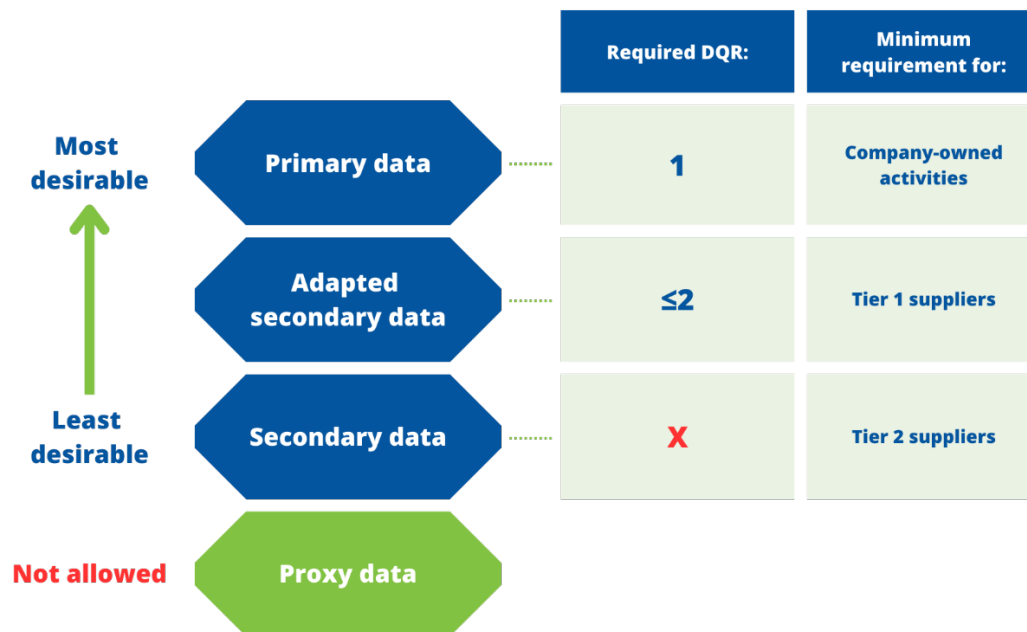|	Required DQR:	Minimum requirement for:
Primary data	**1**	**Company-owned activities**
Adapted secondary data	**≤2**	**Tier 1 suppliers**
Secondary data	**X**	**Tier 2 suppliers**
Proxy data		

Figure 1. Visualization of the desirability of the types of activity data

Concrete examples of what is considered adapted secondary data:

- The available information from the supplier should allow for a selection of country-specific secondary data relevant to the actual origin of imported ingredients, instead of a regional market mix.
- The energy source of the supplier is known. The secondary datasets may be adapted to reflect this.
- A limited amount of primary data can be collected from the tier 1 supplier, but not enough to adhere to the threshold of primary data required (see chapter 1.3.3); in combination with regionalized or secondary data the primary data may be used to create a more representative adapted secondary dataset.

Adapted secondary data will always be subject achieve a limited DQR of equal or less than 2 and may thusly still be rejected if the data fits the concrete examples mentioned above.

1.4 Data sampling at primary data collection

Data sampling for the collection of primary data is described on page 10/70 of GFLI methodology document. Data sampling may be applied for the collection of primary data in case multiple production sites are involved in the production of the same product (e.g., in case the same feed ingredient comes from multiple production sites or in case the same process is outsourced to more than one subcontractor/supplier). Stratified data sampling is often needed to deal with variation in (performance) of technologies. **The applying company should argue and report that their production sites are similar and consider management and output results that could influence the outcome of the sampled sites in comparison with the total production yield/output.** Only when this prerequisite is fulfilled, the square root methodology can be applied. If not, all individual production sites should be included in the assessment.

The following subchapters (1.4.1 and 1.4.2) are the minimally required example methods of sampling. Alternative methods may be considered in discussion with the Technical Management Committee (TMC).

1.4.1 Data sampling method: stratified method by square root

The following approach is a guiding **baseline** for defining the sample size of branded data, but deviation through other approaches is applicable in conjunction with substantiation of representativeness of the product or production line. Any deviation from the baseline sampling method must be approved by the GFLI Technical Management Committee.
The procedure to select a representative sample as a stratified sample is as follows:
1) define the population of operation
2) define if there is variability in (performance) of technologies[2] homogenous sub-populations (stratification)
3) define the sub-samples at sub-population level
4) define the sample for the population starting from the definition of sub-samples at sub-population level, using confidence interval to assure significancy.
5) identify what level of primary data can be gathered within the sample size. Some secondary data may be necessary to fill data gaps, but with a high ratio of secondary data used the sample size should increase.

The baseline approach for defining the sample size for branded data is to use **the square root of the number of operations in the sub-population**, i.e., farm operation (field and harvest management), farm size (hectares), yield, processing operation (drying, processing management), and geography are considered. The sample size for each type of farm operation and land use and each type of geography should be determined. For animal-based products the technology behind the brand should be determined e.g., free range hens or beef finished in a feedlot. For processed products (e.g., chemicals, minerals, and other manufactured products) sub-population sampling requirements should be designed according to the age of a production plant. This indicates that the GFLI branded data method on sampling varies per feed ingredient type. Rounding up to a full number might be applicable, as approached in the PEFCR feed (<0.5 would be rounding down, ≥0.5 would be rounding up).

A simple example of determining sampling size: There are 100 farms providing soybeans as an input. 25 farms are in the US, 25 farms are in Brazil, 25 farms are in France and 25 farms are in Argentina. In this simplified case, the geography is consistent by country of origin. Of the farms in Brazil, 9 of them are organic and 16 are conventional. Of the farms in Argentina, 5 farms have 80 hectares of land. 5 farms have 5 hectares. The other 15 have 40 hectares of land.

Sampling size calculation:

---

[2] variability in (performance) of technologies: for cultivated products three types of farm operation should be identified: conventional, organic and regenerative.

- *Square root of each country farms gives required 5 farms for each country or 20 farms total.*
- *Additionally, for Brazil, 3 of the organic farms (sq root of 9) and 4 conventional (sq root 16) are required for a total of 7 farms in Brazil instead of 5.*
- *The same method applies when a large variability in farm sizes is present. For Argentina, 2 farms of 80 hectares, 2 farms of 5 hectares, and 4 farms of 40 hectares should be included. Total farms to provide primary data of 22 based on above divisions by country and farm operation type.*

Cultivated products need crop (management) data averaging over 3 years, aligned with the regional data-in projects, to consider the fluctuations of yield, irrigation, and other processes. This may not be necessary for processing due to a stable situation but will need the newest data available to provide an accurate current situation.

1.4.2 Sample sizing by weighted average

When using a sample size with farms/facilities from different regions or with different technologies and management, making a weighted average may apply. Assuming post-production does not differentiate their product by facility for production, the weighted average of total population should be considered as opposed to the weighted average of the sample. This decreases the biases with random selections, creates a more accurate share of all emissions from the different facilities under the company, and aligns with the overall methodology.

$$\sum n(countries) \frac{production\ quantity\ country}{Total\ production} \cdot impact\ country$$

Example:
The case: Company X produces product A at 9 sites in 4 different European countries. The sites per country are taken as sub-population and per sub-population the sample is based on the square root.

Table 1. example of making a weighted average in figures

Country	# sites	Production (kton)	Share total production	Impact in CO_2-eq per kton	Impact * share of production
DE	2 sites	250 kton	50%	100 CO_2-eq	50 CO_2-eq
NL	4 sites	100 kton	20%	200 CO_2-eq	40 CO_2-eq
BE	1 site	90 kton	18%	250 CO_2-eq	45 CO_2-eq
UK	2 sites	60 kton	12%	150 CO_2-eq	18 CO_2-eq
Total	**9 sites**	**500 kton**	**100%**	**700** CO_2-eq	**153 CO_2-eq**

Calculating the relative impact per country against the share of total production, the weighted average of table 1 impact is 153 kg CO_2-eq. With the above formula, the numbers per country are indicated in the last column. In this example a country equals a sub-population and the impact per country is based on the impact calculated per sub-population.

1.4.3 Sampling primary data from (tier 1) suppliers: additional guidance

In order to have the most representative LCA results of the product, ideally primary data of each life cycle stage should be collected. GFLI recognizes that obtaining a complete primary data inventory may not be feasible for certain ingredients or suppliers, therefore the GFLI methodology for branded data allows the use of (adapted) secondary data to the defined extent as shown in figure 1.

- Sampling primary data from the (tier 1) suppliers is possible if the representative sample size from suppliers covers at least 70% of the material's volume, in which case the primary data would not need to be supplemented with adapted secondary data. If this sampling percentage cannot be obtained, the remaining material's volume (i.e. at least over 30%) must be covered with adapted secondary data. This percentage must also be indicated in the data collection report.

- In case of limited secondary data availability, such as if only a global secondary dataset available for the material which insufficiently presents the material, the company must improve

the data through supply-chain-specific information in order to achieve a data quality rating of ≤2. In case this poses an impossible task, the applying company may share their case to discuss within the Technical Management Committee on how to proceed.

If the primary data collected misses datapoints in which the inventory would remain incomplete, these datapoints may be completed through (adapted) secondary data as long as they do not hold a significant impact on the product's emissions and would not be subject to scrutiny due to sensitivities (e.g. relying on secondary data for deforestation-free ingredients). The GFLI methodology for branded data upholds the Data Quality Rating (DQR) to achieve high-quality branded data, thus tier 1 supplier data should always achieve this quality and representativeness requirement.

1.5 Data quality measurement

Data quality rating (DQR) shall be conducted based on the quality matrix, being developed in the EC feed database project (see Annex 2 of GFLI methodology document). The DQR information needs to be gathered during the data collection process. Where primary data are required according to these methods, data must be rated DQR 1. Where adapted secondary data are allowed according to these methods, the DQR must be at least a 2. Within the GFLI branded data methodology, data provided from the supplier(s) of the different materials/carriers/processing aids, from which the data-in provider is purchasing (chain of custody), should have a DQR of 2 or lower. The score of criterion Precision (P) cannot be higher than 3 while the score for Time Representative (TiR), Technical Representative (TeR), and Geographical Representative (GR) cannot be higher than 2 (the overall DQR score shall be ≤1.6) as aligned with the PEFCR Feed (February 2020) update, chapter 9.4.1 Company specific datasets.

The following DQR table is aligned with the PEFCR Feed table, but further clarified to with the scope of the GFLI branded data. Please note that the DQR is currently under development by the EC-JRC (PEFCR) and will likely result in changes before the end of 2023. The DQR calculations will be considered during the GFLI internal review, but a recalculation may be requested when the update has taken place. The GFLI cannot be held accountable for any costs necessary for recalculation.

Table 2 DQR interpretation specified towards GFLI branded data

DQR	Precision	Time Representative	Technology Representative	Geographical Representative[3]
1	Measured/ calculated and verified through third party verification or through the internal or external review process	Data (at collection) is maximum 2 years older than the end of the running year	Technology of source data is the same as described in the title and meta data of the GFLI dataset.	Geography of source data is the same as geography stated in the "location" indicated in the meta data of the GFLI dataset
2	Measured/calculated/literature and plausibility checked by reviewer	Data (at collection date) is maximum 4 years older than the end of the running year	Technology of source data is very similar as to what is described in the title and meta data (use of generic technology data instead of modelling all the single plants)	Geography of source data is representative for the geography stated in the ''location'' indicated in the meta data
3	Measured/ calculated/ literature and plausibility not checked by reviewer, or Qualified estimate based on calculations plausibility checked by reviewer	Data (at collection date) can be maximum 6 years older than the end of the running year	Technology of source data is similar to what is described in the title and meta data but merits improvements. Some of the relevant processes are not modelled with specific data but using proxies.	Geography of source data is sufficiently representative for the geography stated in the "location" indicated in the meta data. E.g., the represented country differs but has a very similar electricity grid mix profile.

---

[3] the definition of geography to determine geographical representative: climate zones/ecozones, with similar climate and soil characteristics.

There might be an issue regarding the geographical representation if data is collected from a very specific location. At that stage it is very important that the geographical representation can be covered within the sampling method used.

To evaluate the DQR, a division needs to be made on the type of data and how they are interrelated. Data quality evaluation shall consider the contribution of the data points to the overall environmental impact. The DQR evaluation includes activity data and the background data they relate with (e.g., production of goods such as transport and electricity and combustion of fuels or other chemical conversion during processing).

1.6   Background data

The GFLI database uses background data from Agri-footprint and Ecoinvent. All branded datasets shall be modelled with this same background data, unless primary data can be presented with a higher representativeness (such as for fertilizer use and transport). Any deviations must go through the GFLI Technical Management Committee for approval.

1.7   Peat emissions

Peat emissions are currently not a mandatory part of data collection within the GFLI branded data methodology, but the default method for peat emissions will apply (page 22 of the methodology). National or regional topographies for peat lands to represent the specific region the cultivated ingredient is sourced from may be presented to the GFLI Technical Management Committee (TMC) for a more accurate depiction of peat emissions of that ingredient. Alternative options may be considered as this topic is still under development for the branded data methodology.

1.8   Land Use change

The GFLI methodology follows the PAS 2050-1 (BSI, 2012) method for calculating land use change (LUC) (page 22 of the methodology). Branded data participants may use the default method with the use of the "Direct Land Use Change Assessment Tool version 2021" by Blonk Sustainability; but are encouraged to gather primary data. If a higher tier model is desired to use, the model shall be discussed for its use in the GFLI Technical Management Committee (TMC). Please contact the GFLI Secretariat if this is applicable.

1.9   Criteria for each ingredient

To create a full inventory that aligns with the branded data methodology, criteria is set up for the amount of primary data necessary. For tier 1 suppliers, adapted secondary data or better (primary) data must be collected. For tier 2 suppliers, secondary data or better (adapted secondary or primary) data must be collected. It's desirable to have higher quality/more representative data than the secondary data to obtain a better Data Quality Rating (DQR).

Each ingredient should include nutritional data regarding the caloric value, dry matter content, crude energy content, nitrogen, and phosphorus content. If primary data is not available, defaults may be used from relevant feed ingredients nutritional tables or the Feedipedia from FAO (this data must be sourced by the applying company).

1.9.1   Example of tier 1 and tier 2 suppliers

The following figure 2 shows the interpretation of tier 1 and 2 suppliers for feed additive manufacturing and processing. In the figure, the following terms are defined as follows:
- **Materials:** substances, including raw materials and other inputs (excluding processing aids and carriers) used for the manufacturing process (definition extracted from the ICCF). This definition would cover a chemical used in the chemical synthesis, thyme plant to produce thyme extract, thymol, or oregano essential oil in a mixture of flavouring, extracted material from mining (for trace elements), ingredient entering in the fermentation broth composition.

- **Processing aids:** any substance, not including apparatus or utensils, and not consumed as a feed ingredient by itself, intentionally used in the processing of materials, feed or feed ingredient, to fulfil a certain technological purpose during treatment or processing and which may result in the non-intentional but unavoidable presence of residues or derivatives in the feed ingredient or its ingredient market formulation, provided that these residues and derivatives do not have an adverse effect on animal health, human health or the environment (same source as above). This definition cover for example solvents for extraction /extraction step in the manufacturing process.
- **Carriers:** A feed ingredient or water used to physically facilitate handling of the feed ingredient under assessment and its incorporation into ingredient market formulations, premixtures, feeds or water. The use of a carrier does not alter the feed ingredient's intended effect and purpose (same source as above). This definition would cover wheat middling used in a mixture of additive coated agents, potentially antioxidants or emulsifiers.
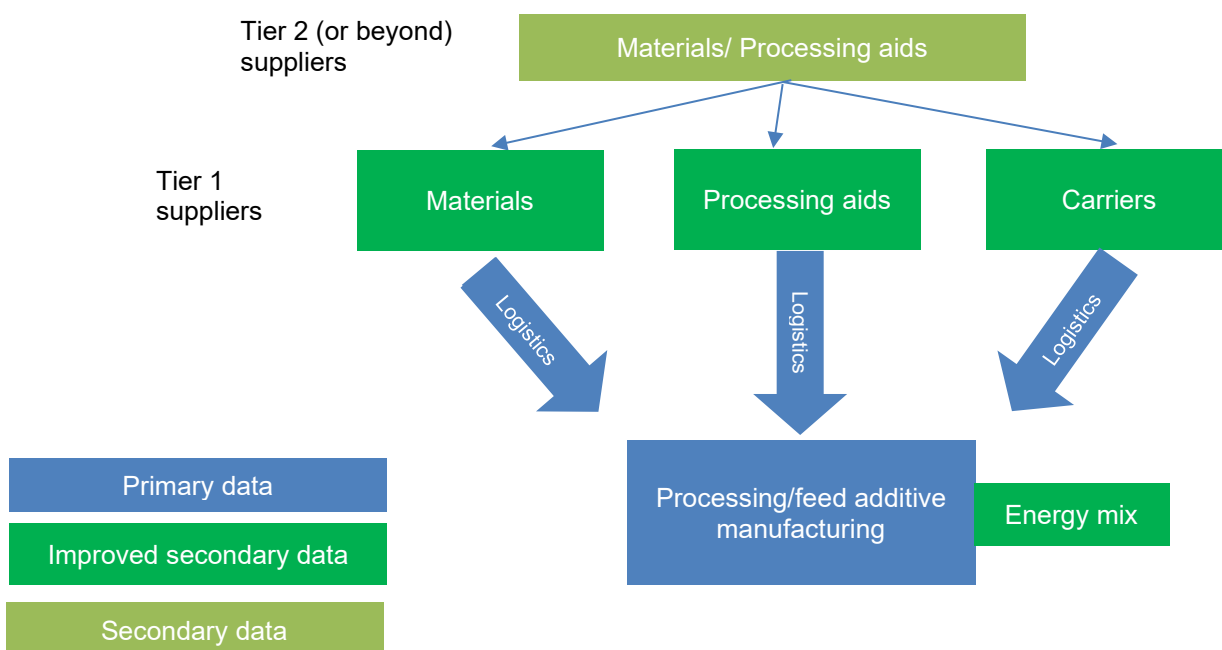

Figure 2. Interpretation of tier 1 and 2 suppliers for feed additives

1.9.2 Cultivated feed ingredients

Cultivated ingredients are all crops produced on land or in a greenhouse, and are the main feed sources, such as soy and maize.

Table 3. Criteria for cultivated ingredients

	Data provider (Cultivation)	**Tier 1 supplier (inputs)**	**Tier 2 supplier**
Required	Primary data (3-year average) available on: - Fuels use - Electricity use - N, P, K Fertilizer use - Organic fertilizer (manure and others) use - Lime use - Use of organic fertilizers or soil improvers - Use of Pesticides on the field and at storage - Use of irrigation water Secondary or better data on:	Secondary or better data on: - Packaging of fertilizers and pesticides - Chemical input	

	- Depreciation of capital goods for machinery and storage - Seed use - Land use (change)		
Example of actor	Soybean grower	Tier 1 supplier is seeds company, organic/chemical fertilizer provider, pesticides provider	

1.9.3 Processed plant-based products

Processed plant-based products are all from cultivation originated ingredients processed in some form or another to get the final product. This may be drying, milling, crushing, flaking, heating, or other processes.

Table 4. Criteria for processed plant-based products

	Data provider (processer)	**Tier 1 supplier**	**Tier 2 (cultivation farm)**
Required	Primary data on: - Fuels use - Electricity use - Water use - Wastewater treatment only for wet processes - Organic waste and losses - Auxiliary materials (processing aids) Adapted secondary data on: - Mass balance/ market mix and prices - Inbound transport (distance per transport means)	If tier 1 supplier produces a cultivated ingredient, look at 1.9.2 cultivated ingredient for which data would need to be adapted secondary or better data. If tier 1 supplier is another processor, see the left column (data provider (processor) for which data would need to be adapted secondary or better data.	Secondary or better data on cultivation (3-year average): - Fuels use - Electricity use - N, P, K Fertilizer use - Organic fertilizer (manure and others) use - Lime use - Use of organic fertilizers or soil improvers - Use of Pesticides on the field and at storage - Use of irrigation water
Example of actor	Soybean meal producer	Tier 1 supplier can be the soybean producer, the crusher facility, another processor, or a trader.	Tier 2 supplier can be the soybean producer, processor if tier 1 supplier is a trader
Example of actor	Mixer/processor of former foodstuffs	The tier 1 supplier are food processing facilities, bakeries, or others, where the former foodstuffs are sourced from, or the tier 1 supplier can be a trader	If tier 1 supplier is a trader, the tier 2 would be the food processing facilities, bakeries, or others

1.9.4 Processed animal-based products

This are all animal co-products becoming available during the slaughter and rendering, such as animal fat, processed animal protein and fish meal. The rendering and processing should contain primary data.

Table 5. Criteria for processed animal-based products

	Data provider (processer/ renderer)	**Tier 1 supplier (slaughterhouses)**	**Tier 2 supplier (livestock farmer)**
Required	Primary data available on: - Animal species (bovine, porcine, poultry, fish) and characteristics - Input: output mass balance of animal (co) products	Adapted secondary or better data on: - Heat/electricity use - Fuels use	If tier 1 supplier is a processor Secondary or better data on: - Data input of live animal

	- Allocation data (3-year average for prices in case relevant) - Fuels use - Heat/Electricity use - Auxiliary materials (for fish e.g., anti fouling, baits) Secondary or better data on: - Mass balance/ market mix and prices		
Example of actor	Feather meal processor	Tier 1 supplier is the slaughterhouse	Tier 2 supplier is the livestock farmer

Secondary data may be used for data input of the live animal. Adapted secondary data on slaughterhouses, such as based-on-industry-data GFLI data (upon approval of the intellectual property owner), may be used. The data for fisheries should be collected for a specific zone (FAO catch zone and subdivisions) and being representative for a 3 year-period (averaging out yearly variations in catches) and fishing technology.

1.9.5 Feed additives

Feed additives are any manufactured products through natural or chemical processes for the enhancement of animal health, productivity, or environment. Please note that the scope of the GFLI database and methodology is until processing gate, thus the impact a feed additive may have on the environmental output, animal performance, or otherwise, is not taken into account. This may be considered in the future, but at this point of time is not feasible.

Table 6. Criteria for feed additives

	Data provider (feed additives company)	Tier 1 supplier (Processor)	Tier 2 supplier (start material/processor)
Required	Primary data available on: - Logistics - Feed additive manufacturing which involves mass balance, fuels use, electricity Auxiliary materials etc. please use same structure as for processing plant and animal based products - Natural or chemical processes: biomass extraction, mining, fermentation	Adapted secondary or better data on: - Materials - Carriers - Processing aids - Auxiliary materials - Heat/electricity use - Fuels use	Secondary or better data on: - Originating ingredient cultivation/production values - Materials - Carriers
Example of actor	DL-methionine producer	Tier 1 suppliers are those where the ingredients are sourced from for the manufacturing of the amino acid, either through direct purchase or after processing	Tier 2 suppliers are those where the ingredients are sourced from the tier 1 supplier for either direct purchase or after processing.

1.9.6 Novel feed

Novel feed is not formally included in the GFLI methodology and the following suggestions have not been tested. Novel feed are not yet commercially available alternative feedstock, such as insects and algae. Some caution may be applied for the inclusion of alternative feed sources as GFLI branded data.

Insects and novel feed ingredients do not have a formal methodological approach as of the 2022 GFLI methodology. With insects, the animal system should be included, such as system boundaries, which emissions come out of the system, what outputs do they have (frass). Modelling should include the

insect breed and nutritional composition as a result of the feed composition fed to the insect. Ideally, this requires an average of 3 years of composition and processing but may also be an average of 1 year if proven representative. For decentralized/not-integrated insect systems, outsourced productions e.g., breeding, rearing, processing, may be (adapted) secondary data.

Table 7. Criteria for novel feed

	Data provider (Insect processor/integrated farm)	Tier 1 supplier (substrate provider/feed)	Tier 2 supplier (breeder)
Required	Primary data available on: - Logistics - Fuels use - Electricity use - Auxiliary materials In case of animal farming (insects), primary data available on: - Insect breed and nutritional composition of the insect - Substrates - Feed composition for the insect (consistent feed or otherwise averaged over 1 year)	Adapted secondary or better data on: - Feed/substrate production - Heat/electricity use - Fuels use	Secondary or better data on: - Data input on animal
Example of actor	Insect meal processor	Tier 1 supplier may be the substrate/feed provider. Not-integrated systems may also have tier 1 suppliers be breeders or rearing facilities.	

2 Procedures for a Branded data-in project

The following chapter details the procedures to start a branded data-in project.

2.1 Eligibility

Branded data has a stricter methodology and review process than non-branded data. This is to ensure as accurate and representative data as possible. Through this inclusion, a product can be verified for its carbon footprint and other environmental impacts related to its production to comply with the GFLI branded data methodology. Through a standardized method and data verification, the key element to achieve is high quality data.

Some expectations to consider,
- Supplier data is required to meet the Chain of Custody standards of the GFLI methodology. The approach of 'I buy from everywhere' requires at least adapted secondary data on a regional level. This is specifically relevant for plant-based processed products and animal rendering products.
- Having branded data compliant with the GFLI methodology will need frequent (once every two years) updates to continuously ensure time and technology representativeness as indicated by the Data Quality Rating (DQR). Likewise, any updates to the branded data methodology and/or new inclusions of data collection in the overall GFLI methodology, will require an update of the datasets to remain compliant to the latest GFLI methodology.

The branded data table on the GFLI website indicates the primary data sourced beyond the minimal requirements (primary data from applying company). Alongside the DQR, this should provide the relevant information to data users in order to make informed decisions.

Table 1. How a branded data result will be published publicly

| Ingredient | Company name | Reference year | Overall DQR | Inventory data sources | | | Chain of custody type | Results (LCIA) availability | LCI availability in .csv |
				Company-owned activities	Tier 1 suppliers	Tier 2 suppliers			
Beet pulp meal, dried, at storage/DE	XYZ	2023	1,38	Primary	100% adapted secondary data	100% secondary data	Mass balance	XYZ publishes their data online	Available upon request
Rapeseed meal, dried, at storage/DE	XYZ	2023	1,45	primary	20% adapted secondary data, 80% primary data	100% adapted secondary data	Mass balance	XYZ publishes their data online	Available upon request
Feed additive ABC, at storage/DE	XYZ	2023	1,6	primary	80% adapted secondary data, 20% secondary data	100% secondary data	Segregated	XYZ publishes their data online	Available upon request
Feed additive ABC, at storage/DE	XYZ	2023	1,6	primary	100% primary data raw material X (representative sample covers 70% of volume), 100% adapted secondary data raw material Y	90% primary data, 10% secondary data	Identity preservation	XYZ publishes their data online	Available upon request

2.2 Brand specificity

Branded data is defined as, *"Lifecycle inventory/impact assessment (LCI/A) data for a feed ingredient marketed under a certain brand and owned by a company or other entity and should present representative data of the company and its chain of custody"*.

The specificity of the branded product should be considered for the sample size necessary as well as the production and/or processing facilities to collect data from. Branded data should not be used to encourage competitiveness within a company's different production and/or processing locations but should present an average value of the brand. For example, a company produces product A in 3 processing facilities, one based in the Netherlands, one in Belgium, and one in Germany. The company itself may have the data collected from various plants to consider ways to improve individual plants, as well as present their customers plant-specific footprint for premium marketing strategies.

Regionally specific branded data may be included depending on the scope (property) of the product, for example cultivated ingredients with a specified distinguishable region or limited growth elsewhere.

2.3 Data use conditions

Upon participation of branded data, the data provider signs the license agreement and states in what forms their datasets may become available.

The information which can be derived from the Feed/Food LCI/A Dataset(s) of the Client will, at the discretion of GFLI, be categorized and made accessible as follows:
1) aggregated data (environmental impact data (LCIA) and inventory data (LCI)) will be freely accessible to the public; and
2) aggregated data may be licensed for use in tooling (i.e., commercially sold LCA calculation tools).

GFLI requires the prior approval of the Client to share the outcome and calculation of the disaggregated Feed/Food LCI/A Dataset(s) of the Client as described in items 3) and 4) below:
1) the disaggregated Feed/Food LCI/A Dataset(s) (unit process inventory data) of the Client presented in accordance with the GFLI methods and procedures and the calculations made in accordance with the GFLI methods and procedures; these data will only be made available to parties which procure a license for the use of the disaggregated Global Feed LCA Database. In this option both the outcome and calculation will be shared.
2) Disaggregated data may be licensed for use in tooling (i.e., commercially sold LCA calculation tools).

2.4 Data confidentiality

The company, data provider, maintains ownership of the datasets and is responsible for the timely update and maintenance of the representativeness of the dataset. Upon signature of the license agreement, the company approves upon their datasets possibly being used as starting blocks for sectoral datasets. The GFLI will only use a multitude of such branded datasets in its aggregated form, meaning the datasets cannot be traced to a specific company and/or dataset, and will notify the owner of the dataset prior to integration of the database. The final sectoral dataset will be owned by GFLI.

2.5 System boundaries

The GFLI branded data adhere to the same system boundaries as the sectoral/regional database. This means the system boundary is cradle-to-farm-gate for cultivated ingredients, and cradle-to-processing-gate (at plant) for processed ingredients, meaning the emissions are calculated for the entire production chain. The company may provide additional transport calculations separately for a full system boundary until animal farm gate and/or end consumer.

2.6 Process and next steps

The GFLI welcomes new branded data on a continuous basis. The schematic approach of table 2 is followed, with more details following. Successfully finalizing a branded data-in project largely depends on timely data gathering, correct data modelling in line with the GFLI branded data methodology, good communications with suppliers for achieving all relevant information of the product and/or ingredient's life cycle and having the right reports ready for the reviews. The procedures for a branded data-in project are mostly similar to the GFLI sectoral and regional data-in project procedures, with small alterations as a branded data-in project is under the control of 1 company and would therefore be a timelier undertaking.

Table 2. Procedures for a branded data-in project

	Steps
Defining the data-in project	Define scope: goal and approach, which feed ingredients, which country or region, data sources, which method, which granularity.
	Define data use conditions, sign the license agreement with GFLI (also includes how the data will become available after the external review).
	Define project plan (planning, data collection and LCA consultant, connection and plan for collection of data from suppliers, budget, etc.).
	Submit the data-in proposal to the GFLI for approval and feedback from the Technical Management Committee (info@globalfeedlca.org).
Data collection and modelling	Execute activities defined in the project plan.
	Collect data from defined sources (databases, primary data from production and/or processing facilities, data from suppliers).
	Modelling the data.
	If any questions, possible discrepancies, or alternative methods are arising during the data collection or modelling process, the applying company should reach out to GFLI for discussion within the Technical Management Committee.
	Write a report on the data collection, which activities were collected and a clarification of approaches used, and LCA report of how the data is modelled and calculated
Review and integration	Submit the data and reporting to GFLI for the internal review (info@globalfeedlca.org), costs are applicable.
	After approval of the internal review, contact one of the external reviewers indicated on the GFLI website. During the external review, points for improvement may be brought up before final approval.
	Upon final viewing of the GFLI Secretariat, you receive a letter of confirmation that the data is GFLI-compliant and, if agreed upon, will be published on the GFLI website.

2.7 Naming of datasets

All ingredients applicable for branded data shall include a name in line with the wording of the overall GFLI database, which is as follows: Product name and product general or chemical name, property/concentration, company name, at farm/processing, region/country, allocation.

A mixed ingredient with a specific brand name or a name that is not available in English, should have a name specified for common understanding, with additional information on its mixture in the process description. The use of other languages in the title may be specified in brackets. In the source, the company name and its original branded name are included. See examples in table 3.

Table 3. Example of naming ingredients

Product name	Source
L-Lysine HCL 98% [company A], at plant/DE Economic Allocation	Company A, L-Lysine HCL
Soybean [Company B], dried, at farm/BR Economic Allocation	Company B, soybean
Hybrid rye [Company C] ''hybridroggen'', at storage/DE Economic allocation	Company C, hybridroggen

For questions about the naming of the ingredient, please reach out to info@globalfeedlca.org. Note that the naming should be compatible for integration in LCA software tools.

2.8 Deliverables

Deliverables after going through a successful branded data verification:
- Letter of confirmation that the specified ingredient is calculated through the GFLI branded data methodology and is in line with the latest version of the database in order for it to be used in calculations. The letter will have an expiration date of its validity included (2 years), upon expiration the data provider should update their data to remain aligned with the methodology.
- The Lifecycle impact assessment format of the datasets, which upon approval of the data provider can be published on the GFLI website.
- (Optional) the lifecycle inventory may be shared by the LCA consultant with the data provider. At the current stage, the GFLI website will not offer a download for this format, but data providers may share its availability on the website with a contact or weblink on their own website. This may be relevant for database users to have more insight on the composition of the dataset.

2.9 Data claims and disclaimers

After finalization of a GFLI branded data-in project, data owners will receive notification from GFLI when their project has been approved. This notification will include the relevant data file and any information that needs to be included in any communication about the calculation (e.g., the version of the GFLI branded data methodology used). Only after receipt of this notification may it be stated that the product's LCA, emissions, or environmental footprint is calculated based on GFLI branded data methodology version [Version Number] as verified through the internal and external review processes outlined in the methodology.

The following disclaimer (**Article X** of the Branded Data License Agreement) refers to any and all branded data calculated per the GFLI methodology: [Licensee] is responsible for the footprint calculated using version [Version Number] of GFLI's database and methodology. GFLI makes no warranties about the accuracy of its data or claims based on its data. GFLI is not responsible or liable for reliance on GFLI data or methodology. GFLI does not endorse this or any product.

In order to make a formal footprint claim (as opposed to sharing the footprint information with an interested party) for marketing and communications on a webpage, label, package or literature, the GFLI Logo and Claim License Agreement must be signed. Please contact info@globalfeedlca.org to request information about this agreement.

2.10 Updating datasets

Following the data quality rating metrics, all primary data should be updated every 2 years, with (adapted) secondary datasets needing updates every 4 years. A request for update should be made 3 months before expiry for a small update (updating primary data), and 6 months before expiry for a large update (updating primary data and updating (adapted) secondary data. The letter of confirmation will be valid for 2 years, but datasets will remain existing on the GFLI webpage in order to provide full transparency when companies communicate their GFLI-compliant environmental assessment. The company must notify GFLI of any material changes of data and/or product processes during the 2-year validity of the confirmation letter.

The branded data methodology is a working document and may have small to large changes in the coming years. Additional methodological approaches may be included. Existing datasets that require an update after a methodological change receive a transition period of in order to collect the additional data requirements. The length of such transition period will be dependent on the methodological update and the efforts required to gather additional data and/or change the calculations. This will be communicated with applying company beforehand. At any point, the applying company will have a time window to review the methodological changes and decide whether they will continue with the publication of the branded data. If there is no response to a call for change, the GFLI is within its rights to remove the dataset(s) and their validity.