

Preliminary proposal to add emoji characters for food allergens

To: UTC

Author: Hiroyuki Komatsu <komatsu@google.com>









Last updated: 2015-07-26

Disclaimer










This document is a proposal whose contents are subject to UTC review, and any and all details are therefore subject to change. Any characters that are contained in this proposal, along with their attributes, including, but not limited to, their names, code points, and representative glyphs, should not be construed as final. There is also no guarantee that the characters will be added to the Unicode standard.

Proposal

We propose to add the following Emoji characters to represent food allergens.





| Code point | Name | Color Sample | To represent products containing: |
|------------|------------|---|---|
| U+1F962 | PEANUT |  | Peanuts |
| U+1F964 | SOYBEAN |  | Soybeans |
| U+1F95A | BUCKWHEAT |  | Buckwheat |
| U+1F963 | SESAME |  | Sesame seeds |
| U+1F95F | KIWI FRUIT |  | Kiwi fruit |
| U+1F95B | CELERY |  | Celery |
| U+1F960 | LUPIN |  | Lupin beans |
| U+1F961 | MUSTARD |  | Mustard (could also use jar of mustard, or yellow squeeze-bottle) Maybe also the plant. |















The following have existing Emoji representations that might be close enough to use, but the Emoji subcommittee recommends encoding new characters:











| Code point | Name and annotations | Color Sample | To represent products containing: | Alternative, use existing character | Color sample |
|------------|---|--|-----------------------------------|-------------------------------------|---|
| U+1F965 | WALNUT <ul style="list-style-type: none"> tree nuts in general |  | Tree nuts | CHESTNUT U+1F330 |  |
| U+1F95C | EGGS <ul style="list-style-type: none"> egg products |  | Eggs | COOKING U+1F373 |  |
| U+1F95D | GLASS OF MILK <ul style="list-style-type: none"> dairy products shown with pitcher pouring in |   | Milk products | BABY BOTTLE U+1F37C |  |
| U+GRAIN | GRAIN <ul style="list-style-type: none"> contains gluten |  | Gluten | BREAD U+1F35E |  |

Background

The following list shows the current coverage of Emoji characters for food allergens. The details are described later.

| Codex recommendations | representations | status | code | b&w | Android | name | annotations |
|---------------------------|-----------------|------------------|---------|---|--|----------|---|
| peanuts | peanuts | not available | n/a | | | | https://en.wikipedia.org/wiki/Peanut |
| soybeans | soybeans | not available | n/a | | | | https://en.wikipedia.org/wiki/Soybean |
| tree nuts | almond | not available | n/a | | | | https://en.wikipedia.org/wiki/Almond |
| | nuts | weak alternative | U+1F330 |  |  | chestnut | chestnut, nature, plant |
| cereals containing gluten | wheat | weak alternative | U+1F35E |  |  | bread | bread, food, loaf, nature, object, restaurant |

| eggs | egg | weak alternative | U+1F373 |  |  | cooking | cooking, egg, food, frying, nature, object, pan, restaurant |
|---------------------|-----------------|------------------|---------|---|--|--------------|---|
| milk | milk | weak alternative | U+1F37C |  |  | baby bottle | baby, bottle, drink, milk, nature, object |
| crustaceans | shrimp | weak alternative | U+1F364 |  |  | fried shrimp | food, fried, nature, object, prawn, restaurant, shrimp, tempura |
| | crab | available | U+1F980 |  | missing | crab | crab |
| fish | fish | available | U+1F41F |  |  | fish | animal, fish, nature, pisces, zodiac |
| sulphite > 10 mg/kg | n/a | | | | | | |
| | | | | | | | |
| other allergens | representations | status | code | b&w | Android | name | annotations |
| buckwheat | buckwheat | not available | n/a | | | | https://en.wikipedia.org/wiki/Buckwheat |
| sesame | sesame | not available | n/a | | | | https://en.wikipedia.org/wiki/Sesame |
| celery | celery | not available | n/a | | | | https://en.wikipedia.org/wiki/Celery |
| lupin | lupin | not available | n/a | | | | https://en.wikipedia.org/wiki/Lupinus |
| mustard | mustard | not available | n/a | | | | https://en.wikipedia.org/wiki/Mustard_(condiment) |
| kiwi fruit | kiwi fruit | not available | n/a | | | | https://en.wikipedia.org/wiki/Kiwifruit |
| mollusks | squid | not available | n/a | | | | https://en.wikipedia.org/wiki/Squid |
| | octopus | available | U+1F419 |  |  | octopus | animal, nature, octopus |
| orange | orange | available | U+1F34A |  |  | tangerine | food, fruit, nature, object, orange, plant, tangerine |
| peach | peach | available | U+1F351 |  |  | peach | food, fruit, nature, object, peach, plant |

| | | | | | | | |
|--------|---------|-----------|---------|---|--|-----------|--|
| apple | apple | available | U+1F34E |  |  | red apple | apple, food, fruit, nature, object, plant, red |
| banana | banana | available | U+1F34C |  |  | banana | banana, food, fruit, nature, object, plant |
| | beef | available | U+1F42E |  |  | cow face | animal, cow, face, nature |
| | pork | available | U+1F437 |  |  | pig face | animal, face, nature, pig |
| meat | chicken | available | U+1F414 |  |  | chicken | animal, chicken, nature |

Background

Emoji should cover characters representing major food allergens. It enables people to understand what nutrients are used in foods even in foreign countries and safely select meals.

Codex Alimentarius, a collection of internationally recognized standards, codes of practice, guidelines and other recommendations relating to foods, food production and food safety (http://en.wikipedia.org/wiki/Codex_Alimentarius), recommends that the following foods and ingredients are known to cause hypersensitivity and shall always be declared. (http://www.codexalimentarius.org/input/download/standards/32/CXS_001e.pdf)

- **Cereals containing gluten**; i.e., wheat, rye, barley, oats, spelt or their hybridized strains and products of these;
- **Crustacea** and products of these;
- **Eggs** and egg products;
- **Fish** and fish products;
- **Peanuts, soybeans** and products of these;
- **Milk** and milk products (lactose included);
- **Tree nuts** and nut products; and
- Sulphite in concentrations of 10 mg/kg or more.

FDA, a federal agency of the United States Department of Health and Human Services (https://en.wikipedia.org/wiki/Food_and_Drug_Administration) uses the same list of food allergens with Codex but treats peanuts and soybeans as different groups. (<http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/Allergens/ucm106890.htm>)

EFSA, the agency of the European Union (EU) that provides independent scientific advice and communicates on existing and emerging risks associated with the food chain

(https://en.wikipedia.org/wiki/European_Food_Safety_Authority) specifies **sesame**, **celery**, **lupin** and **mustard** and treats peanuts and soybeans separately in addition to Codex. (<http://www.efsa.europa.eu/en/press/news/141126.htm>)

JAS, *standards for the agriculture industry maintained by the Japanese Government* (https://en.wikipedia.org/wiki/Japanese_Agricultural_Standard) specifies **buckwheat** and treats **shrimp** and **crab** separately in addition to Codex. (<http://www.caa.go.jp/foods/pdf/syokuhin13.pdf>)

We should have emoji characters representing above food allergens, and also some additional major allergens as described in the above table.

Discussion per character

Tier 1: Codex recommendations, no representations / peanuts, soybeans, almond, nuts

Peanuts and soybeans are food allergens recommended to be always declared. The current emoji set do not have alternatives for them.

Tier 2: Codex recommendations, weak alternatives / cereals containing gluten, tree nuts, egg, milk

Cereals containing gluten are wheat, rye, barley, oats, spelt, etc. For wheat, U+1F35E (🍞, BREAD) could be an alternative, but it is far from cereals containing gluten. U+1F33E (🌾, EAR_OF_RICE) looks like wheat. However it specifies rice and rice is gluten free. We propose a new emoji character for “**grain**”.

Tree nuts are almonds, walnuts, pecans, etc. U+1F330 (🌰, CHESTNUT) could be a representation, but a chestnut may not be common enough to represent tree nuts. We propose “**walnut**” for a new emoji character, or possibly “**tree nuts**”. The images of tree nuts may contains various type of nuts like walnut, cashew, pistachio, etc.

Egg could be alternated by U+1F373 (🍳, COOKING). However it might be confusing in some context. U+1F373 could be misunderstood as ‘baked’ against ‘raw’.

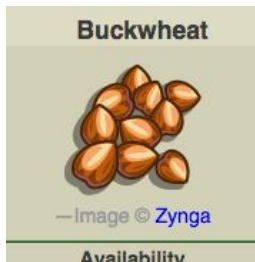
Milk could be alternated by U+1F37C (🍼, BABY_BOTTLE). However it might be confusing in some context. U+1F37C could be misunderstood as edible for babies.

Tier 3: Other allergens, no representations / buckwheat, sesame, celery, lupin, mustard, kiwi fruit

In addition to Codex recommendations, there are several well-known food allergens. In some countries, those foods and ingredients shall always be declared. (e.g. buckwheat in Japan and Korea / sesame, celery, lupin and mustard in EU). Fortunately, most of them are already covered by the current Emoji character set as listed in the above table.

Buckwheat, sesame, celery, lupin, mustard and kiwi fruit are also food allergens. Kiwi fruit is also a good candidate to fill the gap in popular fruits.

For a visual representation of buckwheat, we suggest to use several seeds of buckwheat.



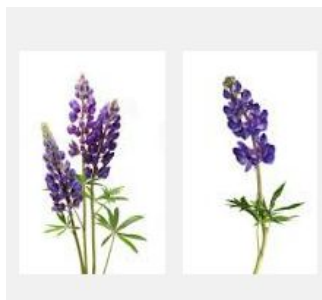
<http://farmville.wikia.com/wiki/Buckwheat>

For a visual representation of sesame, we suggest to use a pile of sesame seeds.



"不老長寿の秘薬"とも呼ばれるゴマ。その理由は: <http://calcium-kazoku.jp/character/gomata.html>

For a visual representation of lupin, we suggest to use flowers of lupin.



<https://goo.gl/E0SdBs> (Google image search)


For a visual representation of mustard, we suggest to use a bottle of mustard


images for mu



<https://www.google.co.jp/webhp?#q=mustard%20bottle>

Tier 4: Good supplemental / shrimp, squid

Shrimp is a representation of “crustaceans” in addition to U+1F980 (CRAB). U+1F364 (, FRIED_SHRIMP) is available as a dish, but there is not shrimp character for animal.

Squid is a representation of “mollusks” in addition to U+1F419 (, OCTOPUS). For squid, there is no emoji character.

In terms of filling the gaps in the existing set of common animals, shrimp and squid are good candidates.

Appendix.

- Emoji survey for food allergens:
https://docs.google.com/spreadsheets/d/1UnaJ4TFVXiVqyNnK5nwkKC_1LeHmTpzNMy9CzXTs-Vo/edit#gid=0
- Codex Alimentarius:
<http://www.codexalimentarius.org/>

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from
<http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before
filling this form.

Please ensure you are using the latest Form from
<http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest *Roadmaps*.

Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04,
1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03,
2008-05, 2009-11, 2011-03, 2012-01)

A. Administrative

| | |
|--|--|
| 1. Title: | Proposal to add emoji characters for food allergens |
| 2. Requester's name: | Hiroyuki Komatsu |
| 3. Requester type (Member body/Liaison/Individual contribution): | Individual contribution |
| 4. Submission date: | 2015-07-23 |
| 5. Requester's reference (if applicable): | |
| 6. Choose one of the following: | |
| This is a complete proposal: | <input checked="" type="checkbox"/> |
| (or) More information will be provided later: | <input type="checkbox"/> |

B. Technical - General

| | | | | | |
|--|---|---|-------------------------------------|------------------------------------|--------------------------|
| 1. Choose one of the following: | | | | | |
| a. This proposal is for a new script (set of characters): | <input type="checkbox"/> | | | | |
| Proposed name of script: | | | | | |
| b. The proposal is for addition of character(s) to an existing block: | <input checked="" type="checkbox"/> | | | | |
| Name of the existing block: | Supplemental Symbols and Pictographs | | | | |
| 2. Number of characters in proposal: | 12 | | | | |
| 3. Proposed category (select one from below - see section 2.2 of P&P document): | | | | | |
| A-Contemporary | <input type="checkbox"/> | B.1-Specialized (small collection) | <input checked="" type="checkbox"/> | B.2-Specialized (large collection) | <input type="checkbox"/> |
| C-Major extinct | <input type="checkbox"/> | D-Attested extinct | <input type="checkbox"/> | E-Minor extinct | <input type="checkbox"/> |
| F-Archaic Hieroglyphic or Ideographic | <input type="checkbox"/> | G-Obscure or questionable usage symbols | <input type="checkbox"/> | | |
| 4. Is a repertoire including character names provided? | | YES | | | |
| a. If YES, are the names in accordance with the "character naming guidelines" | | YES | | | |
| b. Are the character shapes attached in a legible form suitable for review? | | YES | | | |
| 5. Fonts related: | | | | | |
| a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? | | | | | |
| b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): | | | | | |

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

YES

b. Are published examples of use (such as samples from newspapers, magazines, or other sources)

of proposed characters attached?

YES

7. Special encoding issue

Does the proposal address other aspects of character data processing (if applicable) such as input,

presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

NO

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see UAX#44: <http://www.unicode.org/reports/tr44/> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?

NO

If YES explain

2. Has contact been made to members of the user community (for example: National Body,

user groups of the script or characters, other experts, etc.)?

NO

If YES, available relevant documents:

3. Information on the user community for the proposed characters (for example:

size, demographics, information technology use, or publishing use) is included?

Over 110 million people in China use emoji daily, over 44 million in the U.S. use emoji daily

Reference:

| | |
|--|---------------|
| 4. The context of use for the proposed characters type of use; common or rare) | common |
| Reference: | |
| 5. Are the proposed characters in current use by the user community? | NO |
| If YES, where? Reference: | |
| 6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? | NO |
| If YES, is a rationale provided? | |
| If Yes, reference: | |
| 7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? | NO |
| 8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? | NO |
| If YES, is a rationale for its inclusion provided? | |
| If Yes, reference: | |
| 9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? | NO |
| If YES, is a rationale for its inclusion provided? | |
| If Yes, reference: | |
| 10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? | NO |
| If YES, is a rationale for its inclusion provided? | |
| If Yes, reference: | |
| 11. Does the proposal include use of combining characters and/or use of composite sequences? | NO |
| If YES, is a rationale for such use provided? | |
| If Yes, reference: | |
| Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? | |
| If Yes, reference: | |
| 12. Does the proposal contain characters with any special properties such as | |

control function or similar semantics?

NO

If YES, describe in detail (include attachment if necessary)

[Redacted area]

13. Does the proposal contain any Ideographic compatibility characters?

NO

If YES, are the equivalent corresponding unified ideographic characters identified?

If Yes,
reference:

[Redacted area]