THE TFCTN SYSTEM

A food codification and interconnectivity proposal for FCDBs within EuroFIR-AISBL

Ioannis Koutsias, Agricultural University of Athens

Aims of this presentation

- A short introduction to FoodEx2
- A presentation of EFSA's catalogue browser tool and it's significance in effective Foodex2 usage
- A proposal for a unified, standardized, multi - purpose FoodEx2 - based food codification system, that utilizes EFSA's catalogue browser and would enhance interconnectivity between different data of the same FCDB, as well as different countries FCDBs.

FoodEx2

- Food description and classification system
- Foods are organized in hierarchies of progressively narrower groups, until basic food term is reached
- Multiple hiearchies available
- Coding is implemented using terms (ideally extended or core) combined with facets
- Different term detail levels exist
- Different facet categories exist

Hierarchies in FoodEx2

- Hierarchies represent different views of the basic food list and include a base of common terms and specific hierarchy groups.
- Input hierarchies: used to organise foods for data reporting (generation and storage of data in a database)
- Output hierarchies: used for data analysis (grouping in meaningful classes when analysing the data and creating summary tables or documents.

Available hierarchies

- Eight hierarchies are currently available:

- Master hierarchy (entire terminology, for technical use only)
- Reporting hierarchy
- Exposure hierarchy
- Pesticide residues hierarchy
- Zoonoses hierarchy
- Feed hierarchy
- Veterinary drugs residues hierarchy
- Botanicals hierarchy

Exposure hierarchy

- During the development of Gr-FCDB,exposure hierarchy was selected as the codification hierarchy by the AUA FCDB developing team.
- Food items grouping are thus faciliated for exposure calculations. This is also the preferred hierarchy for reporting consumption data.
- The exposure hierarchy is currently structured in six levels (seven in cakes) with 22 groups at the top level.

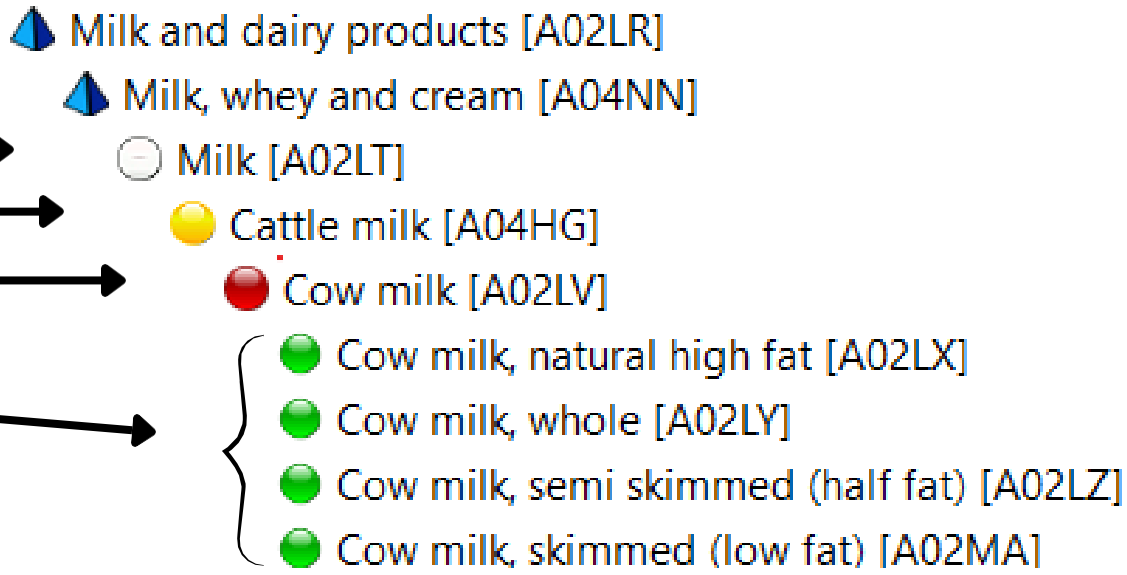
Exposure hierarchy top - level food groups

- Grains and grain-based products [A000J]
- Vegetables and vegetable products [A00FJ]
- Starchy roots or tubers and products thereof, sugar plants [A00ZR]
- Legumes, nuts, oilseeds and spices [A011X]
- Fruit and fruit products [A01BS]
- Meat and meat products [A01QR]
- Fish and seafood [A026T]
- Amphibians, reptiles, and terrestrial invertebrates [A02KP]
- Milk and dairy products [A02LR]
- Eggs and egg products [A031E]
- Sugar and similar, confectionery and water-based sweet desserts [A032F]
- Animal and vegetable fats and oils and primary derivatives thereof [A036M]
- Fruit and vegetable juices and nectars (including concentrates) [A039K]
- Water and water-based beverages [A03DJ]
- Alcoholic beverages [A03LZ]
- Coffee, cocoa, tea and infusions [A03GG]
- Food products for young population [A03PV]
- Products for non-standard diets, food imitates and food supplements [A03RQ]
- Composite dishes [A03VA]
- Seasoning, sauces and condiments [A042N]
- Major isolated ingredients, additives, flavours, baking and processing aids [A046L]
- Other ingredients [A0F0S]

Levels in exposure hierarchy

- Level 1
- Level 2
- Level 3
- Level 4
- Level 5

- Level 6



Terms in FoodEx-2

Hierarchy Term		Represents aggregated groups. Used for system browsing and summarising data analysis results in different food safety domains.	Strongly discouraged
Aggregation Term		Related to a food nature and not to a defined origin (i.e. source, source raw commoditiesor characterising ingredients).	Strictly when more information is not available
Non - specific term		Intermediate aggregated groups used to facilitate the navigation in complex parts of the food list.	Limited to cases where the level of core terms cannot be reached and the non-specific terms provide more detail than the generic term
Core term		These terms represent the minimum recommended level for coding duringdata collection.	Suitable, if the level information of an extended term is not reached
Extended term		More specific and detailed than core term. Not all core term foods meet an extended term’s criteria.	Ideally, if present and detail level is reached

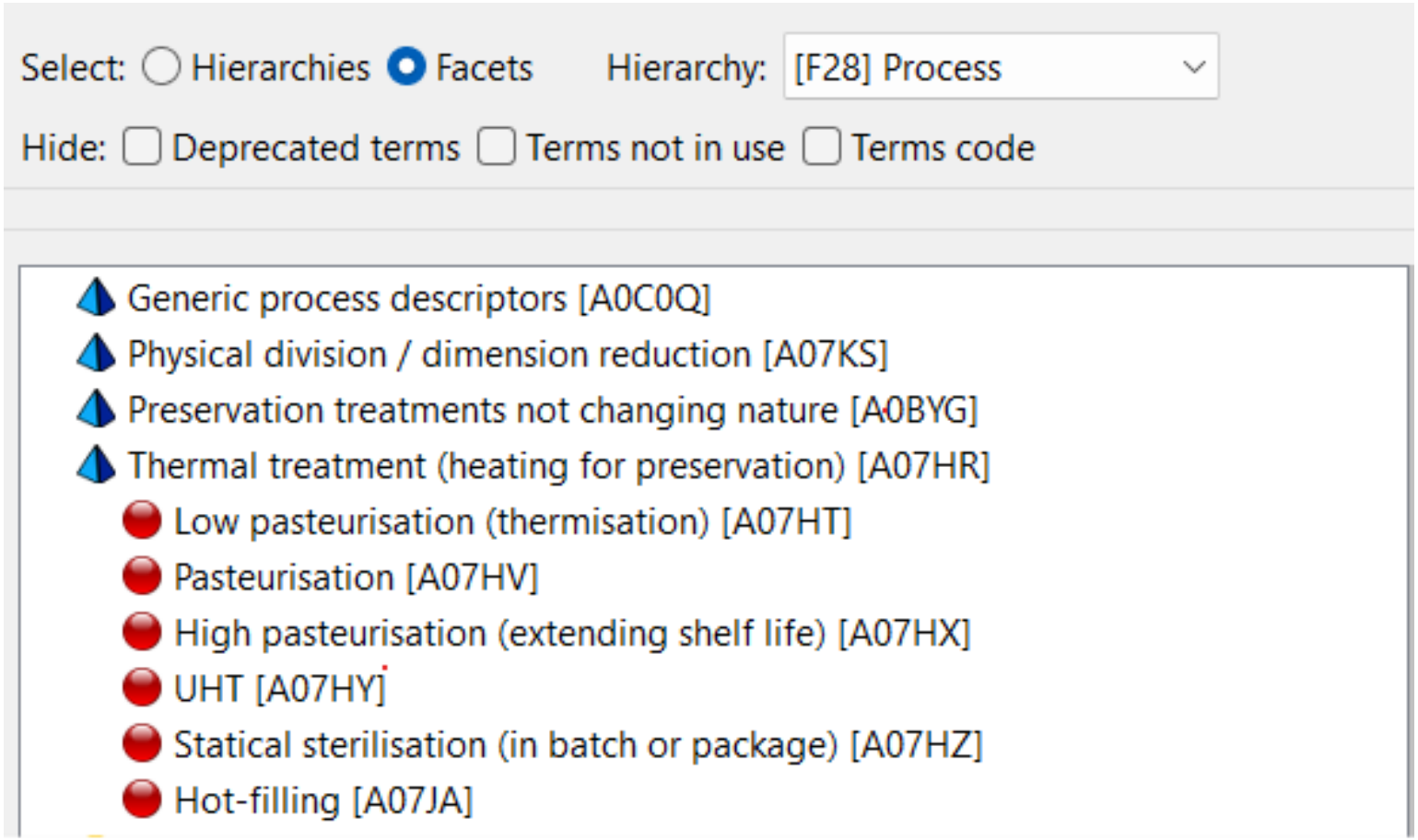
Facets in FoodEx2

- Facets are single descriptors from defined points of view
- Facets may be implicit or added
- Implicit facets: facet descriptors (e.g. source or part nature) implicitly entailed in some terms in the basic food list,and not written when coding.
- Added facets: not implicitly added, added to furtherly describe terms, increasing detail level without the need for creating new terms.

Facets categories

- [F02] Part-nature
- [F01] Source
- [F27] Source-commodities
- [F28] Process
- [F04] Ingredient
- [F06] Surrounding-medium
- [F08] Sweetening-agent
- [F09] Fortification-agent
- [F10] Qualitative-info
- [F17] Extent-of-cooking
- [F26] Generic-term
- [F21] Production-method
- [F18] Packaging-format
- [F19] Packaging-material
- [F03] Physical-state
- [F07] Fat-content
- [F11] Alcohol-content
- [F12] Dough-Mass
- [F20] Part-consumed-analysed
- [F22] Preparation-production-place
- [F23] Target-consumer
- [F24] Intended-use
- [F25] Risky-Ingredient
- [F29] Purpose-of-raising
- [F30] Reproductive-level
- [F31] Animal-age-class
- [F32] Gender
- [F33] Legislative-classes
- [F34] Host-sampled

Adding a specific facet



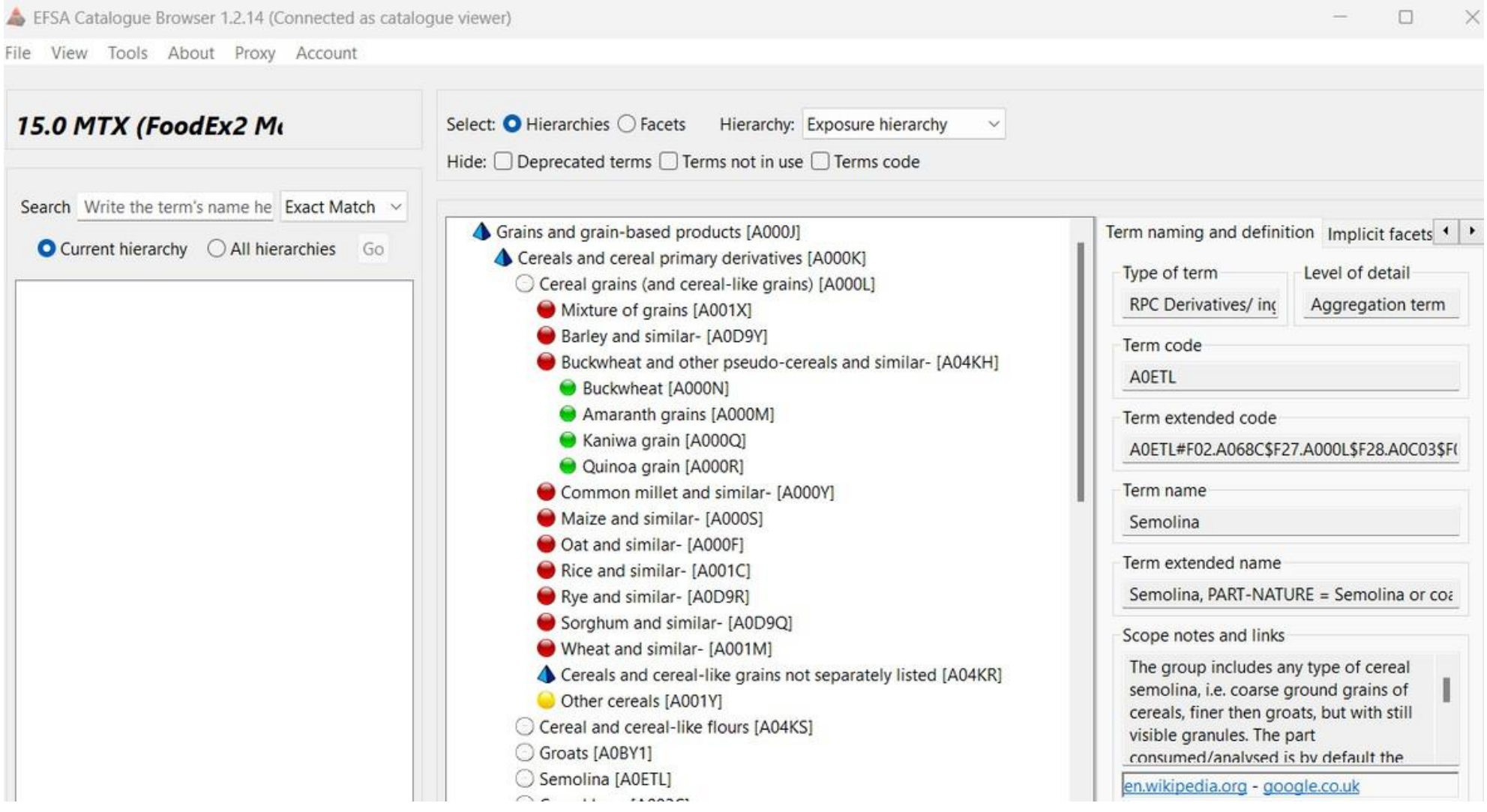
EFSA's Catalogue browser

- Java® based application which allows the browsing, analysis and maintenance of EFSA catalogues.
- It is directly connected with the Data Collection Framework.
- It consists of a number of catalogues, in which individual entries are aggregated within a hierarchical parent-child structure.
- One of the catalogues hosted by the EFSA Catalogue browser is the MTX (FoodEx2 Matrix)

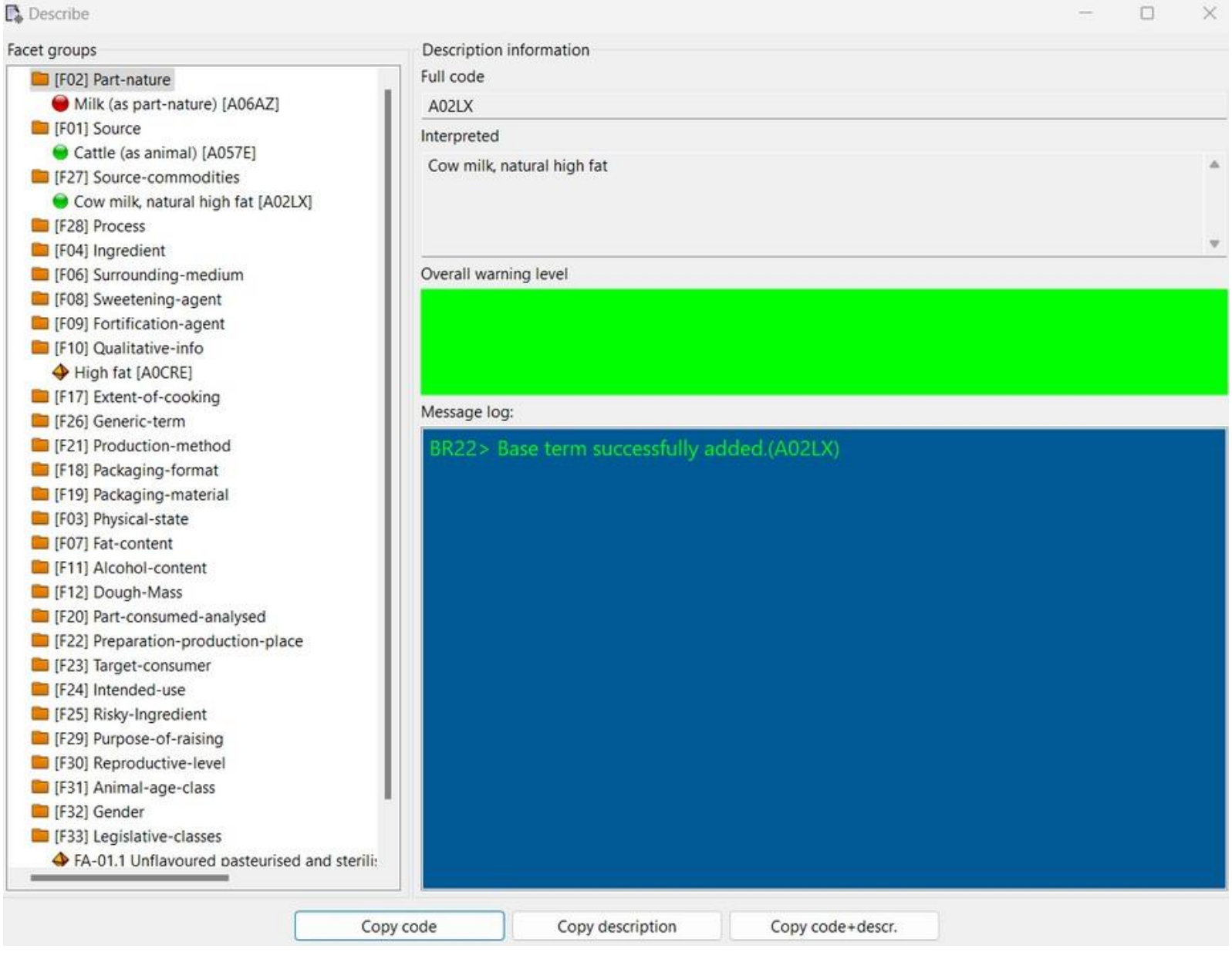
EFSA's Catalogue browser

- The catalogue browser simplifies browsing Foodex2 terms and facets
- Offers guidance in coding
- Gathers all the useful information parameters of a food
- Constantly updated

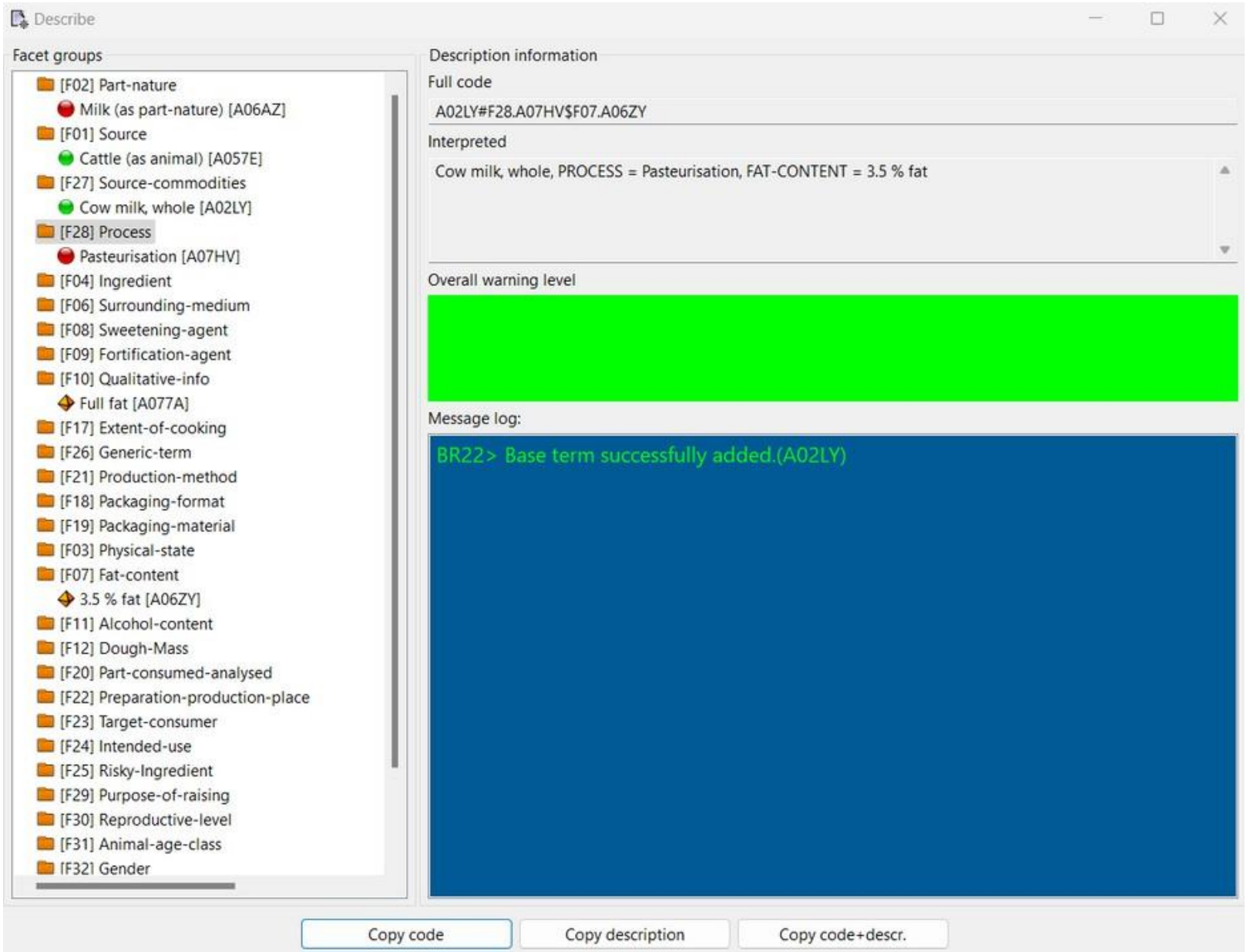
Browsing and available information



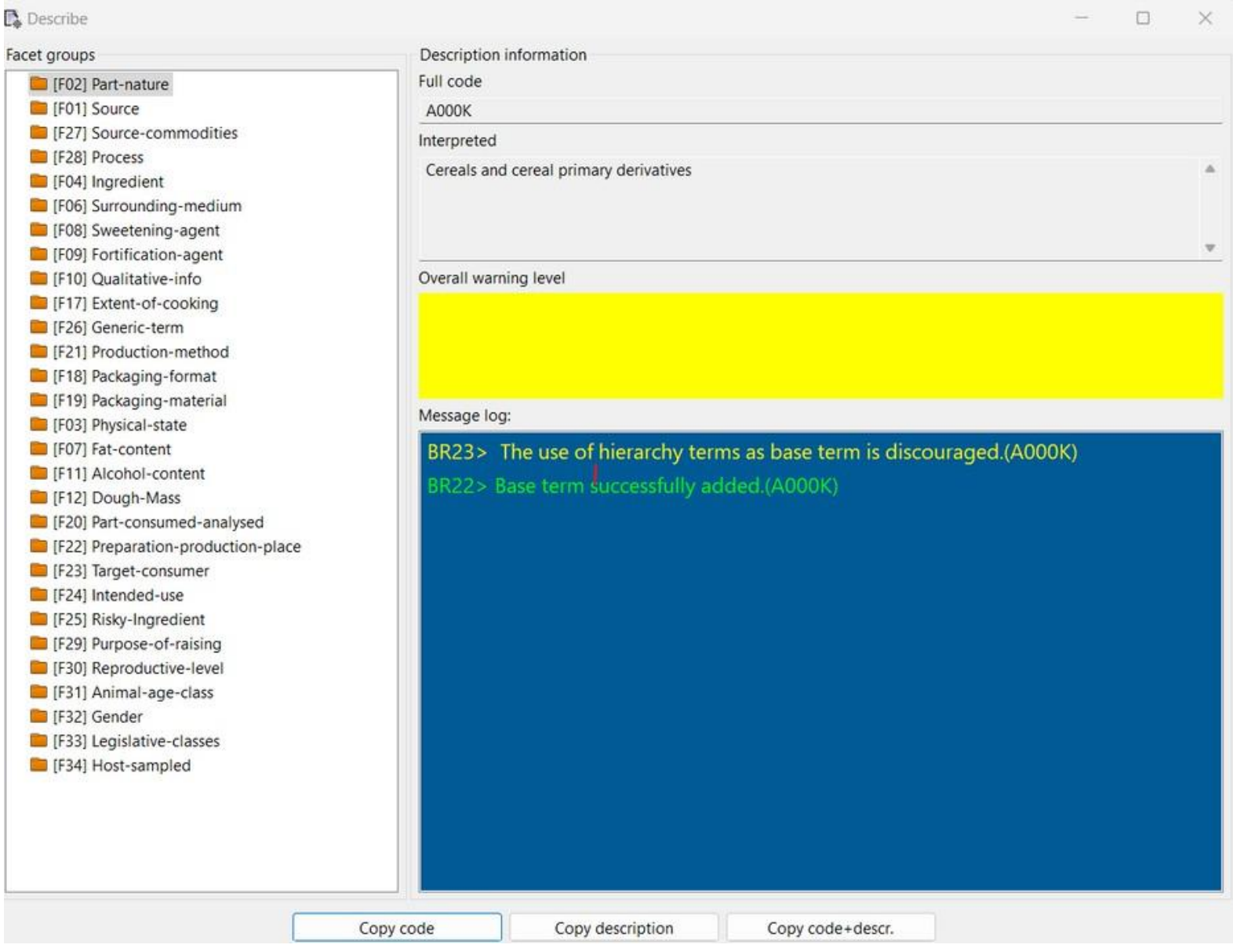
Coding with Catalogue browser-term with its implicit facets



Coding with Catalogue browser-term plus added facets



Coding with Catalogue browser-discouraged coding example



THE TFCTN SYSTEM

- The TFCTN food code is an alphabetical and numerical characters sequence assigned to every FCDB food entry, in the form shown above.
- The different colours indicate distinct areas of food description
- This code allows for universal use and interconnectivity among different dataset types of the same organization, different datasets within a country and different countries within EuroFIR - AISBL.

THE TFCTN SYSTEM

- Internal FCDB codification system
- 16 alphabetical and numerical digits
- FoodEx2 - based
- Data and metadata connectivity
- Generic and branded databases connectivity
- Diffecent same country FCDBs association possibility
- Multiple country FCDBs association possibility
- Standardization

THE TFCTN SYSTEM

Code digits sequence

- Food term (1-6)
- Added facets (7-9)
- FCDB country (10-11)
- FCDB and data type (12-14)
- Data entry serial number (15-16)

Digits 1 - 6; Alphabetical: Food term digits

- Six digits, as many as the levels (almost always) in exposure hierarchy
- Every digit represents a level
- Alphabetical values selected because they provide with more possible values per parameter
- The available options in each level are assigned an alphabetical value, depending on the row they are presented in the Catalogue broswer
- An appendix (1.) should be created, that will list all food terms 6 digit sequence.

Food Term Digits

- A — Grains and grain-based products [A000J]
- B — Vegetables and vegetable products [A00FJ]
- C — Starchy roots or tubers and products thereof, sugar plants [A00ZR]
- D — Legumes, nuts, oilseeds and spices [A011X]
- E — Fruit and fruit products [A01BS]
- F — Meat and meat products [A01QR]
- G — Fish and seafood [A026T]
- H — Amphibians, reptiles, and terrestrial invertebrates [A02KP]
- I — Milk and dairy products [A02LR]
  - IA — Milk, whey and cream [A04NN]
    - IA A — Milk [A02LT]
      - IA A A — Cattle milk [A04HG]
        - IA A A A — Cow milk [A02LV]
          - IA A A A A — Cow milk, natural high fat [A02LX]
          - IA A A A B — Cow milk, whole [A02LY]
          - IA A A A C — Cow milk, semi skimmed (half fat) [A02LZ]
          - IA A A A D — Cow milk, skimmed (low fat) [A02MA]

The term digit value "Z"

- Z is an alphabetical character implying "value not applicable"
- Core and extended terms are found from Level 4 to Level 6
- It is important that all foods have the same code length and therefore all terms from level 4 and below are given a six - digit codification
- If a term is found on the fifth level, the last digit is given a Z value.
- If a term is found on the forth level, the last two digits are given the values ZZ.
- Examples: Cattle milk: IAAAZZ
  Cow milk: IAAAAZ

Digits 7 - 9; Numerical: Added facets digits

- Every generic food should have universally agreed added facets in up to three categories, and in a specific order
- Each digit representsa specific value of this parameter, e.g. if for "cow milk, whole" food term, the first agreed facet represents "Fat content", a value "1" on the seventh code diget would mean "fat content 3.5%", whereas a value "2" would mean "fat content 1.5%"
- An appendix (2.) shall be made for universally agreed facets parameters on generic foods and assigned values.

Generic foods with < 3 facets

- In generic foods with less than three assigned given added facets, any unused facet will be indicated using a value "0" in the corresponding coding digit
- Example 1: a generic food with only 1 added facet, could have a sequence "200" in digits 7 - 8 - 9. (a value 1 for the first parameter, and no other parameter existent.)
- Example 2: a generic food with only 2 added facets, could have a sequence "230" in digits 7 - 8 - 9.

Additional facets in branded foods

- Branded foods may have additional distinction parameters among them, and as such, may make use of more than three added facets.
- To keep the code smaller and for interconnectivity reasons, it was decided to assign facet digits only for facets that generic and branded products have in common.
- Products with additional facets more commonly used in branded products (e.g. packaging), shall be dictated through food serial number (digits 15 - 16)

Digits 10 - 11: FCDB country and compiling organisation

- Digit 10: Alphabetical if country in Europe; numerical if country outside of Europe.
- If all EuroFIR participating countries are arranged in alphabetical order, each country shall be assigned one letter, i.e. the first country in this order is assigned the letter A is its 10th code digit, etc. Same applies to countries outside Europe.

Digit 10: Alphabetical or numerical

- Possible 10th digit assignments:
- Austria A, Belgium B, Bulgaria C, Czech Republic D, Denmark E, Estonia F, Finland G, France H, Germany (if made available) I, Greece J, Iceland K, Ireland L, Italy M, Latvia N, Lithuania O, Netherlands P, Norway Q, Poland R, Portugal S, Serbia T, Slovakia U, Slovenia V, Spain W, Sweden X, Switzerland Y, United Kingdom Z
- Canada 1, Japan 2, New Zealand 3, Turkey 4, United States of America 5
- In case of non-usability due to new country entries, ISO 3166-1 alpha-2 2-digit codes could be used, adding a digit to 17 digits in total.

Digit 11: Numerical Different FCDBs, same country

- In some cases, more than one FCDBs are created by different organizations.
- Existing FCDBs within the same country will be assigned numerical value on digit 11, based on their alphabetical order.
- Any newer entry shall be assigned the smallest possible number at the time of its entering.
- Caution: this digit distincts only different organisations, and not different FCDBs (i.e. generic, branded etc of the same organisation, as this will be addressed by digit 12.

Digits 10 - 11: Examples

- Greece, Agricultural University of Athens: J1
- Greece, Hellenic Health Foundation: J2
- Greece, Medical School of Crete (if made available on Food Explorer): J3

- Italy, CREA, Council for Agricultural Research and Economics: M1
- Italy, IEO, European Institute of Oncology: M2

Digit 12; numerical: Type of data

- Digit 12 addresses the type of the data in discussion.
- Generic food: value "1"
- Branded food: value "2"
- Analytical sample, bibliographical or borrowed value: value "3"

Digit 13; Alphabetical: Type of metadata

- Digit 13 depends on digit 12.
- If digit 12 takes a value "1" or "2", digit 13 takes a value "Z".
- If digit 12 takes a value "3", meaning the entry of discussion is a value for a nutrient, digit 13 letter values will indicate what nutrient the value refers to:

- Energy (kJ): A, Energy (kcal): B, etc (upon agreement)

Digit 14; numerical: Source of metadata

- Digit 14 also depends on digit 12.
- If digit 12 takes a value "1" or "2", digit 14 takes a value "0".
- If digit 12 takes a value "3", meaning the entry of discussion is metadata value for a nutrient a, digit 13 numbers indicate:

- Sample of chemical analysis: 1
- Branded food of same organisation. for generic food calculation through mean values: 2
- Branded food of same country, different organisation: 3
- Branded food of different country:4
- Borrowed same country, other organisation generic food: 5
- Borrowed other country generic food: 6
- Generic from branded mean value: 7
- Peer - reviewed research paper: 8

Digits 15 - 16; numerical: Serial number

- Digits 15 and 16 corresponds to serial numbers 01 - 99 of branded food that correspond to a single generic food on the same organisation's FCDB (existent or not)
- If digit 12 is "1", indicating generic food, digits 15 and 16 take the values 00.

Digits 15 - 16; numerical: Serial number

- Caution: if some branded food data of one organisation, are also used as metadata for generic food values calculation of the same organisation, this means that

1.) there is one branded food code and a metadata code referring to the same food, and still both codes should exist

2. then the digit 15 and digit 16 serial number of the food as metadata, should be the same of the food as branded food, even if that means that there are gaps among branded foods metadata. This should enable traceability of branded metadata as branded foods.

Codes examples

- 1. IAAAAB120-J1-1Z000: This is cow's milk, whole, 3.5% fat(value 1 for fat content facet), pasteurized (value 2 for thermal process), no third facet, Gr - FCDB (AUA's integrated FCDB)entry, generic product (GenGr FCDB). The first six digits indicate an extended term.

- 2. IAAAAZ020-J1-1Z000: Same milk, if we only know it is cow milk, but not its fat content. The first six digits indicate a core term

Codes examples

- 3. IAAAAB120-J1-2Z017: A product in HelTH (Gr - FCDB branded dataset) that is a branded version of 1., which is found in GenGr (Gr - FCDB generic dataset). This specific one is the 17th of its kind in the list.

- 4. IAAAAB120-J1-3C900: This is metadata for 1. calculations. Specifically, digit 12 value "3", means it is metadata, digit 13 value "C" could mean (upon agreement), that this refers to total carbohydrates calculation, digit 14 value "9", that it was calculated by difference and last digits 00 that no multiple samples apply.

Codes examples

- 5. IAAAAB120-J1-3F109: This is metadata for 1. calculations. Digit 13 value "F" could mean protein value, digit 14 value "1", calculated via chemical analysis, and this was sample number 09 used for this purpose.

- 6. IAAAAB120-J1-3J214: This is metadata for 1. calculations. Digit 13 value "J" could indicate saturated fat (upon agreement). Digit 14 value 2 "calculated using HelTH branded values and value "14" that this was HelTH's 14 product of the same category, so this value used for means calculations belongs to HelTH product IAAAAB120-J1-2Z014.

Codes examples

- 7.IAAAAB120-M2-1Z000: The respective product of 1., in the Italian FCDB IEO, which is a generic dataset.

Challenges in TFCTN system use

- TFCTN system offers important standardization and connectivity issues, but it is demanding in terms of complexity. Periodical checking on input errors, or possible updates should be carried out in a systematic way.

- Separate appendixes should be developed and maintained, as:
  - Appendix 1.: Digits 1-6: Food terms
  - Appendix 2.: Digits 7-9 Selected facets and values per food term

- Universal agreement should exist in both indexes, if interconnectivity is thegoal.

Challenges in EFSA's Catalogue Browser use

- EFSA's Catalogue browser is a crucial tool in TFCTN system development, as well as in FCDB Food description in general.
- It provides the compiler with an extensive list of food terms and facets
- Further information is also available, such as term detail level, and term name, code, extended name and extended code.
- Feedback on coding is also offered.

Challenges in EFSA's Catalogue Browser use

- However, there are some issues associated with inconsistencies and/or updates.

- Some subcategories are not always listed in the same row, e.g. in rye bread (Level 4 term), refined flour rye bread comes before wholemeal whereas in rye crisp bread, wholemeal comes first. This becomes an obstacle in the same value meaning the same subcategory (type of flour) for similar but distinct categories (type of bread)

Challenges in EFSA's Catalogue Browser use

- Catalogue browser is constantly online updated, and sometimes entries or groups change places. This means that it would be natural for the first six digits of one entry to constantly have changed values. Possible solutions could be agreement on food coding using a specific date listing of Catalogue Browser as a static lead, or communication with EFSA on the matter of list changes.

- In (only so far) the case of cakes, 7th level terms were found. It was decided that cheese cakes (formerly level 6), was taken out of cream cakes (level 5), and were considered also level 5, so that the 6 level codification would be maintained.

Challenges in FoodEx2 use

- Most FCDBs within EuroFIR - AISBL have been developed using EuroFIR Food Classification as viewed in Langual Thesaurus. For connectivity reasons, Gr-FCDB will preserve any LanguaL code (e.g. A0780 for liquid milk), that best suits a Foodex2 term.

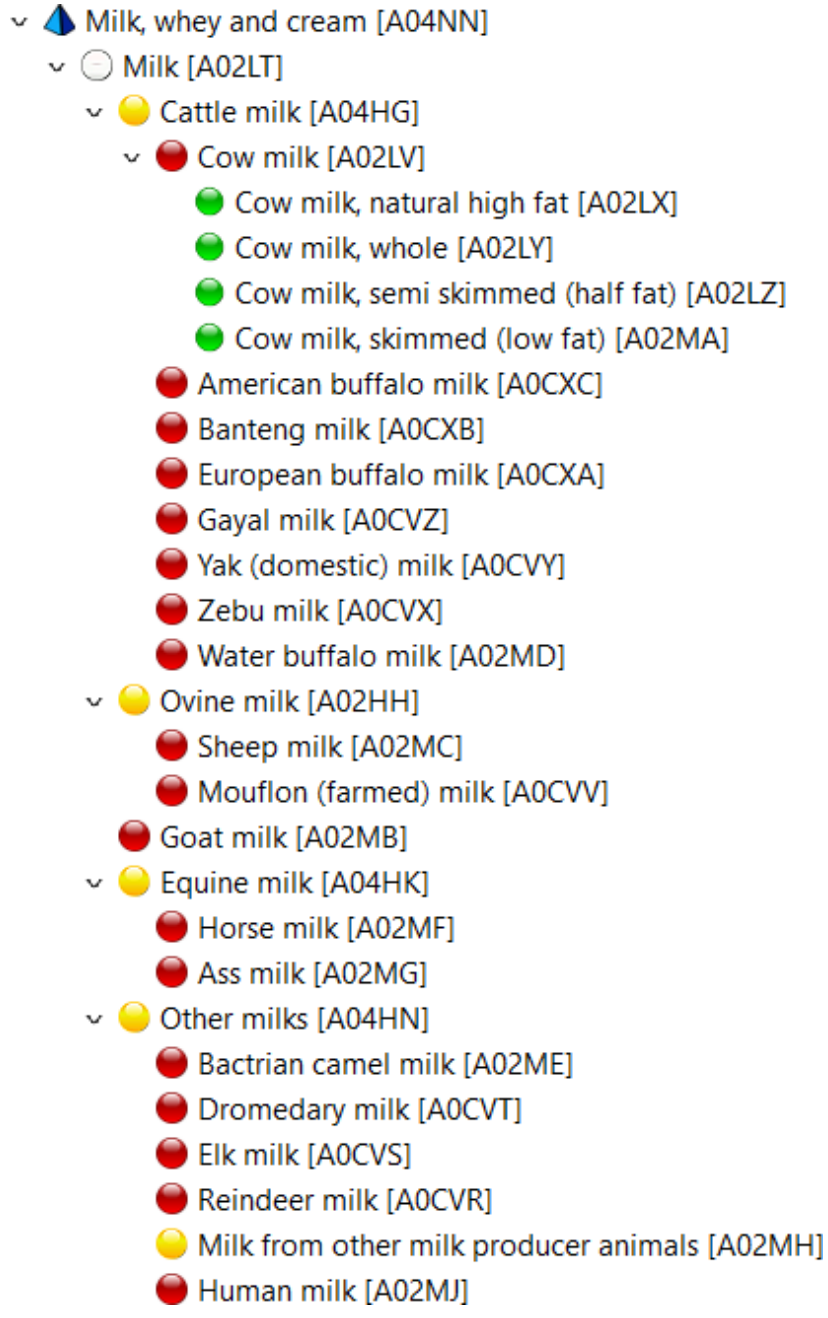
- However, there are differences between Langual EuroFIR food classicication and Foodex2 that should be taken into account.

Challenges in FoodEx2 use

- Some subcategories (levels), might fall in different categories. E.g. "Immitation milk products" fall in the "Milk, milk product or milk substitute" within the EuroFIR classification as seen in Langual Thesaurus, while in Foodex2 immitation products and dairy products are separated.

- One food in Langual EuroFIR food description, might be several possible within Foodex2. E.g.,milk is A0779 and liquid milk is A0780 in EuroFIR food description, while Foodex 2 recognizes a more detailed spectrum of terms for the same products, as shown in the following slide (those are separare terms, without any added facets):

Challenges in FoodEx2 use



- Each term should have its own first 6 digit sequence and should be used, depending on what information is available (as explained in slide 9).

Conclusions

- The TFCTN system is a food codification system that could be used to connect different kinds of data and their metadata within a specific FCDB, as well as possibly interconnect any FCDBs within EuroFIR - AISBL.

- There is a systematic approach in providing information about the food term hierarchy, added facets, country and specific organisation of origin, type of data and metadata obtainment, as well as numerous possible connections among respective entries of different FCDBs, all using a single code.

- Challenges associated with the system itself, EFSA's Catalogue browser and Foodex2 in general exist and should be taken into account when implementing the system.

Resources

- EuroFIR - AISBL: https://www.eurofir.org/
- Langual Thesaurus: https://www.langual.org/langual_Thesaurus.asp
- FoodEx2: https://www.efsa.europa.eu/en/supporting/pub/en-804
- EFSA Catalogue browser: https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/sp.efsa.2019.EN-1726

The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the 4th Call for HFRI PhD Fellowships (Fellowship Number: 011172).

Ioannis Koutsias, Chemist MSc, PhD candidate, Department of Food Science and Human Nutrition, Agricultutal University of Athens
ioanniskoutsias@aua.gr
Supervisor: Dr Maria Kapsokefalou, Professor, Agricultural University of Athens
kapsok@aua.gr