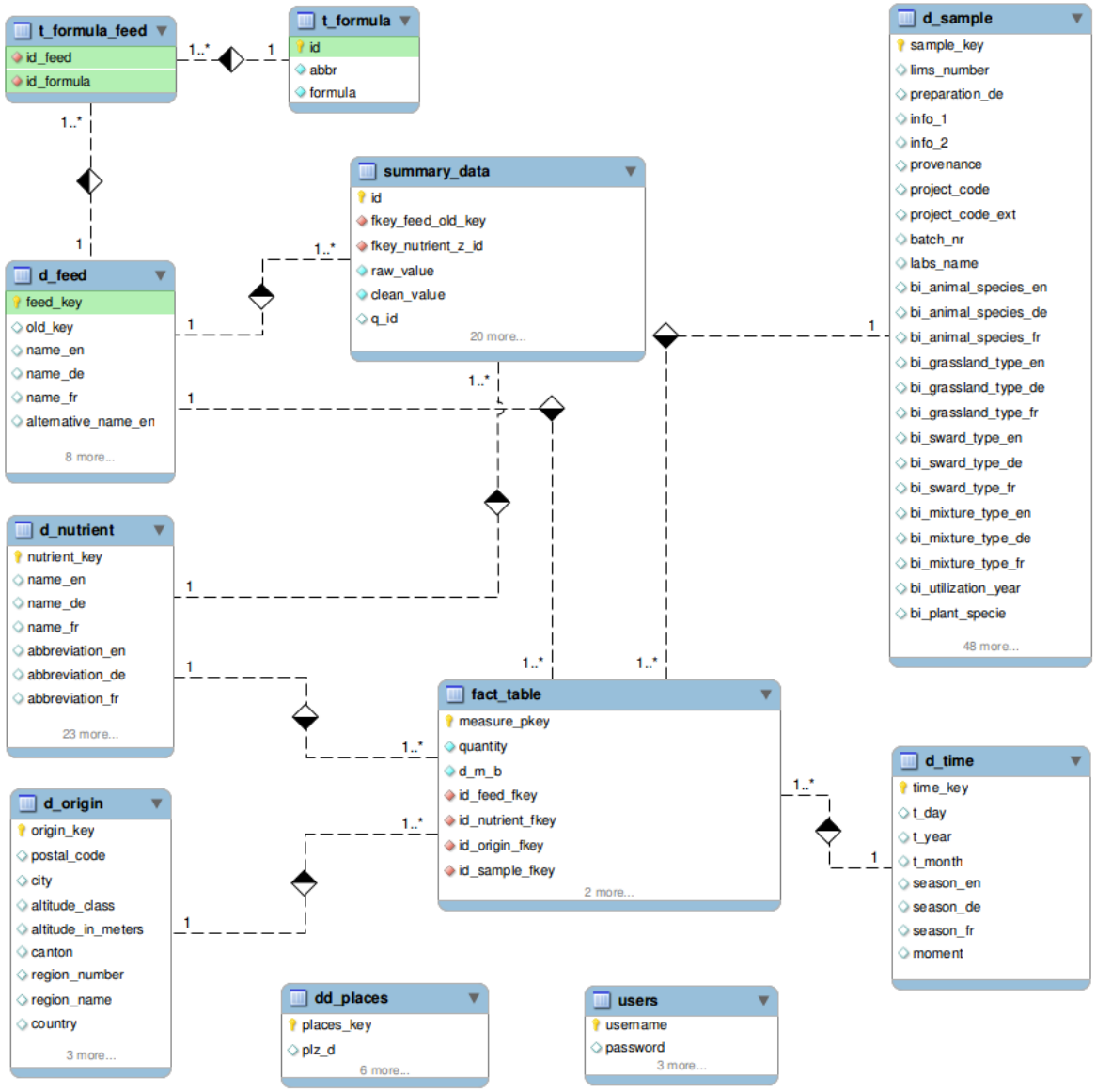

Figure 1: ER-Diagram of the Swiss Feed Database

Table 1: feed categories. The Swiss Feed Database contains data for about one thousand feeds. Feeds are identified by an artificial id and have the name in three languages (English, German and French). About 60% of all feeds form a hierarchy with the root occupied by the general feed categories and with the children of the same parent node distinguished by the biological or technical properties. We refer to such feeds as 'classified'. Unclassified feeds are feeds with incomplete information. For example, the laboratories do not use complete feed names but their own codes which are helpful to determine the general feed category but are not useful to classify the given feed into more detailed feed category.

Table 2: nutrients. The swiss database provide information on the containment of more than 900 nutrients which name, abbreviations and description are stored in *d_nutrient* table in three possible languages: English, German and French. The major part of the nutrients are grouped according to the animal species they are relevant to and, according to the chemical category (as amino acids or mineral). Then, these groups are used in the on-line web application to simplify navigation for the end user.

Tables 3 and 4: geographical data. The origin has strong impact on the quality of animal feeds and in many cases helps to determine the right diet for the farmers and to adjust the production methods for the feed industry. The well known and the most crucial parameters are production intensity in the region, altitude and animal density.

For each feed sample the Swiss Feed Database stores its origin, however, the level of detailedness varies depending on how accurately the sample is described by its provider. The common trend across Switzerland is that the feed producers provide only their name or the name of the region and, without specifying the city or village name. For the feed samples that come from abroad the country name is often the only available information. In general the import of geographical data consists of two steps. In the first step the description of the sample is investigated by the domain expert and out from the detected codes, names or abbreviations the complete location name is extracted. In the second step we use Google services in order to derive the complete geographical information out of the location name. At last, the achieved latitude, longitude, altitude and postal code together with the location name are stored in the database tables.

Table 5: feed sample. Even for the same feed type, the feed samples are different in many properties as production method, the number of consecutive years of utilization of the field, technique used to prepare the sample before the analyses and etc. In total, we distinguish 30 unique properties of a feed sample that fall into the following three classes. The general class of properties which are present for each feed sample regardless feed type are descriptions of the sample by its provider and laboratory, project code that initiated the analyses, laboratory name, batch number and preparation method. All properties of this category are stored in one language(German or French) that was used by the provider to describe the sample. The second class are biological properties that include characteristics of the grass, sward, variety type, plant specie, botanical composition, stage of maturity and etc. In addition to these categorical properties that are stored in three languages (English, German and French), biological properties encapsulate numerical data as percentage of the grass, legumes,cob and forbs in the feed sample. The third class are technical properties that describe the production system used to produce the feed, the use intensity of the land, conservation, storage and feed processing methods.

All this detailed data about the feed samples is highly demanded by the feed industry and is critical for the analyses of the feed quality. Inside the Feed Database we use this information for cleaning, i.e., for detection of samples which are assigned to the wrong feed categories or removal of nutrient measurements that are incorrect due typing mistakes.

Table 6: time information. Each sample can have up to 4 dates: harvesting, sampling, arrival and analyses dates. The availability of these date varies: harvesting date is present for the most of the feed samples, however, sampling, arrival or analyses dates are rarely tracked or provided by the feed producers. In addition to that, dates are given at different granularity levels and, therefore, we use four attributes to store time information. Attribute *t_day* is used for the case then complete date i.e., the year, month and day, is known. For other cases, then only the year or the year and month are given, we only fill attributes *t_year* and *t_month*. The derived attribute *season* stores a season of the year, i.e., winter, spring, summer or autumn.

Tables 7 and 8: nutrient measurements. The Feed Database provides two views on the containment of nutrients in the animal feeds: detailed and aggregated. Detailed view allows to investigate nutrients of individual feed samples. In this case, no search restriction are applied: it is possible to choose samples based on location, time, biological or technical properties. Nutrient measurements of individual samples are organized into *fact_table* that in addition to the numeric value contains foreign keys to tables *d_sample*, *d_nutrient*, *d_feed*, *d_time* and *d_origin*.

In aggregated view for each feed type and nutrient we provide an aggregated value that accurately quantifies the containment at the current date for the whole Switzerland. Such the aggregated measurements are highly demanded by farmers(and feed industry) which are not interested in the analyses but in forming the right diet for their animals based on the feeds that are currently present in the market. Computation of the aggregated measurements is a complex procedure. Since the feed data is collected irregularly and at different granularity levels, a simple averaging of measurements within the fixed time interval results on a high error. Therefore, the domain expert manually chooses measurements from most recent history at certain locations, computes their average and

adjusts the achieved value based on known regressions. Aggregated measurements are stored separately from the individual measurements in table *summary_data*. In addition to the aggregated value, *summary_data* contains statistical information(as the number of aggregated measurements and standard deviation) and detailed information about the measure units in three languages (English, German and French).

Tables 9 and 10: derived nutrients. The derived nutrients are computed from measurements of other nutrients using predefined mathematical expressions, i.e., formulas. The most common examples of derived nutrients are digestibility and energy value of a given feed and animal specie. Computation formulas are often recursive, i.e., to compute one derived nutrient the value of other derived nutrient is needed. We use two tables to represent formulas in the database. Table *t_formula* stores formulas together with their abbreviations. Abbreviation are unique and are used inside recursive formulas. Table *t_formula* maps formulas with the feeds, i.e., it models n to n relationship between tables *t_formula* and *d_feed*.

Computation of the derived nutrients are computationally expensive and in extreme cases it can take up to 50 sec. Therefore, the feed database do not provide on-line computation of derived nutrients for the detailed data, i.e., at level of individual sample. However, for the aggregated data we store precomputed derived nutrients in summary_data *table.*

Table 11: users. The Swiss Feed Database offers extended services for the subscribed users which include data export to the Excel tables, access to the yearly feed catalogs and enhanced search across the detailed data. Each subscribed user is uniquely identified by its *username* that can be up to 50 characters length. For security reasons all passwords are encoded with MD5 encryption within the database. To distinguish between different types of users we assign *userlevel* attribute. The default value is 2 and higher values are reserved for database administrators. We also store minimal usage statistics. Attribute *login_counter* keeps track of the number of times a user logged in into the system and, attribute *last_log* records the time of the last login.

The future development of the Swiss Feed Database aims to dynamically adjust the search interface for the needs of each individual user and to collect more detailed statistics. We plan to approach this with the user profiles that will record search parameters of each request and, at any time, will provide to the user the list of the top and most recent queries.

d_feed				
nr	**attribute**	**type**	**constraints**	**description**
1	feed_key	integer	serial,not null	the primary key of the relation
2	old_key	integer		the key that uniquely determines the feed in the non-temporal(old) database
3	name_en	string	max 255 characters	the name or abbreviation of the feed in English
4	name_de	string	max 255 characters	the name or abbreviation of the feed in German
5	name_fr	string	max 255 characters	the name or abbreviation of the feed in French
6	alternative_name_en	string	max 100 characters	alternative name or abbreviation of the feed in English
7	alternative_name_de	string	max 100 characters	alternative or abbreviation of the feed in German
8	alternative_name_fr	string	max 100 characters	alternative or abbreviation of the feed in French
9	source	string	max 100 characters	determines the primary institution that provides analyses of the samples of the given feed
10	feed_group_id	integer		the id of the group that contains the given feed
11	parent_feed_group_id	integer		the id of the group to which belongs the group of the given feed
Relation *d_feed*, continued in the next page...				

12	feed_group_en	string	max 100 characters	the English name of the group that contains the given feed
13	feed_group_de	string	max 100 characters	the German name of the group that contains the given feed
14	feed_group_fr	string	max 100 characters	the French name of the group that contains the given feed

Table 1: Relation *d_feed* consists of 14 attributes in total, out of which 4 are integers and 10 are of string type.

d_nutrient				
nr	**attribute**	**type**	**constraints**	**description**
1	nutrient_key	integer	serial,not null	the primary key of the relation
2	name_en	string	max 100 characters	the complete long name of a nutrient in English
3	name_de	string	max 100 characters	the complete long name of a nutrient in German
4	name_fr	string	max 100 characters	the complete long name of a nutrient in French
5	abbreviation_en	string	max 20 characters	the short name of a nutrient in English
6	abbreviation_de	string	max 20 characters	the short name of a nutrient in German
7	abbreviation_fr	string	max 20 characters	the short name of a nutrient in French
8	group_en	string	max 100 characters	the name of the group that contains the given nutrient in English
9	group_de	string	max 100 characters	the name of the group that contains the given nutrient in German
10	group_fr	string	max 100 characters	the name of the group that contains the given nutrient in French
11	description_en	string	max 100 characters	description of a nutrient in English
12	description_de	string	max 100 characters	description of a nutrient in German
13	description_fr	string	max 100 characters	description of a nutrient in French
14	dryable	boolean		indicates if the dry matter base of the nutrient can be calculated
15	z_id	integer		the id of the nutrient for the aggregated data
16	z_abbreviation_en	string	max 50 characters	short name of the nutrient for the aggregated data in English
17	z_abbreviation_de	string	max 50 characters	short name of the nutrient for the aggregated data in German
18	z_abbreviation_fr	string	max 50 characters	short name of the nutrient for the aggregated data in French
19	z_group_id	integer		the id of the group in context of the aggregated data that contains the given nutrient
20	z_name_en	string	max 100 characters	complete long name of the nutrient for the aggregated data in English
21	z_name_de	string	max 100 characters	complete long name of the nutrient for the aggregated data in German
22	z_name_fr	string	max 100 characters	complete long name of the nutrient for the aggregated data in French
23	in_lims	boolean		true if the nutrient belongs to the detailed data, false if the nutrient belongs to aggregated data
24	z_specie_id	integer		the id of an animal specie that is relevant for the given nutrient in context of aggregated data
Relation *d_nutrient*, continued in the next page...				

25	z_specie_name_en	string	max 100 characters	the name of an animal specie that is relevant for the given nutrient in context of aggregated data, in English
26	z_specie_name_de	string	max 100 characters	the name of an animal specie that is relevant for the given nutrient in context of aggregated data, in German
27	z_specie_name_fr	string	max 100 characters	the name of an animal specie that is relevant for the given nutrient in context of aggregated data, in French
28	z_order	integer		an integer that determines the priority/order of the nutrients to be displayed in the tree while searching across the aggregated data
29	z_group_order	integer		an integer that determines the priority/order of the nutrients groups to be displayed in the tree while searching across the aggregated data

Table 2: Relation *d_nutrient* consists of 29 attributes in total, out of which 6 are integers, 2 are boolean and 21 are of string type.

d_origin				
nr	**attribute**	**type**	**constraints**	**description**
1	origin_key	integer	serial,not null	the primary key of the relation
2	postal_code	integer		the postal code of the feed sample origin
3	city	string	max 100 characters	the city name in the official language of that city
4	altitude_class	string	max 100 characters	the altitude as classes, i.e, '> 600'.
5	altitude_in_meters	integer		the altitude of the sampling location in meters
6	canton	string	max 50 characters	the name of the canton in the official language of that canton
7	region_number	integer		th id of the region according to the division of Switzerland into regions by Agridea
8	region_name	string	max 100 characters	the name of the region according to the division of Switzerland into regions by Agridea
9	country	string	max 100 characters	the country name in English.
10	latitude	float		latitude of the sampling location
11	longitude	float		longitude of the sampling location
12	animal_denisty	float		the density of animal at the sampling location

Table 3: Relation *d_origin* consists of 12 attributes in total, out of which 4 are integers, 3 are floats and 5 are of string type.

dd_places				
nr	**attribute**	**type**	**constraints**	**description**
1	places_key	integer	serial,not null	the primary of the relation
2	plz_d	integer		the postal code
3	city_short_d	string	max 50 characters	the short name of the city
4	city_long_d	string	max 50 characters	the long name of the city
5	canton_de	string	max 50 characters	the canton name in German
6	canton_abbreviation	string	max 2 characters	abbreviation of the canton name
7	region_number	integer		the id of the region according to the devision of the Switzerland by Agridea
Relation *dd_places*, continued in the next page...				

8	region_name	string	max 50 characters	the name of the region according to the devision of the Switzerland by Agridea

Table 4: Relation *dd_places* consists of 8 attributes in total, out of which 3 are integers and 5 are of string type.

d_sample				
nr	**attribute**	**type**	**constraints**	**description**
1	sample_key	integer	serial, not null	the primary key of the relation
2	lims_number	string	not null, unique	a string that uniquely identifies the sample within Lims system at Agroscope
3	preparation_de	string	max 200 characters	describes the preparation of the sample before the analyses
4	info_1	string	max 500 characters	description of the sample by the lab
5	info_2	string	max 500 characters	description of the sample by the lab
6	provenance	string	max 100 characters	description of the source of the feed sample
7	project_code	string	max 100 characters	code of the project that initiated the analyses
8	project_code_ext	string	max 100 characters	extended project code
9	batch_nr	string	max 100 characters	the code of the batch from which the sample was taken
10	labs_name	string	max 100 characters	the name of the laboratory that analyzed the sample
biological properties of the sample				
11	bi_animal_species_en	string	max 200 characters	defines the animal that provided the sample in English
12	bi_animal_species_de	string	max 200 characters	defines the animal that provided the sample in German
13	bi_animal_species_fr	string	max 200 characters	defines the animal that provided the sample in French
14	bi_grassland_type_en	string	max 200 characters	stands for the characteristics of the grass (as artificial or natural) in English
15	bi_grassland_type_de	string	max 200 characters	stands for the characteristics of the grass (as artificial or natural) in German
16	bi_grassland_type_fr	string	max 200 characters	stands for the characteristics of the grass (as artificial or natural) in French
17	bi_sward_type_en	string	max 200 characters	stands for sward characteristics (mixtures or pure stands) in English
18	bi_sward_type_de	string	max 200 characters	stands for sward characteristics (mixtures or pure stands) in German
19	bi_sward_type_fr	string	max 200 characters	stands for sward characteristics (mixtures or pure stands) in French
20	bi_mixture_type_en	string	max 200 characters	mixture type (exp. Mixtures of the grass)according to the Swiss standards in English
21	bi_mixture_type_de	string	max 200 characters	mixture type (exp. Mixtures of the grass)according to the Swiss standards in German
Relation *d_sample*, continued in the next page...				

22	bi_mixture_type_fr	string	max 200 characters	mixture type (exp. Mixtures of the grass)according to the Swiss standards in French
23	bi_utilization_year	integer		the number of consecutive years of utilization of the field
24	bi_plant_specie	string	max 200 characters	plant specie in Latin
25	bi_plant_specie_en	string	max 200 characters	plant specie in English
26	bi_plant_specie_de	string	max 200 characters	plant specie in German
27	bi_plant_specie_fr	string	max 200 characters	plant specie in French
28	bi_variety_type_en	string	max 200 characters	variety type in English
29	bi_variety_type_de	string	max 200 characters	variety type in German
30	bi_variety_type_fr	string	max 200 characters	variety type in French
31	bi_variety	string	max 200 characters	variety international name
32	bi_bot_composition_ch_en	string	max 200 characters	botanical composition according to the Swiss system in English
33	bi_bot_composition_ch_de	string	max 200 characters	botanical composition according to the Swiss system in German
34	bi_bot_composition_ch_fr	string	max 200 characters	botanical composition according to the Swiss system in French
35	bi_percentage_of_grass	float		percentage of the grass in the feed sample
36	bi_percentage_of_lolium	float		percentage of legumes on the feed sample
37	bi_percentage_of_forbs	float		percentage of forbs in the feed sample
38	bi_age	integer		the number of days since the last utilization
39	bi_utilization_stage_en	string	max 200 characters	phenological stage at the moment of utilization in English
40	bi_utilization_stage_de	string	max 200 characters	phenological stage at the moment of utilization in German
41	bi_utilization_stage_fr	string	max 200 characters	phenological stage at the moment of utilization in French
42	bi_stage_of_maturity_en	string	max 200 characters	the stage of maturity for the plants at the harvesting time in English
43	bi_stage_of_maturity_de	string	max 200 characters	the stage of maturity for the plants at the harvesting time in German
44	bi_stage_of_maturity_fr	string	max 200 characters	the stage of maturity for the plants at the harvesting time in French
45	bi_cob_proportion	float		percentage of cob in the feed sample
46	bi_quality_level_en	string	max 200 characters	general field to describe the quality of the feed sample in English
47	bi_quality_level_de	string	max 200 characters	general field to describe the quality of the feed sample in German
48	bi_quality_level_fr	string	max 200 characters	general field to describe the quality of the feed sample in French
49	bi_extra_properties_en	string	max 200 characters	other biological properties that do not fall into predefined categories in English
50	bi_extra_properties_de	string	max 200 characters	other biological properties that do not fall into predefined categories in German

Relation *d_sample*, continued in the next page...

51	bi_extra_properties_fr	string	max 200 characters	other biological properties that do not fall into predefined categories in French
technical properties of the sample				
52	te_production_system_en	string	max 200 characters	the production system used to produced the feed in English
53	te_production_system_de	string	max 200 characters	the production system used to produced the feed in German
54	te_production_system_fr	string	max 200 characters	the production system used to produced the feed in French
55	te_use_intensity_en	string	max 200 characters	the use intensity of the land (exp. how many harvest are achieved yearly) in English
56	te_use_intensity_de	string	max 200 characters	the use intensity of the land (exp. how many harvest are achieved yearly) in German
57	te_use_intensity_fr	string	max 200 characters	the use intensity of the land (exp. how many harvest are achieved yearly) in French
58	te_mode_of_use_en	string	max 200 characters	the mode of use such as pasture, cut, etc, i.e., the field type in English
59	te_mode_of_use_de	string	max 200 characters	the mode of use such as pasture, cut, etc, i.e., the field type in German
60	te_mode_of_use_fr	string	max 200 characters	the mode of use such as pasture, cut, etc, i.e., the field type in French
61	te_mode_of_conservation_en	string	max 200 characters	how the feed is conserved such as drying, ensiling and etc. in English
62	te_mode_of_conservation_de	string	max 200 characters	how the feed is conserved such as drying, ensiling and etc. in German
63	te_mode_of_conservation_fr	string	max 200 characters	how the feed is conserved such as drying, ensiling and etc. in French
64	te_feed_processing_en	string	max 200 characters	feed processing such as melting, heat expansion and etc. in English
65	te_feed_processing_de	string	max 200 characters	feed processing such as melting, heat expansion and etc. in German
66	te_feed_processing_fr	string	max 200 characters	feed processing such as melting, heat expansion and etc. in French
67	te_storage_en	string	max 200 characters	how the feed is stored such as tank, silo and etc. in English
68	te_storage_de	string	max 200 characters	how the feed is stored such as tank, silo and etc. in German
69	te_storage_fr	string	max 200 characters	how the feed is stored such as tank, silo and etc. in French
70	te_extra_properties_en	string	max 200 characters	other technical properties of the feed sample that do not fall into predefined categories in English
71	te_extra_properties_de	string	max 200 characters	other technical properties of the feed sample that do not fall into predefined categories in German
72	te_extra_properties_fr	string	max 200 characters	other technical properties of the feed sample that do not fall into predefined categories in French

Table 5: Relation *d_sample* consists of 72 attributes in total, out of which 3 are integers, 4 are floats and 65 are of string type.

d_time				
nr	**attribute**	**type**	**constraints**	**description**
Relation *d_time*, continued in the next page...				

1	time_key	integer	serial,not null	the primary key of the relation
2	t_day	date		if it is available, the complete date
3	t_year	integer		if it is available, the year
4	t_month	integer		if it is available, the month as a number from 1 to 12
5	season_en	string	max 50 characters	the name of the season in English
6	season_de	string	max 50 characters	the name of the season in German
7	season_fr	string	max 50 characters	the name of the season in French
8	moment	integer		indicates the type of the date: value 1 stands for the harvesting time, value 2 stands for the sampling time, value 3 stands for the arrival time and value 4 stands for the analyses time

Table 6: Relation *d_time* consists of 8 attributes in total, out of which 4 are integers, 1 is date and 3 are of string type.

fact_table				
nr	**attribute**	**type**	**constraints**	**description**
1	measure_pkey	integer	serial, not null	the primary key of the relation
2	lims_number	string	max 20 character	the lims number of the sample for which the the given record stores a nutrient's measurement
3	quantity	float		the numeric evaluation of the containment of a nutrient
4	id_time_fkey	integer		the foreign key to the *d_time* relation
5	id_nutrient_fkey	integer		the foreign key to the *d_nutrient* relation
6	id_origin_fkey	integer		the foreign key to the *d_origin* relation
7	id_sample_fkey	integer		the foreign key to the *d_sample* relation
8	id_feed_fkey	integer		the foreign key to the *d_feed* relation
9	d_m_b	boolean		true the quantity is converted to the dry matter basis, false otherwise

Table 7: Relation *fact_table* consists of 9 attributes in total, out of which 6 are integers, 1 is float, 1 is boolean and 1 is of string type.

summary_data				
nr	**attribute**	**type**	**constraints**	**description**
1	id	integer		the primary key of the relation
2	fkey_feed_old_key	integer	not null	determines the feed type by mapping to the *old_key* in *d_feed* relation
3	fkey_nutrient_z_id	integer	not null	determines the nutrient by mapping to the *z_id* in *d_nutrient* relation
4	raw_value	real		raw aggregated value that determines containment of the nutrient in the given feed
5	clean_value	string	max 100 characters	clean aggregated value that determines containment of the nutrient in the given feed
quality parameters of the aggregated value				
6	q_id	integer	not null	the unique artificial id for each group of quality parameters
7	q_standart_deviation	double		standard deviation of all nutrient measurements that contributed to the aggregation
Relation *summary_data*, continued in the next page...				

8	q_number_of_values	integer		number of nutrient measurements that contributed to the aggregation
9	q_decimal_places	integer		precision of the aggregated value in terms of number of digits after comma
10	q_value_type	integer		1 stands for reliable aggregated values, 0 stands for not reliable aggregated values that usually are displayed as 'NA'
11	q_formula	string	max 255 characters	if applicable, contains a formula that was used to aggregate nutrient measurements
12	q_modification_date	string	max 30 characters	the date of the last modification
13	q_auto_remarks	string	max 255 characters	additional comments
units				
14	u_id	integer	not null	the unique artificial id for each distinct unit
15	u_name_en	string	max 255 characters	the name of unit in English
16	u_name_de	string	max 255 characters	the name of unit in German
17	u_name_fr	string	max 255 characters	the name of unit in French
18	u_kg_converter	integer		
19	u_group_id	integer	not null	the unique artificial id for each distinct group of units
20	u_group_name_en	string	max 255 characters	the name of unit group in English
21	u_group_name_de	string	max 255 characters	the name of unit group in German
22	u_group_name_fr	string	max 255 characters	the name of unit group in French
technical characteristics of the feed samples used for the aggregation				
23	t_id	integer		the unique artificial id for distinct characteristic
24	t_name_en	string	max 255 characters	technical characteristics in English
25	t_name_de	string	max 255 characters	technical characteristics in German
26	t_name_fr	string	max 255 characters	technical characteristics in French

Table 8: Relation *summary_data* consists of 26 attributes in total, out of which 10 are integers, 2 are real numbers and 14 are of string type.

t_formula				
nr	**attribute**	**type**	**constraints**	**description**
1	id	integer	serial, not null	the primary key of the relation
2	abbr	string		abbreviation of the formula that is used for the visualization
3	formula	string		mathematical expression, sometimes recursive, that determines computation of the derived nutrients

Table 9: Relation *t_formula* consists of 3 attributes in total, out of which 1 is integer and 2 are of string type.

t_formula_feed				
nr	**attribute**	**type**	**constraints**	**description**
1	id_feed	integer	not null	reference to the *feed_key* in *d_feed* relation
2	id_formula	integer	not null	reference to the *in* in *t_formula* relation

Table 10: Relation *t_formula_feed* consists of 2 attributes in total, out of which 2 are integers.

users				
nr	**attribute**	**type**	**constraints**	**description**
1	username	string	max 50 characters	the primary key of the relation and the name of the user in the system
Relation *users*, continued in the next page...				

2	password	string	max 50 characters	password of the user encoded with MD5
3	userlevel	integer	default 2	2 stands for subscribed/normal user, other values denote administrators
4	login_counter	integer	default 0	the number of times the user logged in into the system
5	last_log	time stamp	without time zone	time of the last login

Table 11: Relation *users* consists of 5 attributes in total, out of which 2 are integers, 1 is time stamp and 2 are of string type.