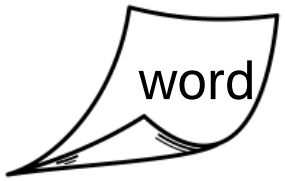
Best practices for building taxonomies

Ekaterina Mitova

ФМИ, Софтуерна документация
2022/2023

we see the world...



think of a word

any word

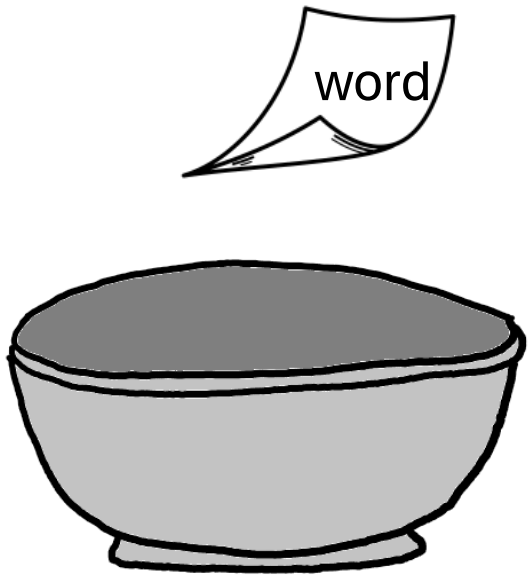
remember it

poll

what is your word?

A. adjective
B. adverb
C. name, noun or pronoun
D. verb
E. other

words in the mind

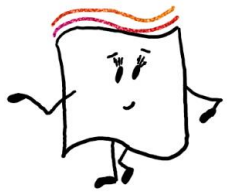


63%	68 nouns	
17%	18 abstract nouns	
6%	6 names	>85% objects
6%	6 interjections	
4%	4 adjectives	
3%	3 adverbs	
2%	2 unknown	<15% all else
1%	1 verbs	

imagine

to run

pink



beautiful

we, humans, see the world in objects

we...

...think in objects, exist in places of objects, shape lives and environments around objects

...recognize and organize objects by their properties

...know, imply, or invent rules for how objects relate to each other and to us

...perform actions on the objects based on the affordances of the objects

basic terms

metadata - definition

- a set of data that describes and gives information about other data

or

- data that describes other data, providing a structured reference that helps to sort and identify attributes of the information it describes

or

- data about data; it's information that's used to describe the data that's contained in something like a web page

or

- enriches the data with information that makes it easier to find, use and manage

metadata - examples

- wide variety of metadata depending on its purpose, format, quality and volume
- some of the widely used categories of metadata:
  - descriptive, structural, administrative and statistical
- example of basic metadata:
  - author, date created, date modified, and file size
  - everything written on a letter envelope to help the actual content – the letter – get delivered to its recipient
  - HTML tags instruct web browsers how to layout the pages to make it easier for humans to read them and follow references to other pages
- also used for unstructured data such as:
  - images, video, web pages, spreadsheets

metadata in a nutshell

- describing physical and digital objects is what metadata is about
- it helps the classification, access and storage of digital assets of all kinds
- it is with metadata that the encoding of knowledge within any data element is possible

metadata - types

- descriptive metadata

  adds information about who created a resource, and most importantly – what the resource is about, what it includes
- structural metadata

  includes additional data about the way data elements are organized – their relationships and the structure they exist in.
- administrative metadata

  provides information about the origin of resources, their type and access rights.

taxonomy - definition

- origin (from the ancient Greek)
  - taxis - arrangement or order
  - nomia – method
- the practice and science of categorization or classification
- the process of naming and classifying things such as animals and plants into groups within a larger system, according to their similarities and differences

thesaurus - definition

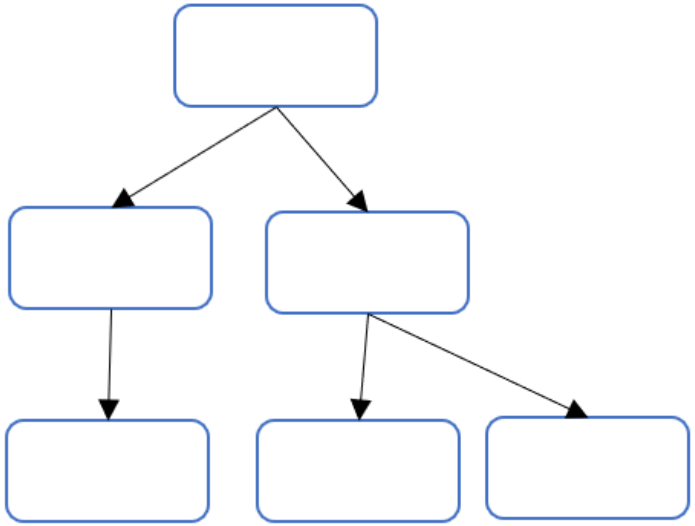
- origin (from antient Greek)
  - thēsauros - treasure, treasury, storehouse
- a book that lists words in groups of synonyms and related concepts
- a reference book in which words that have the same or similar meanings are grouped together

ontology - definition

- origin (from antient Greek)
  - on - being
  - logia – study
- the study of being alive and existing
- a set of concepts and categories in a subject area or domain that shows their properties and the relations between them

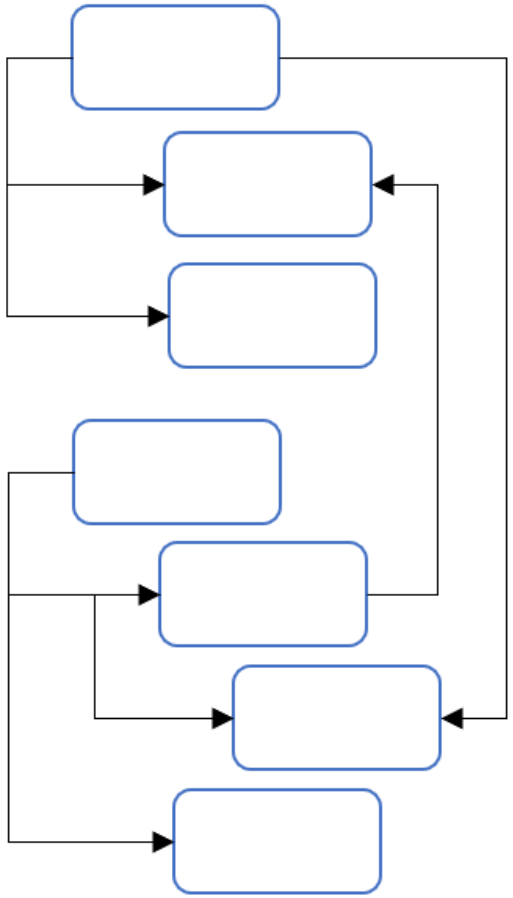
knowledge organization systems

taxonomy



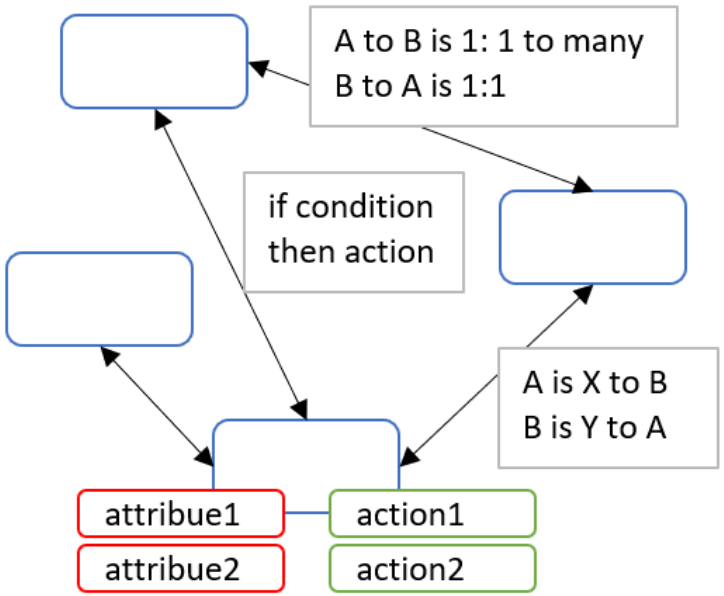
- structure of things and classes of things, some attributes, and their relationships
- always hierarchical – broader-narrower relationships
  - generic-specific: sports-soccer
  - class-named instance: human-Socrates
  - whole-part: theatre-stage
- one or few top concepts – start at top and traverse the whole domain
- just the necessary concepts; can grow and adapt within the hierarchy
- synonyms are optional

thesaurus

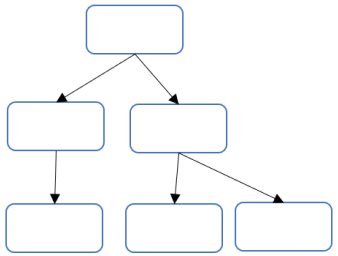
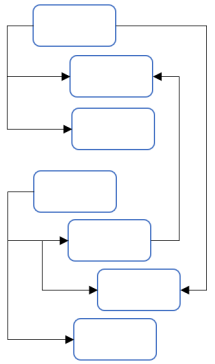
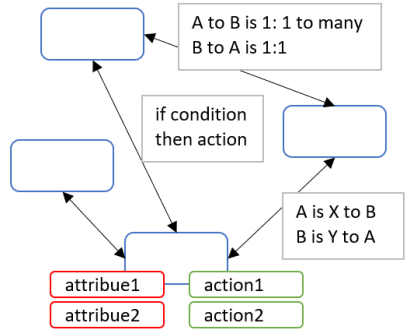


- structure of things and classes of things, some attributes, and their relationships
- hierarchical and associative relationships:
  - cause-effect: covid19-lockdown
  - field-practitioner: soccer-player
  - action-target: writing-book …
- many top concepts > hierarchies > many entry points, neither traverses whole domain
- domain covered from what is necessary to comprehensive; can grow in all directions as needed and cross domains
- complete synonym rings

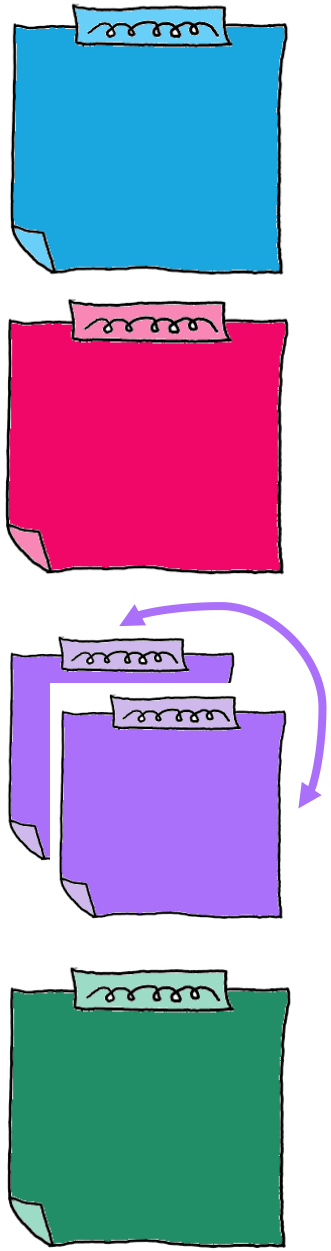
ontology



- structure of classes of things (rarely individual named instances), their attributes, and their relationships.
- semantic relationships, with descriptive labels of objects; hierarchical relationships possible - only generic-specific
    - example: employee - employer
- contains also axioms, rules, restrictions, events, and actions

	Taxonomy	Thesaurus	Ontology
			
Structure	things and classes of things, some attributes, and relationships		classes of things, attributes, and relationships
Relationships	always hierarchical; broader-narrower links	both hierarchical and associative	semantic
Hierarchy	one or few top concepts	many top concepts => many hierarchies	if applicable
Domain coverage	just the necessary concepts	from what is necessary to comprehensive	
Synonyms	optional	complete synonym rings	as needed
Other components	no additional components		axioms, rules, restrictions, events, and actions

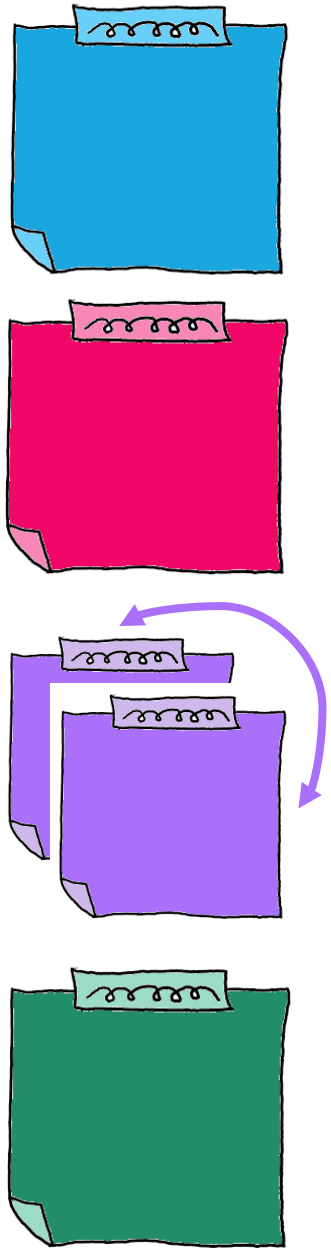
object-oriented approach (OOUX)



the object-oriented approach

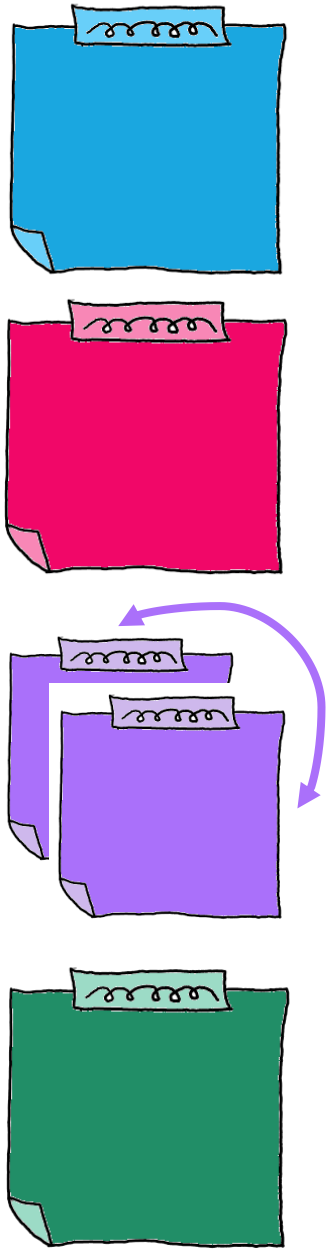
a method for analysis by deconstruction:

- break a domain into its parts (objects)
- discover the object traits (attributes)
- figure out connections (relationships)
- identify what can be done with the objects (actions)



applications

- build products
- structure content
- understand user tasks better
- clarify services
- create chatbot conversations
- model taxonomy and ontology
- …



example

UX Sofia - Meeting

File | Meeting | Scheduling Assi | Insert | Draw | Format Text | Review | Help | PDF-XChange | Tell me

Busy | 15 minutes

Mitova, Ekaterina : Automatic reply:" Hello, I am attending the UX Sofia conference and will have very limited acces...

You haven't sent this meeting invitation yet.
This appointment conflicts with another one on your calendar.

Send

Title: UX Sofia

Required: Mitova, Ekaterina

Optional:

Start time: Wed 18-Nov-20 | 10:00 | All day | Time zones

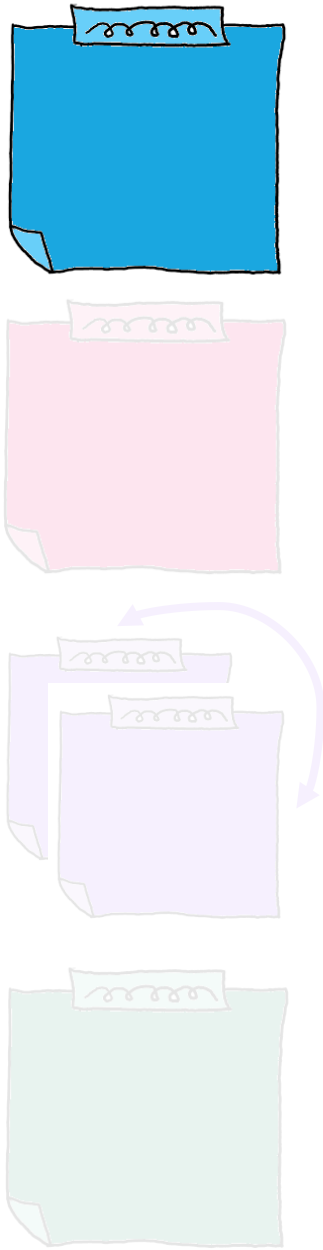
End time: Wed 18-Nov-20 | 10:25 | Make Recurring

Location: Microsoft Teams Meeting | Room Finder

plants.png
93 KB

Hello,¶
Let's·do·Quantum·Design·¶

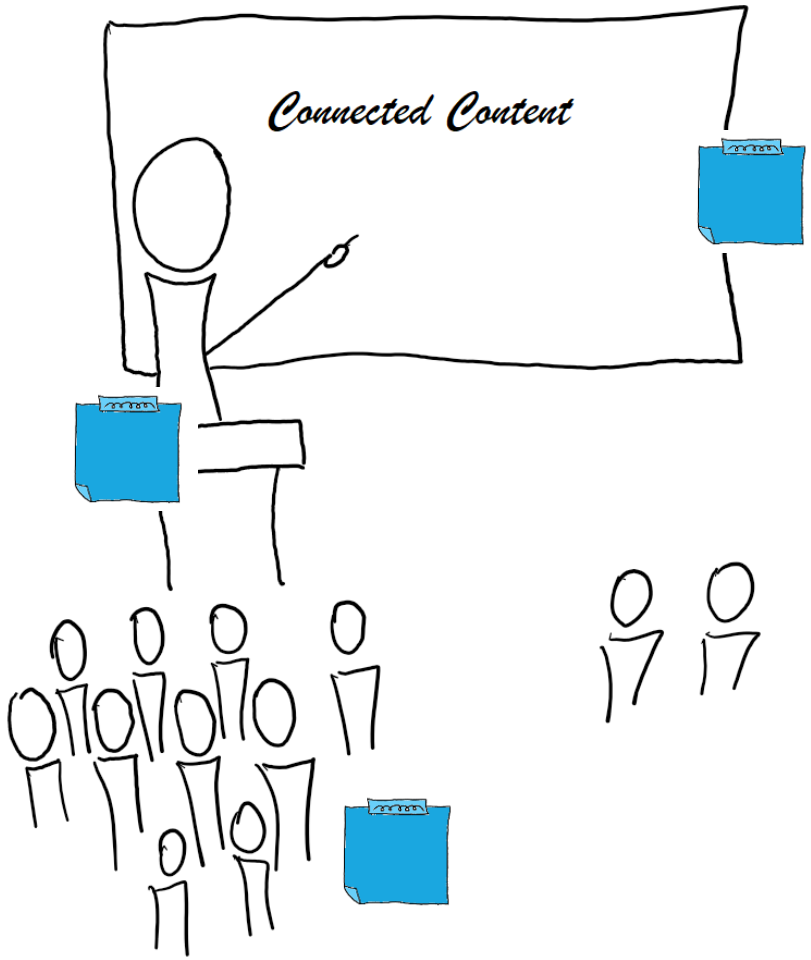
In Shared Folder | Calendar



objects

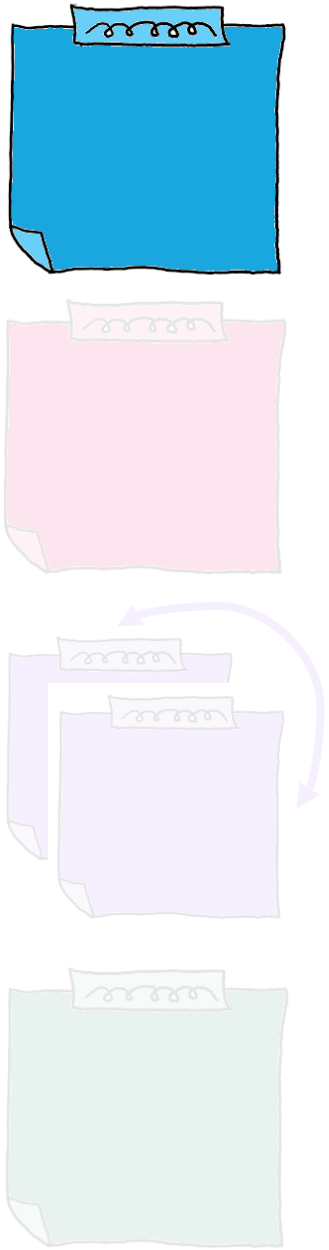
entities (real, virtual, abstract) that:

- have business purpose
- can be manipulated by users for achieving the business purpose
- have structure
- have many instances



analyzing a conference

1. think of a conference website
2. what would be the objects
   - **speaker**
   - ...



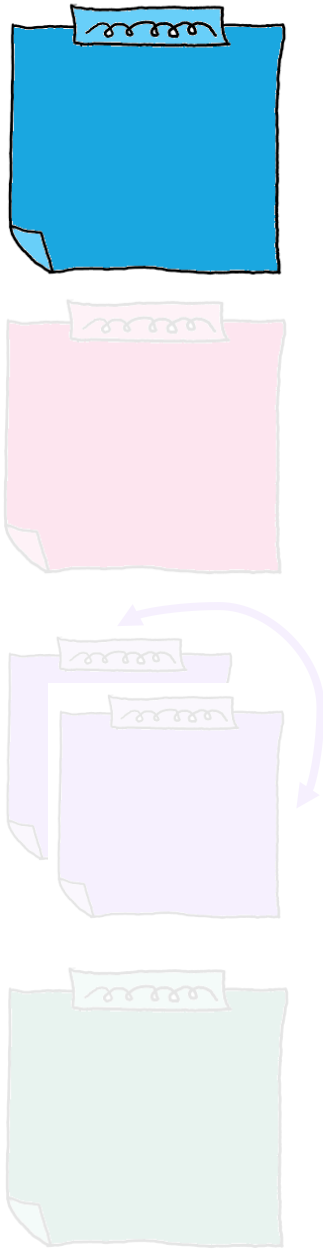
looking for objects

Let me see what workshops are there for Information Architecture.

This year, our line-up includes speakers from all over the world.

Which sessions does Jimmy have?

Noun > Object



objects: example

Speaker: a person presenting at an event

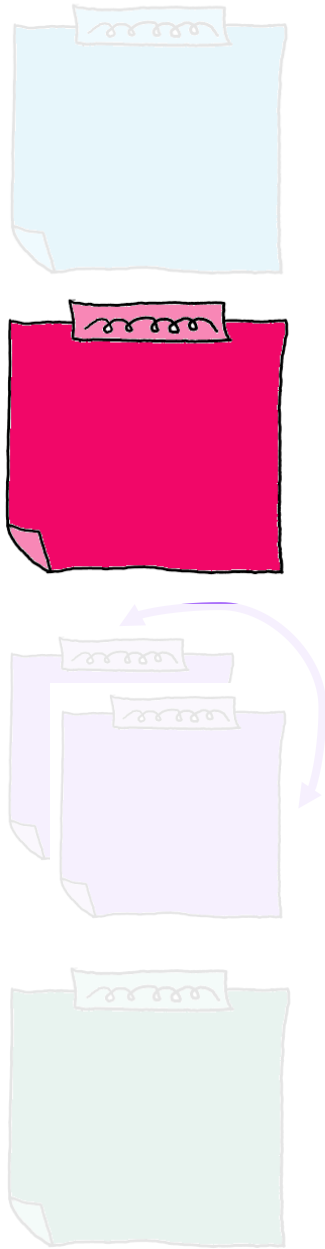
Attendee: every one joining the event

Session: an individual slot in the program - talks, lectures, workshops, discussions...

Room: a place within the venue - meeting room, bathroom, speaker lounge...

*Venue: the place where the event takes place - hotel, exhibition center

*Organizer: company, committee, or people preparing the event

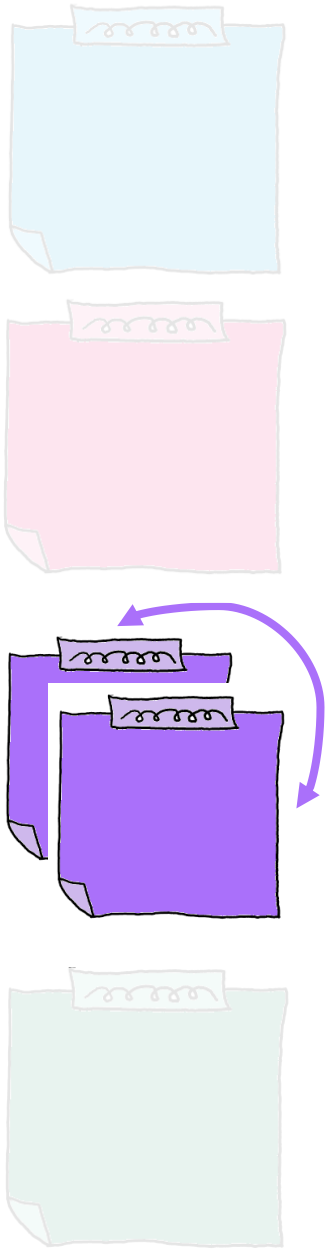


attributes

- a property, trait, or characteristic of an object that is part of the structure of the object
- attributes do not have a purpose on their own, outside the object
- we often use attributes to group, filter, search, and arrange things

attributes:
example

session	speaker	room	attendee
title	name	name	name
abstract	bio	track#	email address
topic	photo	location	profile
start time	company	capacity	program
end time	position		
duration	country		
capacity	Linkedin profile		
	Twitter handle		



relationships

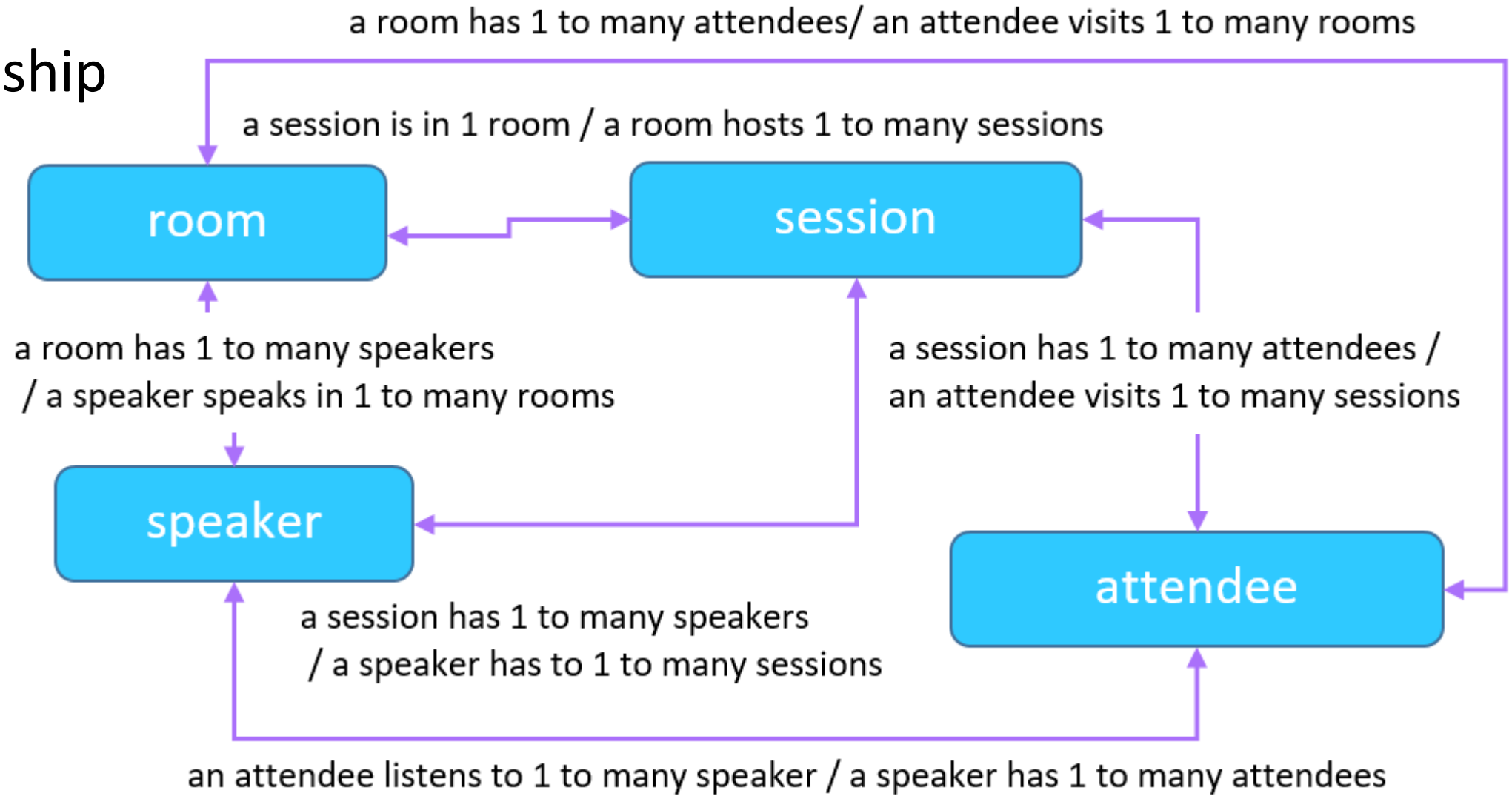
connections between objects

define how two objects relate to each other

relationships can have:

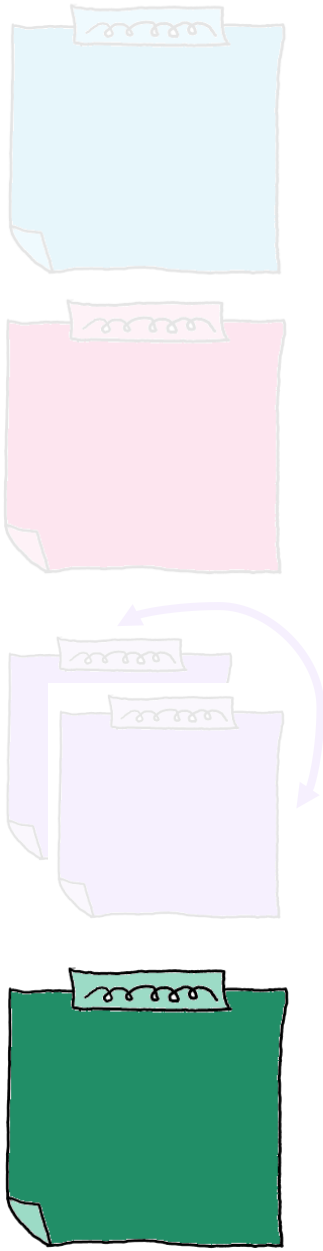
- direction: what is related to what
- cardinality: the number of relations that an object can take; differs per direction

relationships:
example
– relationship
diagram



relationships:
example

session	speaker	room	attendee
title	name	name	name
abstract	bio	track#	email address
topic	photo	location	profile
start time	company	capacity	program
end time	position	>session	>session
duration	country	>speaker	
capacity	Linkedin profile		
type	Twitter handle		
>speaker	>session		
>room	>room		
>attendee	>speaker		
>session			



actions

- things that can be done with or on an object
- expressed by verbs

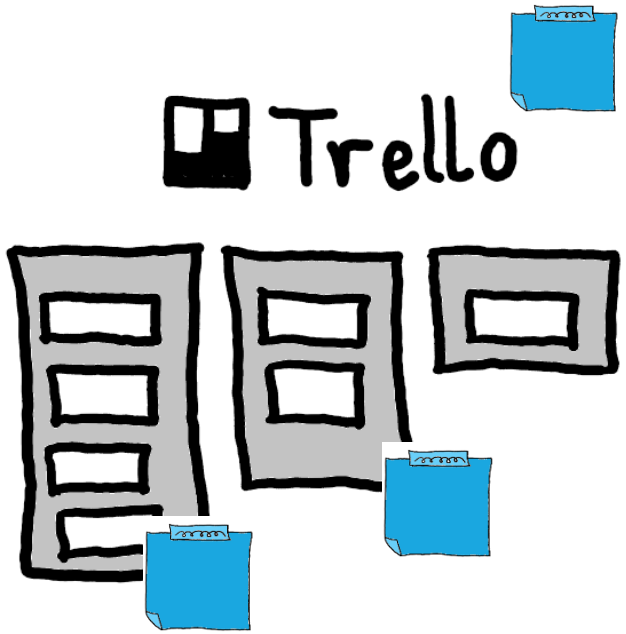
actions:
example

session	speaker	room	attendee
title	name	name	name
abstract	bio	track#	email address
topic	photo	location	profile
start time	company	capacity	program
end time	position	* type	>session
duration	country	>session	register
capacity	Linkedin profile	>speaker	pay
materials	Twitter handle	find room	unregister
type	>session	go to room	contact
>speaker	>room	see info	
>room	>speaker		
>attendee	see info		
>session	contact		
attend			
see info			
compare			
book			

your turn

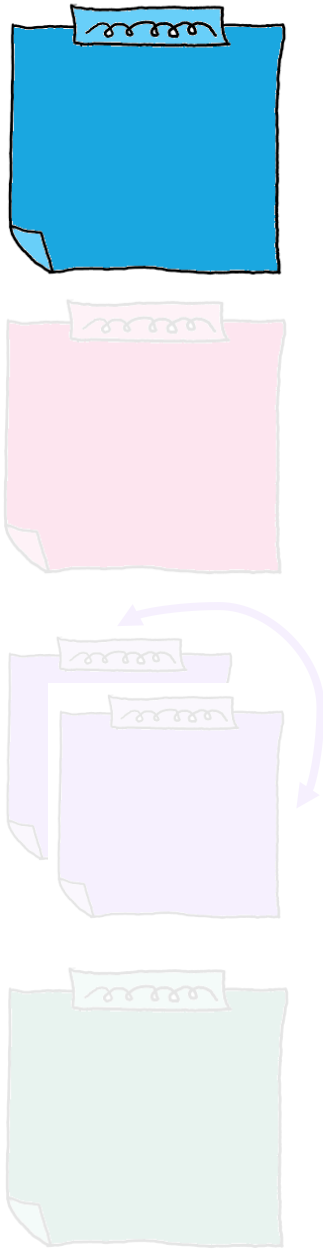
before we start

sign up and log in to trello.com



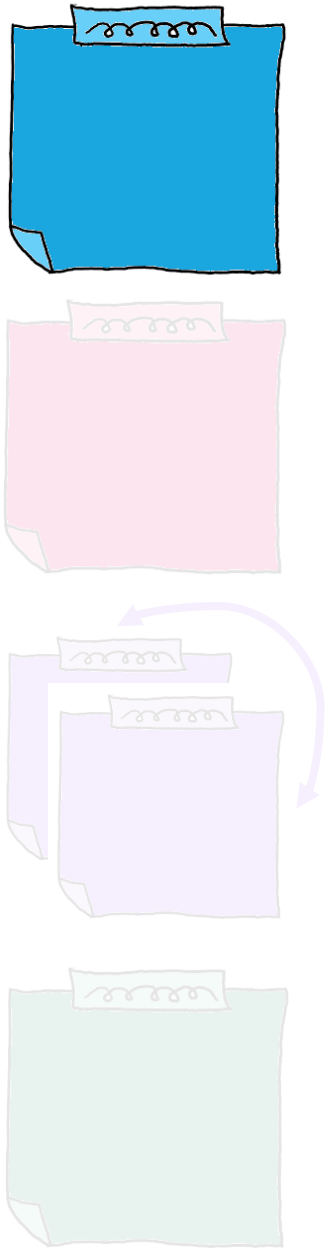
objects: exercise (15 min)

1. log in to trello.com
   you will build a taxonomy (and ontology)
   to organize content and to tag it
2. identify the objects of Trello
   - **workspace**
   - …



objects: tips

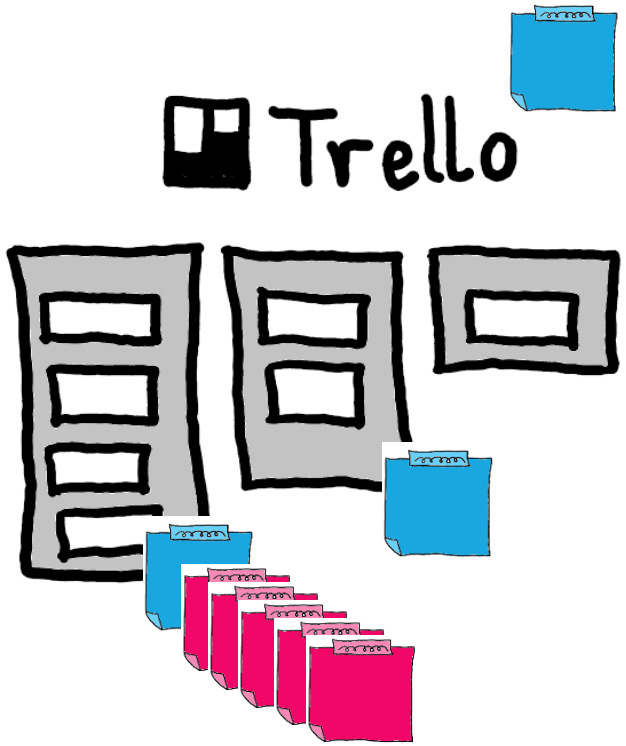
- gather all objects related to our product, our domain and our use case
- use documentation, marketing materials, user interviews, subject-matter experts, forums, blogs…
- write down a definition for each object
- look for business purpose and multiple instances



objects: example

Workspace: organizes your work and teammates
Board: your own workflow
Card: individual task
List: group of cards
Template: sample boards or cards
Member: people in your workspace/board/card
Rule: automated task
Power-up: service

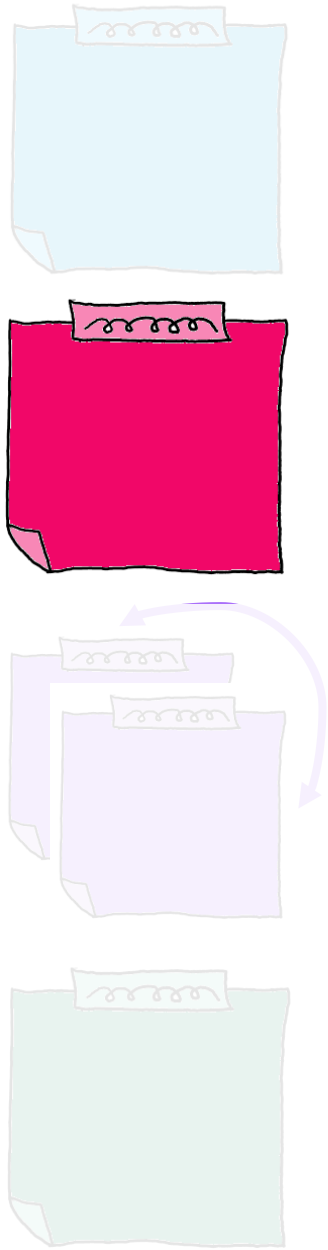
non-object: Automation



attributes: exercise (10 min)

1. think of a **Workspace** in Trello
2. what would be the attributes
    - name
    - description
    - ...

3. when ready, move on to **Board, Card, Power-up**

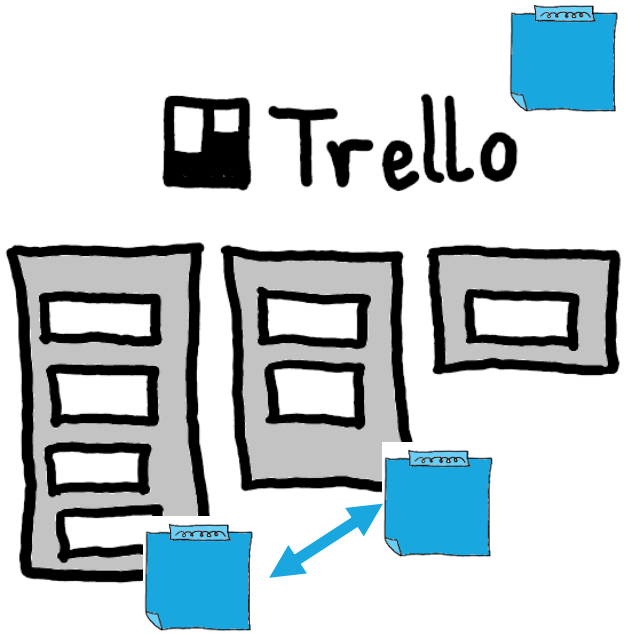


attributes: tips

- list all attributes we could think of
- focus on the attributes that are specific to this use case
- prioritize

attributes: example

Workspace	Board	Card	Power-up
name	view (timeline, table,...)	name	name
type	name	label	description
description	starred (y/n)	start date	featured (y/n)
public/private	description	due date and time	enabled (y/n)
highlights	background	reminder	provider
website			



relationships: exercise (10 min)

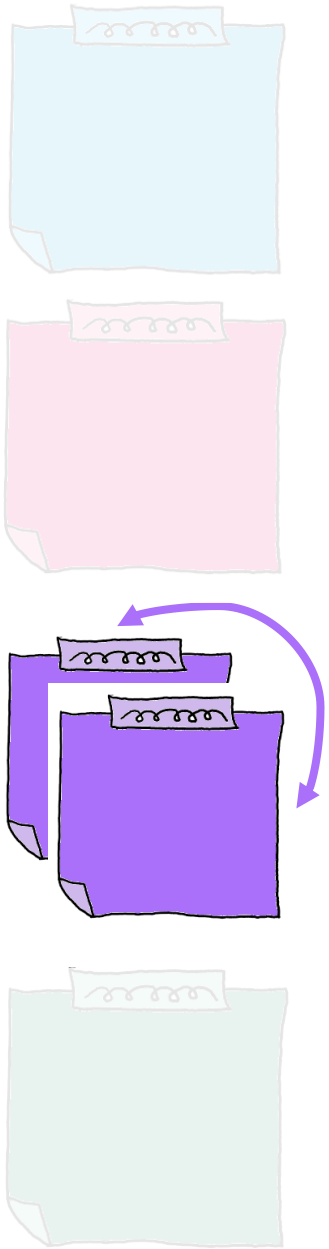
1. think of the **objects** in trello
2. what would be the relationships
   - a workspace has 0 to many boards
     a board belongs to 1 workspace
   - ...

workspace

board

card

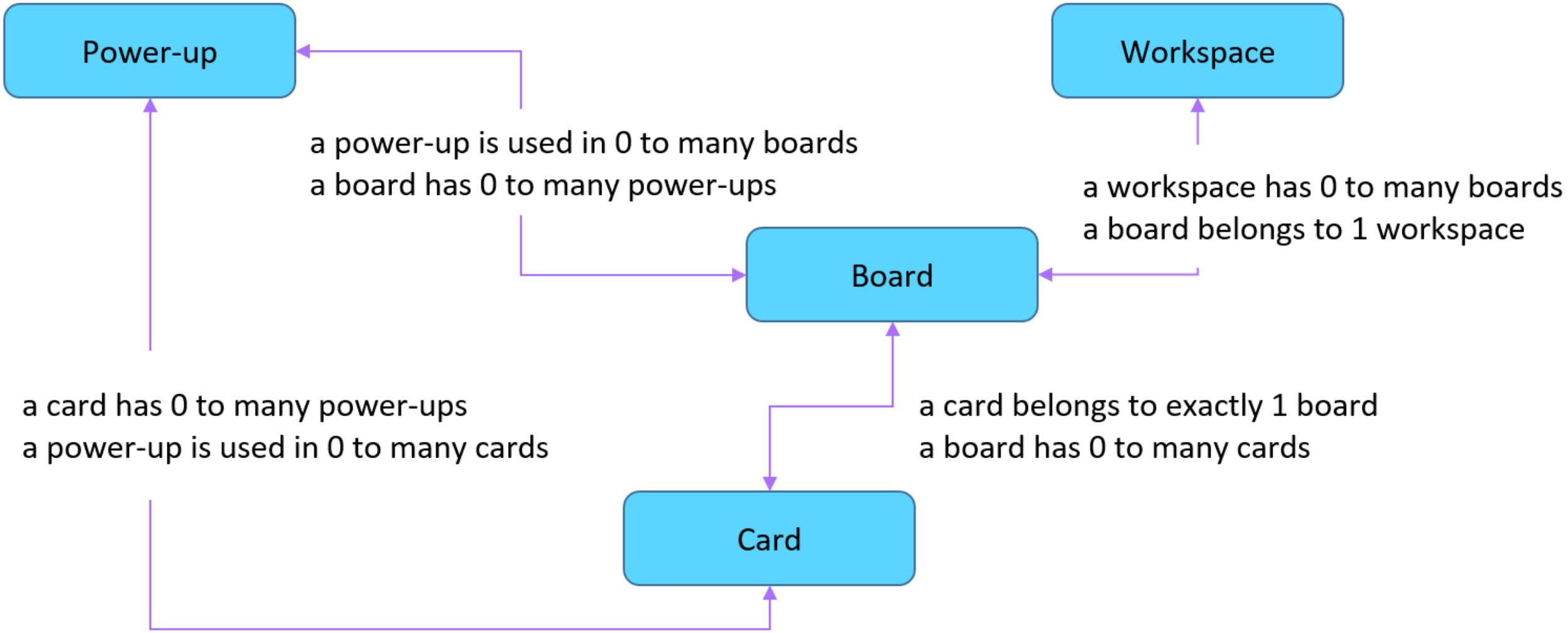
power-up



relationships: tips

- build a relationship diagram
- not all relationships make business or design sense
- pick up only the relationships that are relevant
- bi-directional: if object A is related to object B, then B is related to A as well.
- specify cardinality

relationship diagram



relationships: example

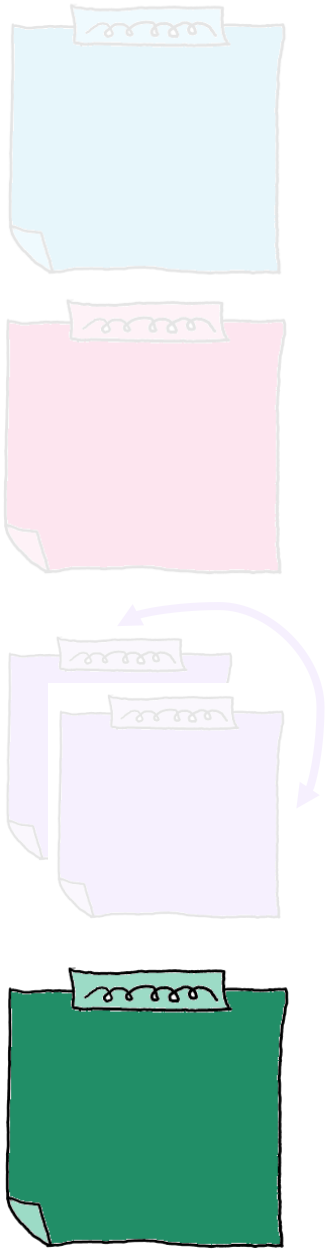
Workspace	Board	Card	Power-up
name	view (timeline, table,...)	name	name
type	name	label	description
description	starred (y/n)	start date	featured (y/n)
public/private	description	due date and time	enabled (y/n)
highlights	background	reminder	provider
website			
> Board (1:0-many)	> Card (1:0-many)	> Board (1:1)	> Card (1:0-many)
	> Workspace (1:1)	> Power-up (1:0-many)	> Board (1:0-many)
	> Power-up (1:0-many)		

relationships: hierarchy included

Workspace			Power-up
name	Board		name
type	view (timeline, table,...)	Card	description
description	name	name	featured (y/n)
public/private	starred (y/n)	label	enabled (y/n)
highlights	description	start date	provider
website	background	due date and time	
		reminder	
> Board (1:0-many)	> Card (1:0-many)	> Board (1:1)	> Card (1:0-many)
	> Workspace (1:1)	> Power-up (1:0-many)	> Board (1:0-many)
	> Power-up (1:0-many)		

actions: exercise (10 min)

1. think of the actions we can apply to a **Workspace**
2. what would be the actions
   - create
   - upgrade
   - ...
3. when ready, move on to **Board, Card, Power-up**



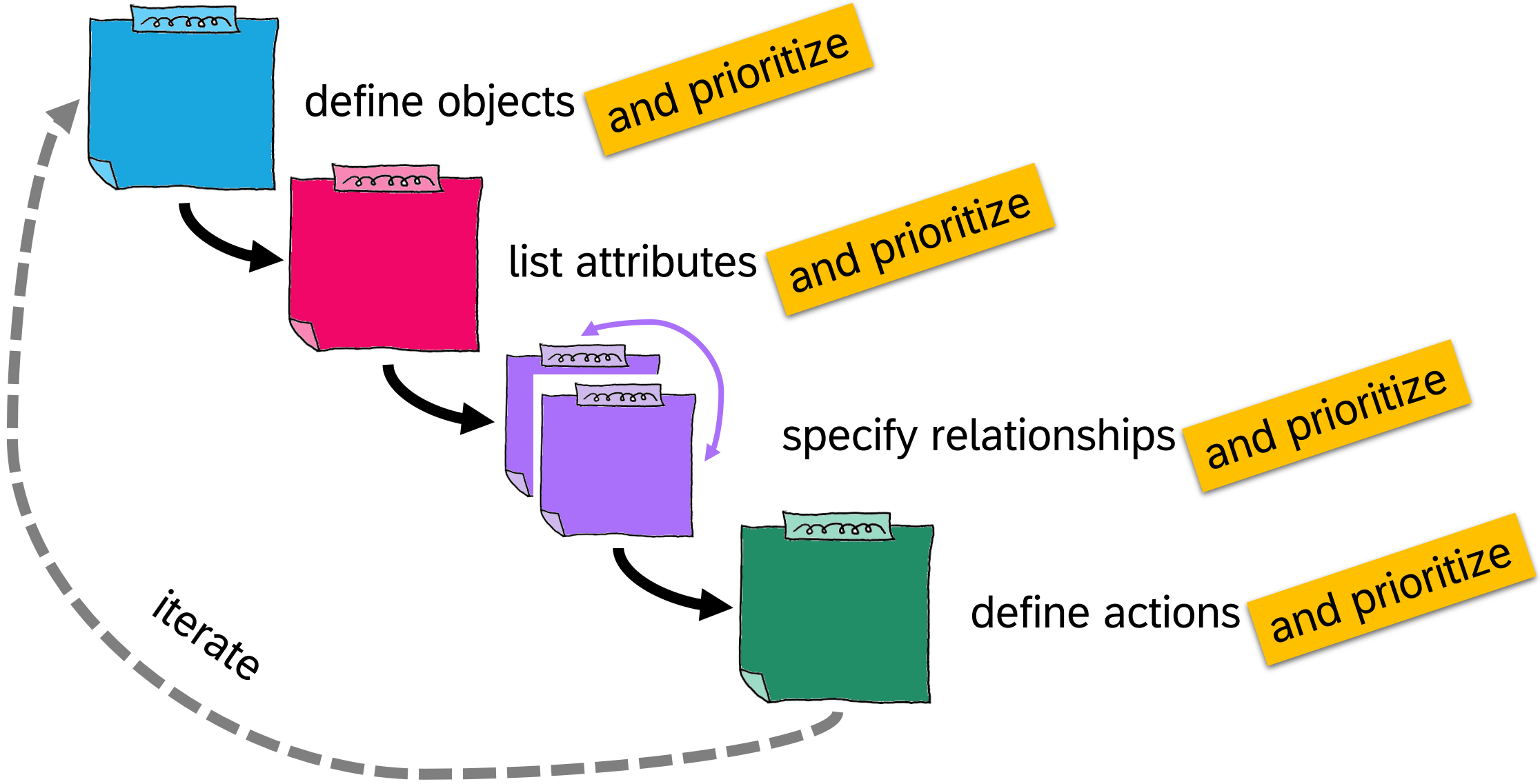
actions: tips

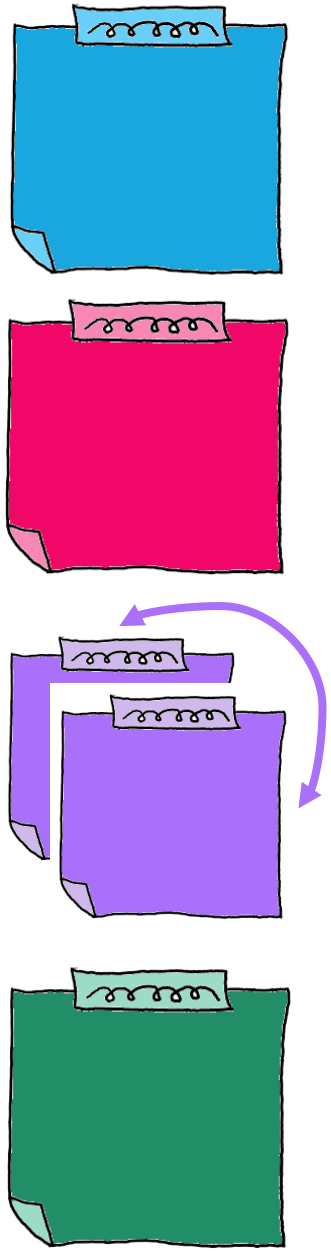
- gather all actions a user might apply to every object
- prioritize
- combine them to form user tasks
- can differ by user role

actions: example

Workspace	Board	Card	Power-up
name	view (timeline, table,...)	name	name
type	name	label	description
description	starred (y/n)	start date	featured (y/n)
public/private	description	due date and time	enabled (y/n)
highlights	background	reminder	provider
website			
> Board (1:0-many)	> Card (1:0-many)	> Board (1:1)	> Card (1:0-many)
	> Workspace (1:1)	> Power-up (1:0-many)	> Board (1:0-many)
	> Power-up (1:0-many)		
create	search	search	add
add boards	star	open card	authorize
invite members	change background	move	disable
upgrade	search cards	copy	
connect apps	add stickers	make template	
add automations (butle	invite members	watch	
		archive	

the OOUX process





exercise

- what changes will you make to the trello documentation based on the object model that we created?
- review https://trello.com/en/guide/trello-101

wrap-up

benefits of taxonomies

1. bring together different wordings to the same concept
2. help people search for information by different names
3. organize information in a logical structure
4. provide context and meaning for concepts for indexing and retrieval

purpose of taxonomies

- consistent tagging
- category browsing
- search (matching search strings to concepts)
- discovery (related concepts)
- filtering results
- sorting results
- automatic linking of relevant topics (recommendation systems)

take home

- use the object-oriented approach to analyze the domain
- gather your use cases and construct the taxonomy structure
- start by building a small taxonomy and then make it grow.
- adapt your taxonomy to the taxonomy repository you are using
- reuse as much as possible the taxonomy that is already there

resources

blog: Accidental Taxonomist

book: The Knowledge Graph Cookbook

OOUX: https://www.objectorientedux.com/