**on\_market.export\_urls.txt**

This file contains data of all the products that are available on the market currently. There is at least one distributors who gets this product on market. Some important attributes of the product are like,

**Product\_id** -icecat's internal product identifier

**Prod\_id** - manufacturer part number

**Quality** = Icecat ( icecat editors have standardized the data-sheet)

= Supplier (data sheet is auto imported from mfg CMS)

= Noeditor (raw data received from channel partners)

= removed (no more exist)

**URL** - link for product full description and specification

**supplier\_id** - manufacturer id

**High\_res\_img** - link to high-resolution image of the product

**Low\_res\_img** - link to low-resolution image of the product

**Thumbnail\_img** - link to thumbnail image of the product

**Category\_ID** - icecat's internal category identifier

**M\_prod\_id**  - manufacturer part number variants used by distributors

**ean\_upcs** - logistics code depending on packaging and geography

**model\_name**  - product name

**product\_view** - how many times current product was requested

**on\_market** - indicates presence of product in market currently

**country\_market\_set** - country id where product is available as per icecat’s record

**categories.xml**

It states the hierarchy information of different categories of product in icecat open catalog dataset. For each category, their attributes gives information about their ID, image url link, score. The score attribute reflects the usage statistics of the category. It has child tags like Description, Keywords, ParentCategory and Name. The <Description> tag contains a short summarization of the category in terms of its definition and usage. The <Keywords> tag contains a list words which are regularly used with the category. The <ParentCategory> tag contains the vital information about the immediate parent of the category. It is worth noting that the parent category is also a category and so like other categories it has its own set of information content. The <Name> tag does the nomenclature of the category. They are specified for different languages mentioned in language list.

**categorization.txt**

This file contains one to one mapping information of category id to category name. It has three fields catid, uncatid and name. These files are available in other languages like English, Dutch and French.

**daily.export\_urls.txt**

This file contains the information about the new or altered products data sheet of the respective day. It is generated on a daily basis. It also lists products to be removed which are no longer available from the day before. Such products can be identified if they have ‘QUALITY = REMOVED’ as their field value.So, it can be used to update the existing catalog of data.

**daily.export\_urls\_rich.xml**

This file details the new suppliers that are added on the respective day. It also lists the new product which have multi valued entries for ‘EAN\_UPCS’ and ‘Country\_Markets’. They represent the logistic code based on packaging and geography and country where products became active, respectively.

**export\_suppliers.txt**

This text file contains the list of new added supplier id along with their name.

**measure.xml**

This file lists all possible measuring units used for different products. Each measure unit has an ID which are used while describing the feature list of products and categories. The description of measure units are available in different languages

**supplier\_mapping.xml**

All supplier, also known as manufacturers, have a unique id and an original name. But the name of the supplier is variant across different distributors. So, the purpose of this file is to provide a one to many mapping of supplier ids to many names as symbolized by different distributors. This is done to have a consistent meaning for different names used.

**supplierlist.xml**

List of supplier or manufacturer with their name and id.

**languagelist.xml**

It contains information about different languages supported by Icecat as mentioned below. Information like language id, code and name are specified.

1. US English

2. Dutch

3. French

4. German

5. Italian

6. Spanish

7. Norwegian

8. Russian

9. Chinese (traditional)

10. Swedish

11. Polish

12. Slovakian

13. Finnish

14. Greek

15. Turkish

16. Romanian

17. Ukrainian

18. Japanese

19. Arabic

20. Vietnamese

21. Lithuanian

22. Korean

23. Latvian

24. Indonesian

25. Estonian

26. Thai

27. Danish

28. Portuguese

29. Chinese (simplified)

30. Czech

31. Hungarian

32. Bulgarian

33. Georgian

34. Serbian

35. Catalan

36. Croatian

37. Hebrew

38. Brazilian Portuguese

39. Indian English

40. Swiss German