



# Metataxis Ltd

Is it a bird? Is it a plane?  
How to develop a taxonomy

Judi Vernau

[judi@metataxis.com](mailto:judi@metataxis.com)



Europe's Largest XML  
& CMS Conference

# What is a taxonomy?

Within information architecture, taxonomy is a collection of terms that enable the classification of content in order to support:

- information organisation and management
- searching
- information discovery (browsing)
- web site usability
- content re-use

# Taxonomy includes...

## File plans

- Asset management
- Financial management
- ▶ Managing people
  - Equal opportunities
  - Recruitment
- ▶ Staff management
  - Appraisal policies
  - Disciplinary procedures
  - Terms and conditions
- Training

## Navigation hierarchies

amazon.co.uk Hello. Sign in to get [personalized](#) recommendations.

Your Amazon.co.uk Deal

Shop All Departments Search Amazon.co.uk

- Books >
- Music, DVD & Games >
- Electronics & Computing >
- Home & Garden
  - Kitchen & Dining
  - Appliances
  - Houseware & Furnishings
  - Garden & Outdoors
  - DIY & Tools
  - All Home & Garden
- Toys, Children & Baby >
- Jewellery & Watches >
- Shoes & Accessories >
- Sports & Leisure >
- Health & Beauty >

## Thesauri

### GAMING CLUBS

Use for Casinos  
Broader term Clubs  
Related term Night clubs

### GAMING MACHINES

Use for Fruit machines  
Broader term Games of chance  
Narrower term Video games  
Related term Amusement arcades

## Controlled vocabularies

### DOCUMENT TYPES

Briefing  
Correspondence  
Form  
Invoice  
Map  
Report

ie any set of terminology or 'business semantics'

a common language

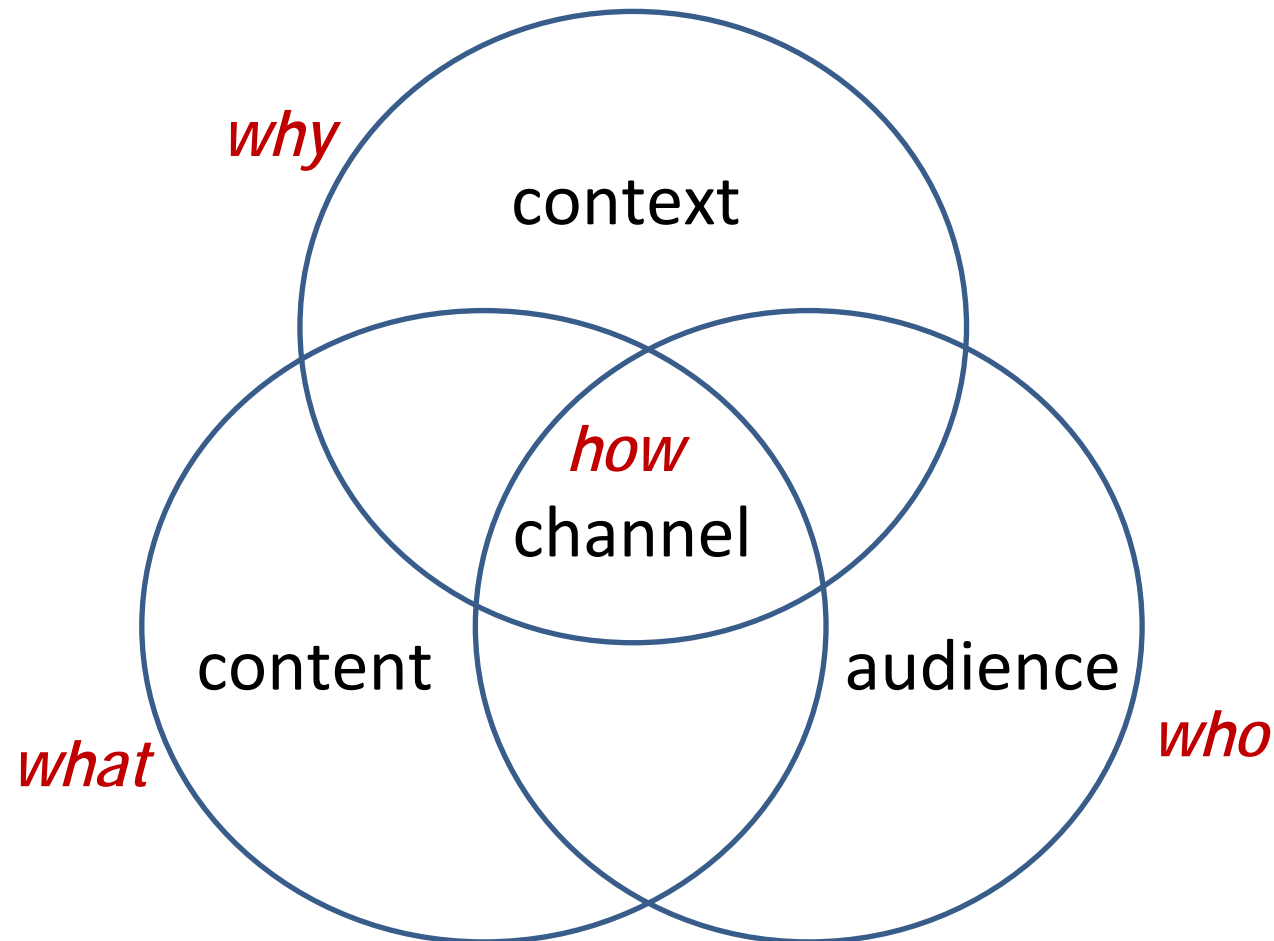
© Metataxis Ltd, 2008



# Taxonomy development : the six phases

- Project scoping
- High level taxonomy development
- Detailed taxonomy development
- Piloting
- Taxonomy revision
- Implementation

# Project scoping





## Project scoping – context (why)

- Confirm overall business goals with senior management;
- Identify stakeholders (senior management, subject experts, key users);
- Establish a recognised stakeholder group to elicit stakeholder needs.

# Project scoping – content (what)

- Gain a clear picture of the intended content:
  - Sources
  - Content types
  - Metrics
  - Workflows
- Take stock of existing vocabularies within the organisation (search logs, indexes, subject headings, etc);
- Identify relevant external vocabularies.

# Project scoping – audience (who)

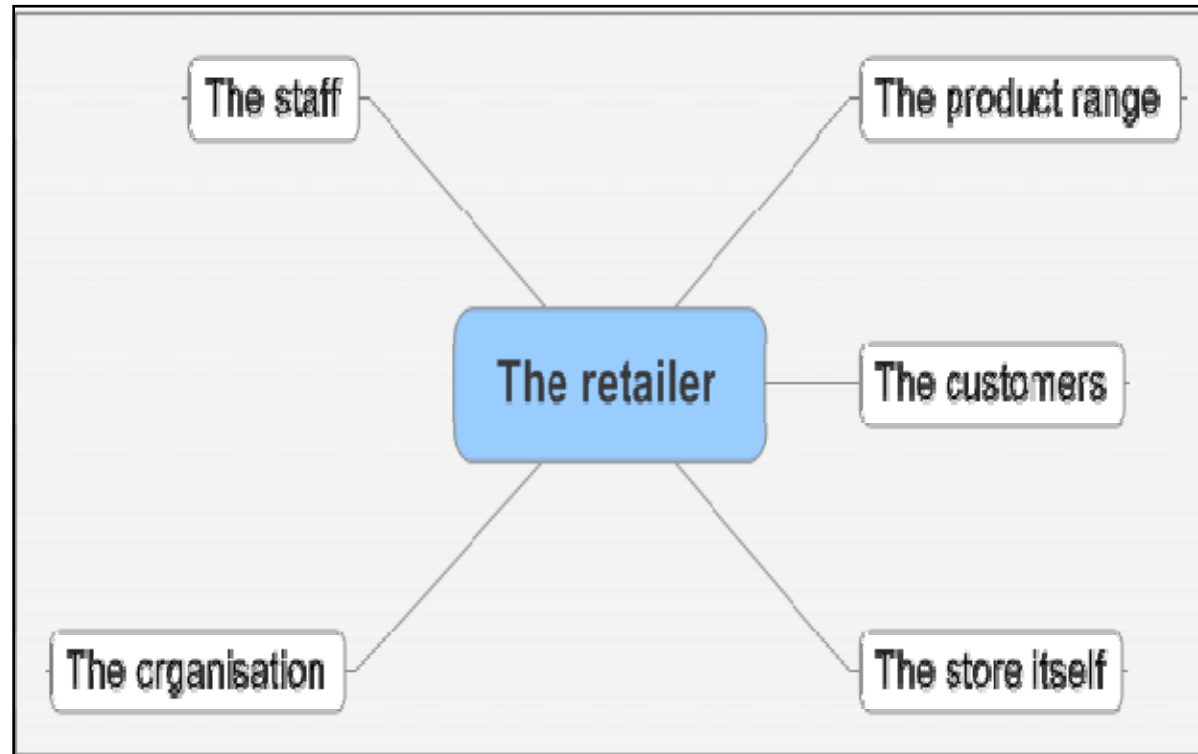
Taxonomy type	Potential audience
Navigational taxonomy	Internal (all staff, researchers) External (motorists, guitar enthusiasts, shoppers...)
File plan	Staff, business partners, members
Thesaurus	All staff, researchers Search engines
Controlled vocabularies	People creating metadata or filling in forms People interpreting forms Search engines



## Project scoping – channel (how)

- What's the delivery mechanism?
- Are there technical constraints? (eg SharePoint can't deal with hierarchical vocabularies)

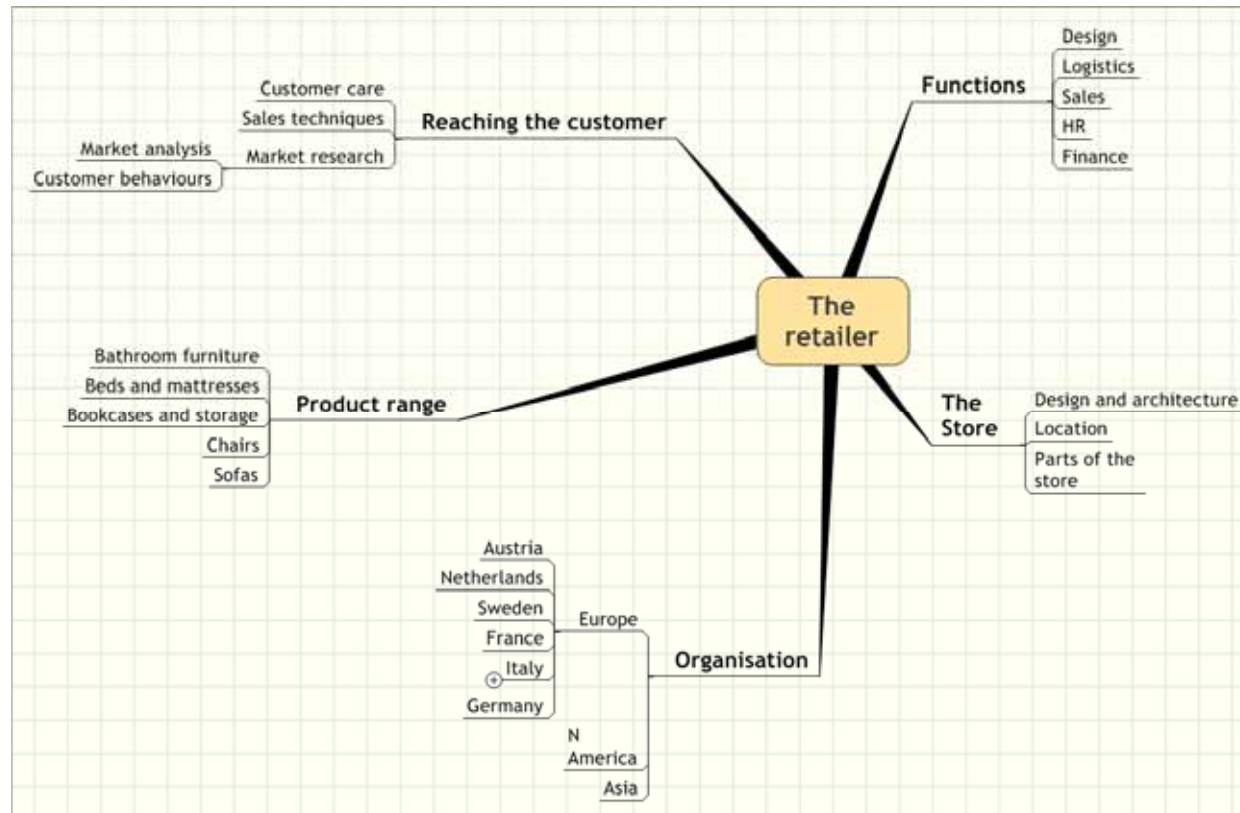
# High level taxonomy development



# High level taxonomy development

- Run a workshop (or number of workshops) for the relevant stakeholders to agree the high level conceptual groupings, by mindmapping terms and structures;
- If possible and appropriate, appoint a Subject Expert for each main subject area covered by the taxonomy.

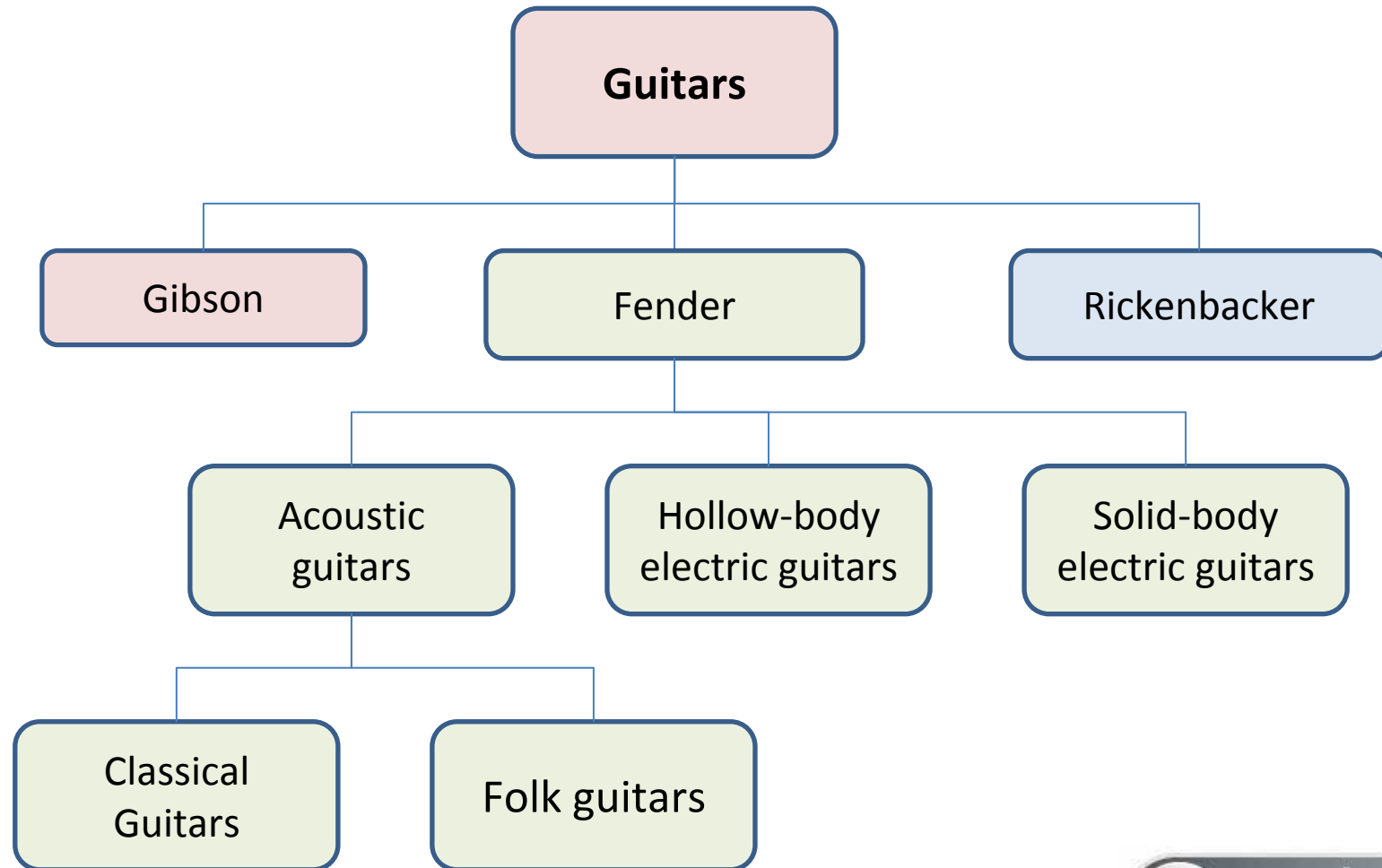
# High level taxonomy development



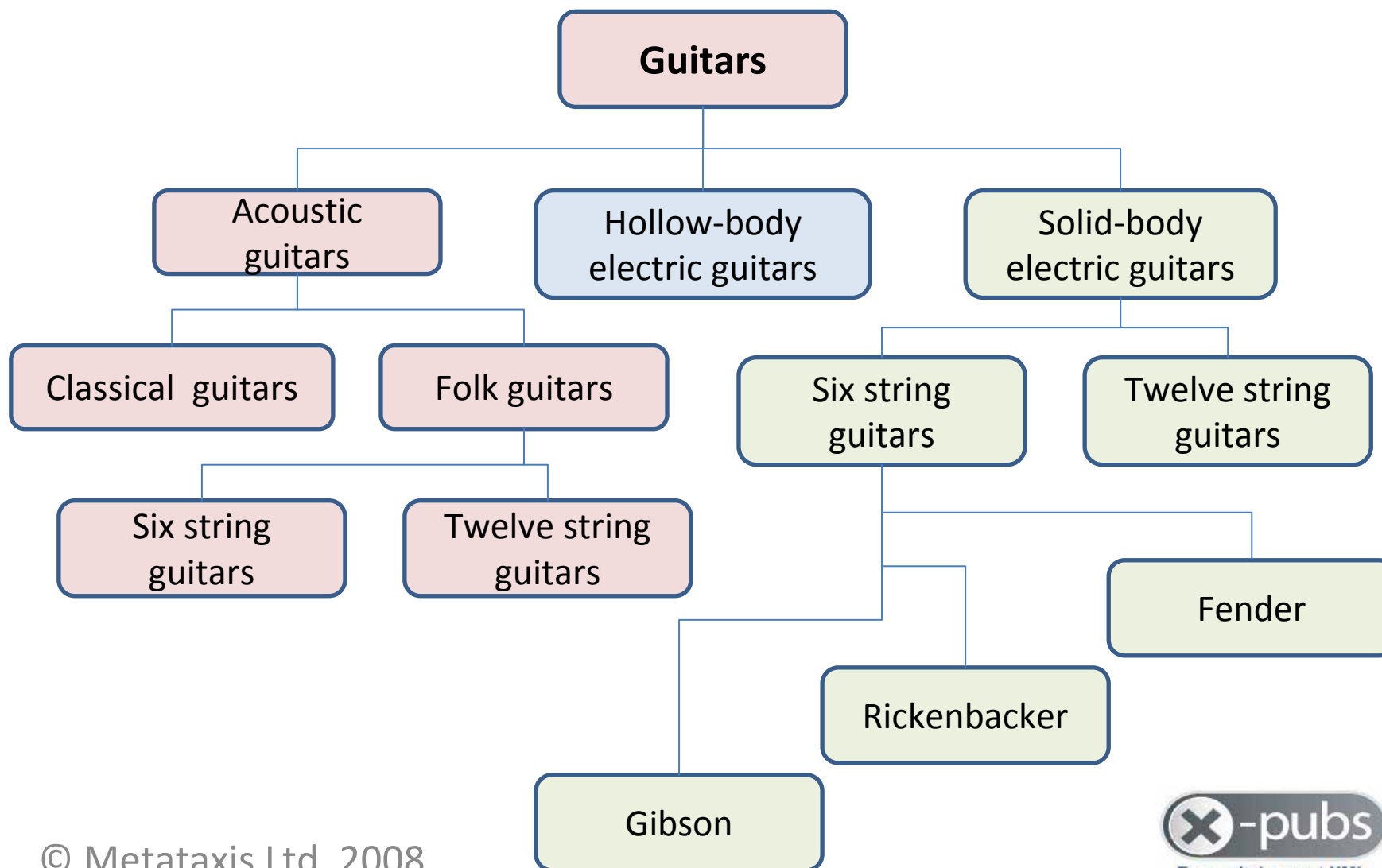
# Detailed taxonomy development

- Work with the subject experts to determine the appropriate structure and terminology for the taxonomy;
- Review all relevant vocabularies to identify useful terms;
- Identify specific taxonomy facets (eg people, places, products);
- Create the appropriate structure (eg broader / narrower terms; synonyms.)

# Taxonomy examples



# Taxonomy examples



Search term:

[ESC] clear [Enter] search

Term	Status	Flag	Relationships	Notes	Categories	Translation
Acoustic guitars	A		NT(2)	-	-	ENG
Classical guitars	A		BT(1)	-	-	ENG
Fender	A		BT(1)	-	MK	ENG
Folk guitars	A		BT(1), NT(2)	-	-	ENG
Gibson	A		BT(1)	-	MK	ENG
Guitar makers	A		NT(3)	-	-	ENG
Guitars	A		NT(1)	-	-	ENG
Hollow-body electric guitars	A		BT(1)	-	-	ENG
Rickenbacker	A		BT(1)	-	MK	ENG
Six-string guitars	A		BT(2)	-	-	ENG
Solid body electric guitars	A		NT(2)	-	-	ENG
Twelve-string guitars	A		BT(2)	-	-	ENG



## Guitar makers

- . Fender
- . Gibson
- . Rickenbacker

## Guitars

- . Acoustic guitars
  - .. Classical guitars
  - .. Folk guitars
  - ... Six-string guitars
  - ... Twelve-string guitars
- . Hollow-body electric guitars
- . Solid body electric guitars
  - .. Six-string guitars
    - ... Fender
    - ... Gibson
    - ... Rickenbacker
  - .. Twelve-string guitars

# Faceted classification

facets = mutually exclusive aspects within the domain\*

abstract entities (ideas)

living entities

artefacts

attributes (eg colour, size)

materials

parts or components

organisations

agents – people

agents – tools

recipients of action (patients, victims)

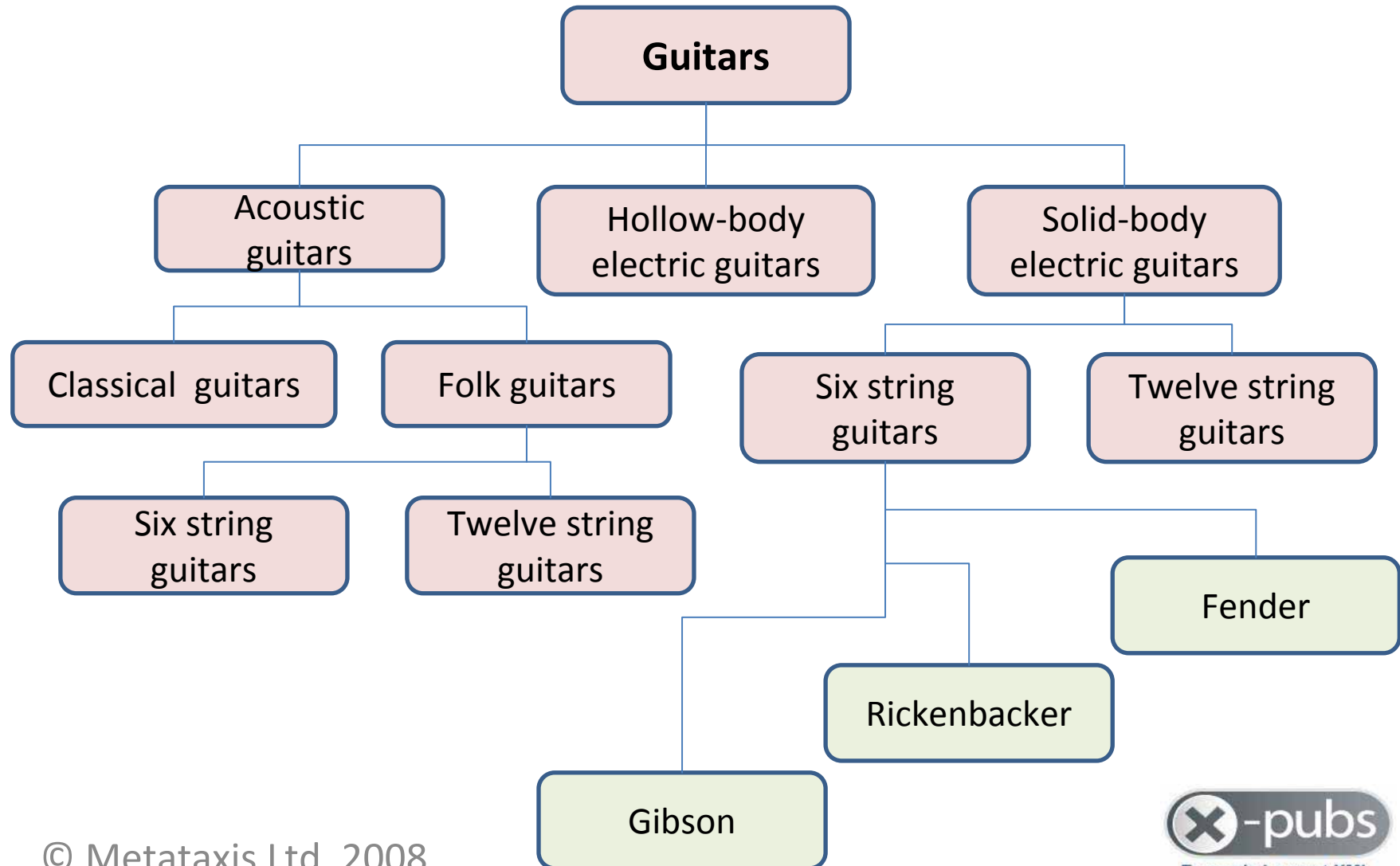
end products (eg furniture, guitars)

place

time (eg century)

processes

# Examples of facets



# Examples of facets

## Type

- Sparkling
- Liqueur wine
- Light wine
- Aromatized wine

## Colour

- Red
- White
- Blush/rose

## Grape type

- Shiraz/syrah
- Granache/Garnacha
- Tempranillo
- Chardonney
- Sauvignon blanc
- Pinot gris
- etc etc

## Regions

- Old world
  - France
    - Bordeaux
  - Italy
  - Spain
- New world

## Producers

- Lindermans
- Penfolds
- Fetzer
- etc etc

## Qualities

- Vin de table
- Vin de pays
- Kabinett
- etc etc

# Piloting

- Test the structure of the taxonomy by looking at the how balanced and consistent it is across the whole design;
- Carry out card sorting and other techniques to review how relevant, intuitive and unambiguous the taxonomy is for end users;
- This should also check whether you are using the appropriate terminology (do they say Personnel or Human Resources?).

# Piloting - card sorting

- Open / closed
- Focus on key topics
- Usually 25-50 cards, but could be more
- Participants are asked to sort them into piles that make sense, and then give each pile a name
- Are any obvious terms missing?
- Results: which cards appear in which categories? which cards appear together each time?

# Piloting – test classification

- Check how easy it is to assign taxonomy terms to the content, and whether the taxonomy covers the subject matter of the content adequately.

# Taxonomy revision

- Discuss user research findings with Subject Experts;
- Update the taxonomy accordingly (major revisions should be referred to the Stakeholder Group);
- Repeat piloting and revision as necessary.

# Implementation

- Work with the system developers to ensure that the taxonomy is presented in a user –friendly way, and that the system is supporting the taxonomy fully;
- Draw up guidelines and scope notes for users assigning terms to information objects;
- Draw up a methodology and rules for keeping the taxonomy up to date.



# Management and maintenance

- Needs ongoing management to ensure up-to-dateness;
- New terms will come on stream, new areas may need to be added (eg new product groups);
- Involve the experts and users in agreeing terms for inclusion and deprecation.

# Management and maintenance – software (1)

- supports basic relationships (eg broader term, synonym)
- supports more advanced relationships (eg ‘is producer of’)
- allows the creation and maintenance of metadata for each term
- supports faceted classification
- good browse functionality (tree view, alphabetical view, relationship view, metadata view)
- allows customisable subsets



# Management and maintenance – software (2)

- impact analysis
- good reporting facilities
- provides standardised XML inputs and outputs
- allows you to easily export into downstream applications (eg CMS)
- good technical support

# Metataxis and its associates have built taxonomies for:

Inter-IKEA Systems bv  
CBI  
BUPA  
UEFA  
The Law Society  
HBOS Treasury Services  
Macmillan Cancer Support  
RNIB  
British Computer Society  
Police Information Technology  
Organisation  
Foreign and Commonwealth Office

OECD  
Learning and Skills Council  
Scottish Natural Heritage  
Chartered Society for Physiotherapy  
UK Sports Council  
Becta (ACLearn)  
Commission for Racial Equality  
JISC  
Croner Publishing  
Philip Morris International  
NHS