# iUnit Definition in 1CLICK-2

An iUnit for a search topic is a factual statement that is *relevant, atomic, credible and temporally valid*. These four properties are defined as follows:

## 1. Relevance:

The factual statement *alone* satisfies the information need behind the search topic either wholly or partially. Thus, it does not require other iUnits to be present in order to provide useful information.

e.g.

1. Tetsuya Sakai was born in 1988.
2. \*Tetsuya Sakai was born.\*

If the information need is "Who is Tetsuya Sakai that keeps posting these silly tweets?", (b) alone is probably not useful and therefore this is not an iUnit.

## 2. Atomicity:

The factual statement cannot be broken down into multiple iUnits. Thus, if it is broken down into several statements, at least one of them does not pass the relevance test. Whereas, it is possible that an iUnit entails another iUnit.

e.g.

1. Takehiro Yamamoto received a PhD from Kyoto U.
2. Takehiro Yamamoto has a PhD.

Clearly, (a) entails (b) and they are both iUnits.

## 3. Credibility:

The factual statement is explicitly stated in at least one document. However, wherever there is a conflict between two statements caused by multiple sources, we take the most "reliable" statement as the iUnit according to the credibility ranking of sources we define elsewhere. Thus, we avoid conflicts among iUnits for a given search topic.

e.g.

1. Kawori Manabe was born on March 31, 1981 [according to Nihon Talent Meikan]
2. Kawori Manabe was born on May 31, 1980 [according to the official blog/wikipedia]

We take the information from the official blog as more reliable, so create an iUnit for (b), but not for (a).

## 4. Temporal Validity:

The factual statement holds true as of the date specified elsewhere ("snapshot date").

That is, the information is not obsolete at the time of the snapshot date.

e.g.

1. Kawori Manabe's official blog is www.manabekawori.cocolog-nifty.com [blog closed on Oct 31, 2010]
2. Kawori Manabe's official blog is blog.livedoor.jp/kaworimanabe [blog active as of Feb 21, 2012]

If the specified snapshot date is Dec 31, 2009, then we create an iUnit from (a), not from (b). If the specified snapshot date is Feb 21, 2012, then we create an iUnit from (b), not from (a).

# Nugget Definition in Previous Work

## 1CLICK-1

(a) A nugget is a short factual statement such that an assessor can judge whether a given text shows or clearly implies that statement to be true.

(b) Information available on official Web pages is considered factual. Information available from other Web sources is considered factual provided that it does not contradict with the official information.

(c) Nuggets are built based on established facts as of December 31, 2010. Events that occur after this date will be ignored.

## Virgil’s WSDM2012 paper

The smallest portion of text that constitutes relevant information in and of itself. However, nuggets are not restricted to text as it appears in the document: slight modifications of the text, e.g. co-reference disambiguation, deleting contextual stopwords, etc. were encouraged.

## GALE Distillation evaluation

They defined nuggets based on the following rules:

1. Nuggets are created out of each core verb and its arguments, where the maximal extent of the argument is always selected.

2. All temporal, locative, causative and other types of modifiers of the verb constitute a separate nugget, including subordinate clauses, where the maximal extent of the modifier is always selected.

3. Noun phrases are not decomposed into separate nuggets, unless they contain temporal, locative, numerical information, or titles.

## TREC QA

* A nugget is defined as a fact for which the assessor could make a binary decision as to whether a response contained that nugget.
* An information nugget is defined as an atomic piece of information about the target that is interesting.