

---

# Best practice for knowledge base creation

---

Organising multiple knowledge bases and  
assessing formal terms for meaning and  
consistency

---

# How many knowledge bases?

- If knowledge acquisition involves two or more groups that MAY have distinct knowledge (e.g. different tribal groups or farm workers and extension workers), assume their knowledge will be different and develop separate knowledge bases.
    - for example, we could assume that farmers and vegetable wholesalers have different knowledge about pesticide pollution. It would be good practice to develop a **farmer** knowledge base and a **trader** knowledge base about effects of pesticides.
    - Informants who use different taxonomic classifications **MUST** have distinct knowledge bases
  - It is better to have two smaller closely related knowledge bases than to have one large 'confused' one
    - Knowledge bases can be merged at a later date if there is no difference between strata's knowledge
-

---

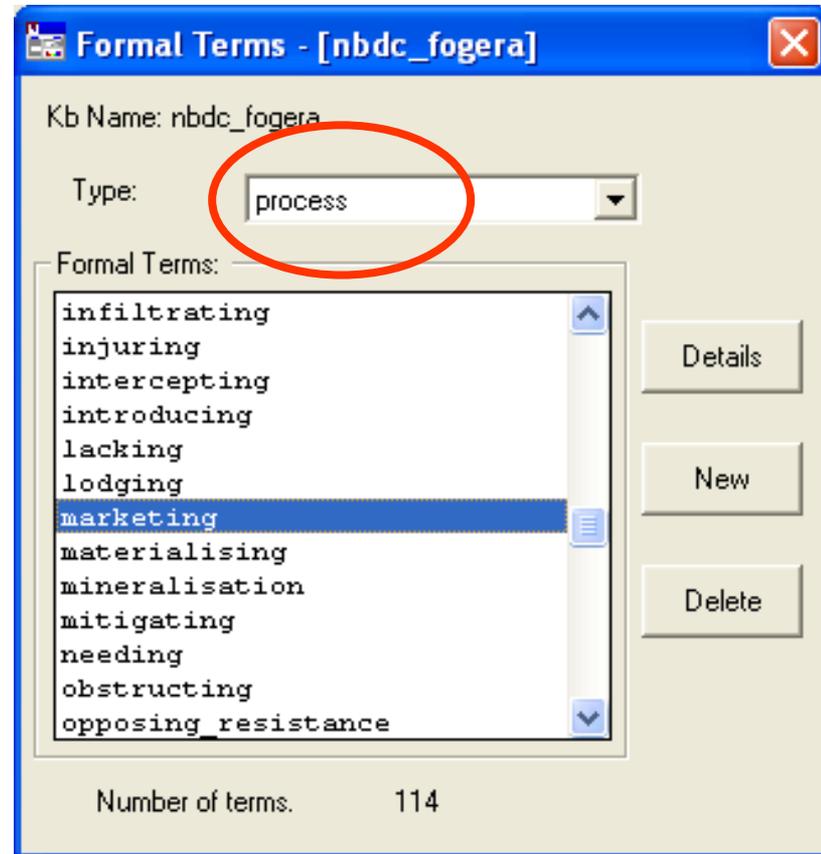
Development of an internally consistent knowledge base requires iterative evaluation of the knowledge base

Key areas of evaluation:

1. **Formal terms**
  2. **Individual unitary statements**
  3. **Sets of unitary statements**
  4. **The relationships between formal terms**
-

# Evaluation of formal terms (1)

Check to ensure that formal terms have been **categorised properly** e.g. “marketing” is not a process, it is an action.



---

# Evaluation of formal terms (2)

## Spelling and consistency

- Be consistent with plurality
    - e.g. choose **either** goat or goats
  - Check for spelling mistakes
    - e.g. gaot
  - Avoid using different terms to mean the same thing
    - e.g. amount / number / quantity
    - use one consistently where possible
-

## Evaluation of formal terms (3)

- Check that object terms have the correct parts
- Add relevant synonyms and ensure consistency of their number order throughout the KB

The screenshot shows a software window titled "Formal Term Details - [ego]". It contains the following fields and controls:

- Formal Term:** A text box containing "koto\_ohamdi".
- Type:** A dropdown menu set to "object".
- Part of:** An empty text box.
- Parts:** A list box containing "rhizome", "root", and "part(rhizome,root)". A red arrow points from the "Check that object terms have the correct parts" bullet point to this list.
- Definition:** A text area containing "A member of the Marantacea family. Its leaves are mainly used for covering kola nuts, and have been known to be".
- Synonym(s):** A list box containing "1. Ataenidia conferta" and "2. ntentrema".
- Controls:** "up" and "down" buttons for the synonym list; "add" and "delete" buttons for adding and removing synonyms; "Save" button; "Show use in statements" and "Show use in hierarchies" buttons.
- Image:** A small photograph of a person in a blue shirt working with green plants in a field.

# Definitions

- Full definition of locally specific terms is very important  
*Remember that readers of the knowledge base may not have visited the research location at all.*
- Add definitions as you add new terms – most people do not and often leave many key terms undefined
- Provide full definitions for terms that may be ambiguous e.g. **'quality'** or **'land\_preparation'** or **'farm'**
- Add photos to your knowledge base



---

## Evaluation of formal terms (4)

- Look carefully at long formal terms, for example:

“human\_control\_over\_plant\_lifecycle”

(This was originally listed as an object!)

This is a ‘scoping’ term – i.e. a holdover from the scoping phase and needs to be broken down further into unitary statements.

- However “in\_upper\_part\_of\_catchment” would be fine.
-

---

# Evaluation of individual unitary statements

- Statements must be genuine unitary statements
  - Statements must be **relevant**
    - e.g. *Coffea arabica* has red berries when ripe
  - Statements must be **useful**
    - e.g. manure production influences crop production?
-

# Complete and explicit specification of meaning

- Clear and unambiguous construction of statements in the formal language

part('Celtis australis,leaves) causes1way process(sickness)

Would be better represented as

'action(feeding, part('Celtis australis', leaves), livestock)  
causes1way att\_value(livestock, health\_status, sick)'

- Use of conditions
  - E.g. the sole\_cropping of Canavalia ensiformis time is major\_season causes the Canavalia ensiformis dry\_matter accumulation is in the range 3.2t/ha to 4.0t/ha  
**if the Canavalia ensiformis location is transition\_zone\_of\_Ghana and the growth of Canavalia ensiformis duration is 4 months**
- Use of memos where necessary
- Use of knowledge categories for location specific information

---

# Evaluating sets of unitary statements

## ■ Repetition

- Are there spelling mistakes that have caused repetition of statements?

Trampling by **goats** causes soil creep

Trampling by **goat** causes soil creep

Trampling by **gaots** causes soil creep

## ■ Contradiction

- Not necessarily wrong!
  - If statements are contradictory, go back to interviewees to see why this might be
-

---

# Completeness

- Through the evaluation of a set of unitary statements apparent gaps in knowledge may be identified which may require further probing and knowledge elicitation (or may be genuine knowledge gaps!)
  - Diagrams are particularly useful for identifying gaps in sets of causal statements
-

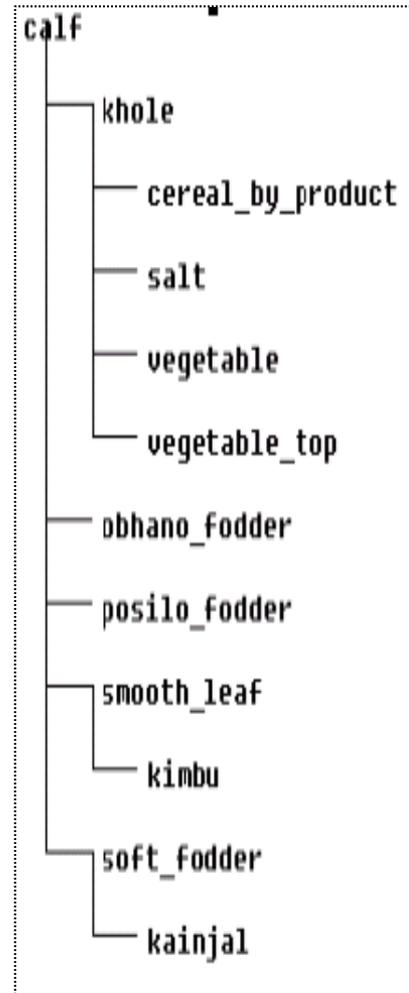
# Use tools to assess attribute statements

- Tool: objects and their attributes (found in - System Tools > Knowledge Analysis > Single Kb)

Object	ability	age	amount	area	location	Mat effect	productivity	Sponge effect	thickness
new_lay	low	<5years	high		on_wet_field		high		
new_lay			increase						
old_lay	high	>10years	increase	decrease		decrease	decrease	decrease	high
old_lay		increase	low	increase		high	low	increase	thick
old_lay						increase	no_change		

# Object hierarchies

- Key and powerful element of knowledge bases
- Use object hierarchies to specify the relationships between objects and to compact the knowledge base
- Take care to use correct syntax in the creation of object hierarchies. Syntax in object hierarchies should conform to the 'is a type of' relationship



**The 'calf' superobject should be changed to something like calf\_feeds**

---

## Welcome memo

- This is a vital component of the knowledge base. Having a clear idea about the purpose for knowledge base development can help reduce the amount of “irrelevant” information recorded in a knowledge base
  - Make sure you understand the purpose of your knowledge base and make sure the statements are addressing the stated knowledge areas
-

---

## Also update:

- Source label details

- As the knowledge base develops some elements may become more important to have as labels and the sources can be edited to reflect this.

- Derivations

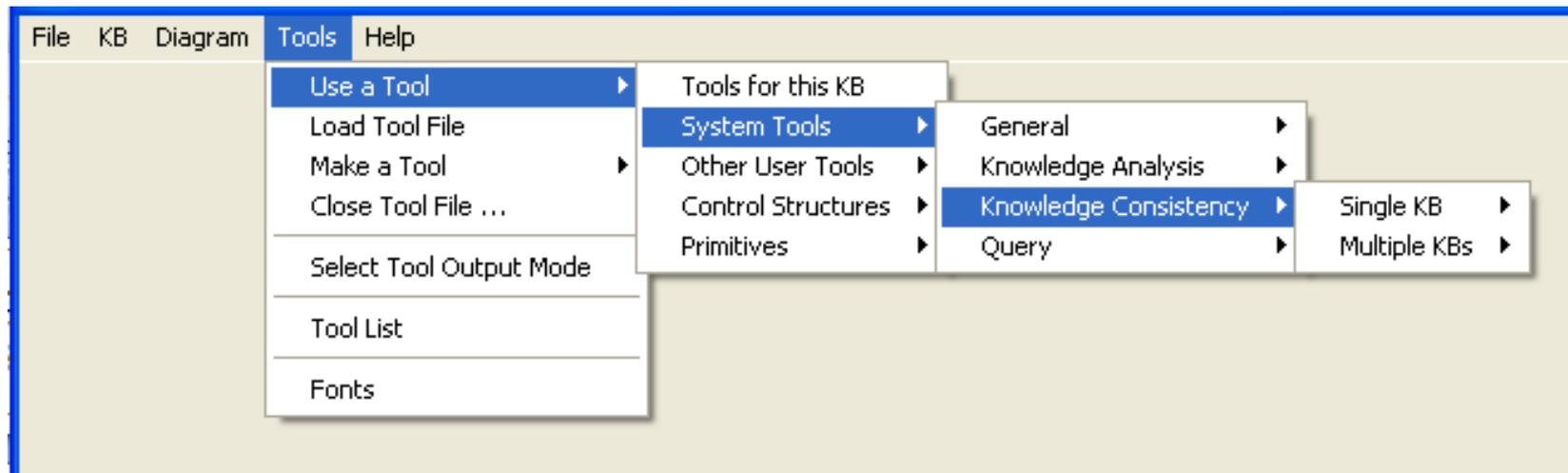
- These may change from 'unknown' to another value as acquisition progresses

- Appending sources

- Add additional sources to statements. What is common knowledge? Is there a sensible threshold?
-

# Knowledge Consistency Tools

Under System Tools there are Knowledge Consistency Tools that should be used throughout the development of a knowledge base



---

# Knowledge vs. Practice

- Essentially 'practice' is what users (farmers) do, and knowledge is why they do it (although there are often other considerations as well). Practice can be **difficult** to represent in knowledge bases and information about practice can usually be 'stored' in other more useful ways e.g. the use of cropping calendars, land use diagrams, ranking data etc.
- Knowledge associated with these practices can be easily stored in knowledge bases. The knowledge base can refer to the causes and results of such practices but should not include detailed statements about practice itself.

**It is a very common mistake to focus on practices rather than focusing on the reasons why**

---