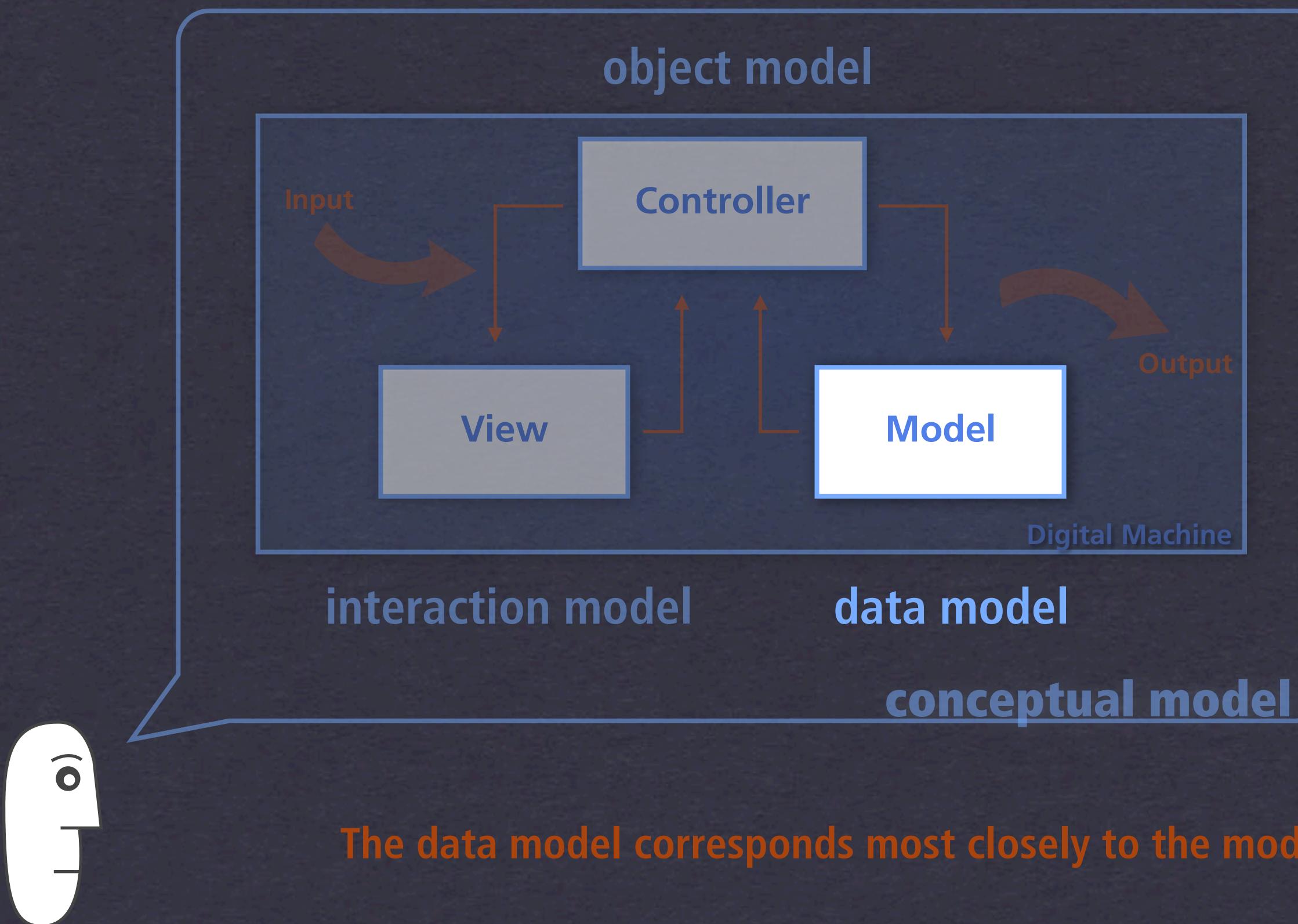


# Data Model

## Who lives here?

**Nº 6, *Design of Digital Machines***  
Tim Sheiner

# The four sub models of the machine



# What is stored where?

The data model defines storage.

- what information is stored?
- how is the store organized?

Storage, in turn, defines inheritance

- how do I get from here to there?
- what attributes will I find when I get there?

**Let's begin to understand the  
data model by starting with the  
data**

# An example: backpacking stoves

The screenshot shows the REI website's product listing for Backpacking Stoves. The page features a sidebar with filters for Categories (Canister Stoves, Liquid Fuel Stoves, Ultralight Stoves, Alternate Fuel Stoves), Features (Auto ignition, Made in USA, New Arrivals), Brand (BioLite, Esbit, Jetboil, MSR, Primus, Snow Peak, Soto, Trekmates), Fuel Type (Canister, Canister / liquid, Liquid, Liquid / tablet, Tablet, Wood), and Weight (LBS). The main content area displays four products: Jetboil Flash Cooking System, Jetboil Sol Titanium Stove, MSR WhisperLite Universal Backpacking Stove, and Jetboil Zip Stove. Each product is shown with its name, star rating, price, and a 'Compare' button. A large orange banner at the top of the page promotes 'THE REI WINTER CLEARANCE' with 'SAVE UP TO 50%'.

Backpacking Stoves

THE REI WINTER CLEARANCE  
**SAVE UP TO 50%**

Shop early for the best selection!

Relevance

Items per Page: 20 100 12 of 2

Compare items

Product	Rating	Price	Award
Jetboil Flash Cooking System	★★★★★ (147)	\$99.95	
Jetboil Sol Titanium Stove	★★★★★ (20)	\$149.95	
MSR WhisperLite Universal Backpacking Stove	★★★★★ (10)	\$139.95	Award Winner!
Jetboil Zip Stove	★★★★★ (29)	\$79.95	

# Specifically, the old skool fav:

The screenshot shows a Mac OS X desktop with a window titled "MSR WhisperLite Shaker Jet" open in a web browser. The URL in the address bar is [www.rei.com/product/708999/msr-whisperlite-shaker-jet-backpacking-stove](http://www.rei.com/product/708999/msr-whisperlite-shaker-jet-backpacking-stove). The page displays the product details for the MSR WhisperLite Shaker Jet Backpacking Stove, including a price of \$79.95, a 4.5-star rating from 93 reviews, and a "FREE SHIPPING" offer with a \$50 minimum purchase. The product image shows the stove with its fuel canister attached. Below the main image are two smaller thumbnail images: one showing the stove with a fuel canister and another showing a video thumbnail for the "MSR WhisperLite Shaker Jet Backpacking Stove Video". The video thumbnail includes icons for "Read reviews" and "View video". The page also features a "DESCRIPTION" section, "SPECS" tab, and "REVIEWS" tab. The REI logo is visible in the top left corner of the page content.

# Data is details



**MSR WhisperLite Shaker Jet Backpacking Stove**

**\$79.95**  
★★★★★ (93)  
Item # 708999

**REI Members** get back an estimated **\$ 8.00** on this item as part of their annual member refund.

1

**THE REI DIFFERENCE** | **100% SATISFACTION GUARANTEED** | **GEAR & ADVICE YOU CAN TRUST**

This item ships for free! [Learn more](#)

**MSR WhisperLite Shaker Jet Backpacking Stove Video**  
The ever-popular MSR WhisperLite stove features a self-cleaning shaker jet to reduce soot build-up and stove maintenance.

DESCRIPTION	SPECS	REVIEWS
<b>MSR WhisperLite Shaker Jet Backpacking Stove Specs</b>		
<b>Specification</b>	<b>Description</b>	
Best use	Backpacking	
Fuel type	Liquid	
Fuel	White gas	
Auto ignition	No	
Burn time (max flame)	(100g of fuel) 28.7 minutes	
Average boil time?	3 min. 58 sec.	
Water boiled per 100g fuel?	7.3 liters	
Weight	(Stove and pump only) 11 ounces	
Dimensions	6 x 4 x 4 inches	
Cold weather use?	Good	
Stove stability	Good	
Pot stability?	Good	
Ability to simmer	Good	

# Name:value pairs are a simple structure



item # 708999

# Name:value pairs are a simple structure



name      value  
↓            ↓  
item # 708999

# Name:value pairs are a simple structure



name      value  
↓            ↓  
**item #** 708999

**brand** MSR

**model** WhisperLite Shaker Jet

**price** \$79.95

**auto ignition** no

**weight** 11 oz

**average boil time** 238 sec

# Name:value pairs are a simple structure



object

name	value
item #	708999
brand	MSR
model	WhisperLite Shaker Jet
price	\$79.95
auto ignition	no
weight	11 oz
average boil time	238 sec

object properties or attributes

# A group of name:value pairs is a vector



variables	values
item #	708999
brand	MSR
model	WhisperLite Shaker Jet
price	\$79.95
auto ignition	no
weight	11 oz
average boil time	238 sec

vector

# Vectors can be written as an array



(item #, brand, model, price, auto ignition, weight, average boil time)

(708999, MSR, WhisperLite Shaker Jet, 79.95, no, 11, 238)

# Vectors can be written as an array



(item #, brand, model, price, auto ignition, weight, average boil time) ← variables

(708999, MSR, WhisperLite Shaker Jet, 79.95, no, 11, 238)

↑  
values

# Vectors can be written in matrix form

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	79.95	no	11	238

# A collection of vectors is matrix or a table

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

# More terminology

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

# More terminology

columns (variables)

rows (values)

record ID  
(must be unique)

a record  
(a vector)

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

# More terminology

columns (variables)

rows (values)

record ID (must be unique)

categorical

a record (a vector)

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

# More terminology

columns (variables)

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

rows (values)

a record (a vector)

record ID (must be unique)

categorical

continuous

# More terminology

columns (variables)

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

rows (values)

a record (a vector)

record ID (must be unique)

categorical

boolean (special kind of category)

continuous

# More terminology

columns (variables)

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	80	no	11	238
722001		MSR	XGK EX	150	no	13	296
813622		Jetboil	Sol	120	yes	11	300

rows (values)

a record (a vector)

record ID (must be unique)

categorical

continuous

boolean (special kind of category)

time (special kind of continuous)

# Statistics are not data

item #	image	brand	model	price (\$)	auto ignition	weight (oz)	average boil time (sec)
708999		MSR	WhisperLite Shaker Jet	79.95	no	11	238
722001		MSR	XGK EX	149.95	no	13.2	296
813622		Jetboil	Sol	119.95	yes	10.5	300
Averages				116.62		11.6	278

data → 278

statistic → 11.6

When used as a presentation tool, tables often mix data and statistics.

Data are observations.

Statistics are methods of aggregating, compressing and transforming data.

Now we understand data.  
What about the data model?

# Data are the observations

image	brand	price (\$)	weight (oz)
	MSR	80	11
	MSR	150	13
	Jetboil	120	11
	MSR	40	3

**EGK EX**



\$149.95  
13.2 oz

**Pocket Rocket**



\$39.95  
3 oz

**WhisperLite Shaker Jet**



\$79.95  
11 oz

**Sol**



\$119.95  
10.5 oz

# The data model is the organization

data model

Activity

Equipment Type

Equipment Subtype

Brand

Model

*metadata*

Camping & Hiking

Backpacking Stoves

Liquid Fuel

MSR

EGK EX



\$149.95  
13.2 oz

Canister

MSR

Pocket Rocket



\$39.95  
3 oz

WhisperLite Shaker Jet



\$79.95  
11 oz

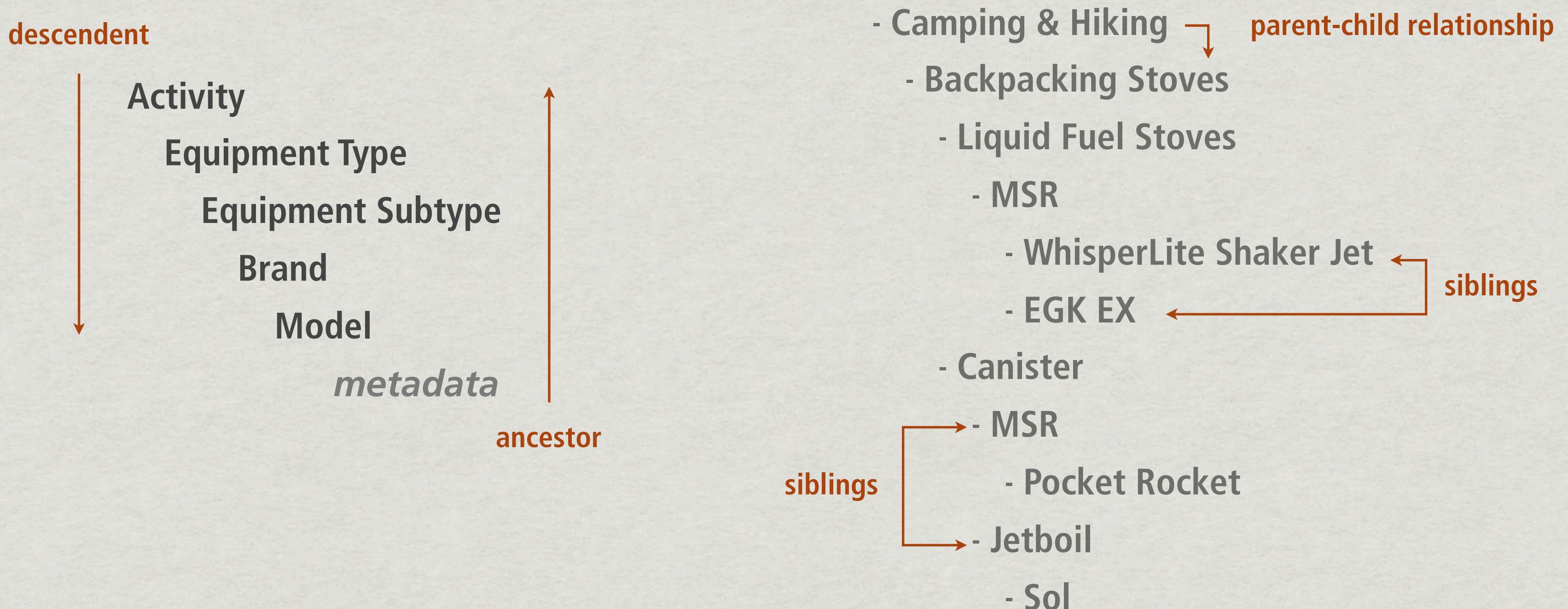
Jetboil

Sol

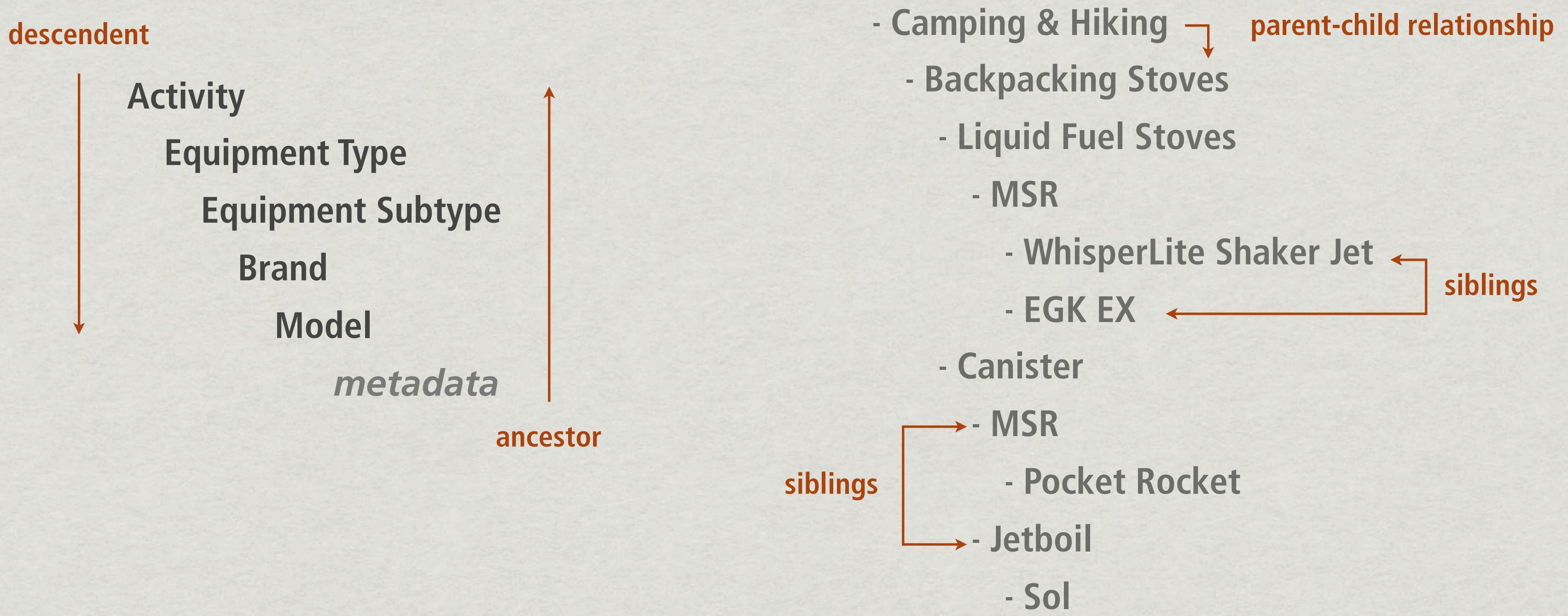


\$119.95  
10.5 oz

# The data model is a point of view about inheritance



# The data model is a point of view about inheritance



# The data model is an general inheritance solution for all system objects

**model**

instance

Camping & Hiking>Backpacking Stoves>Liquid Fuel Stoves>MSR>WhisperLite Shaker Jet



**Activity**

**Equipment Type**

**Equipment Subtype**

**Brand**

**Model**

instance

Kayaking & Canoeing>Canoes>Recreational Canoes>Mad River>Adventure 16



instance

Climbing>Chalk>Chalk Bags>Prana>Eagle Eye



# The data model is an general inheritance solution for all system objects

model

instance

Camping & Hiking>Backpacking Stoves>Liquid Fuel Stoves>MSR>WhisperLite Shaker Jet



Activity

Equipment Type

Equipment Subtype

Brand

Model

instance

Kayaking & Canoeing>Canoes>Recreational Canoes>Mad River>Adventure 16



instance

Climbing>Chalk>Chalk Bags>Prana>Eagle Eye



a solution, not *the* solution

# Communicating the Data Model

## Several standard forms

- hierarchy
- set diagram
- illustration

# Data Model Example

## data model

Activity

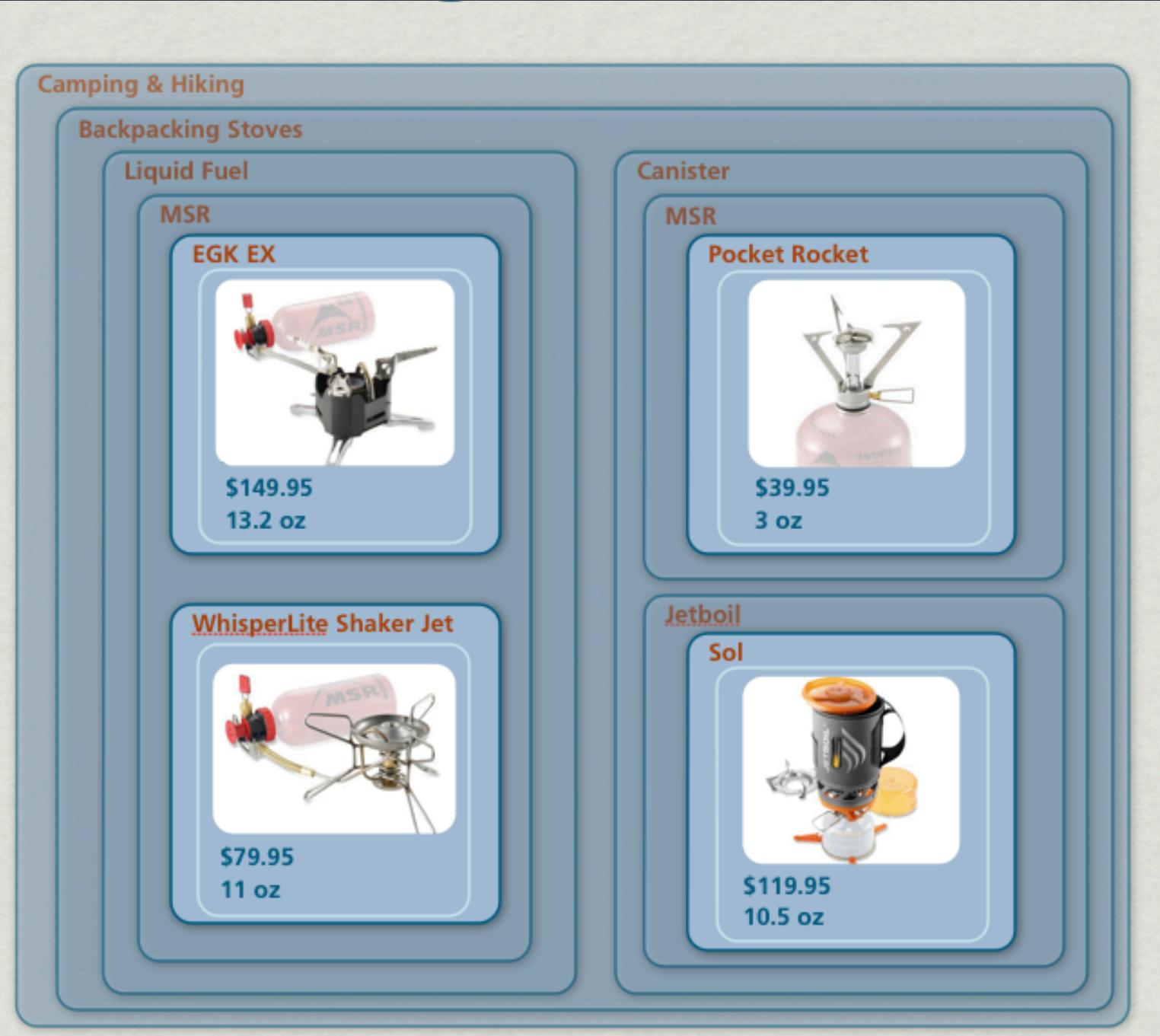
Equipment Type

Equipment Subtype

Brand

Model

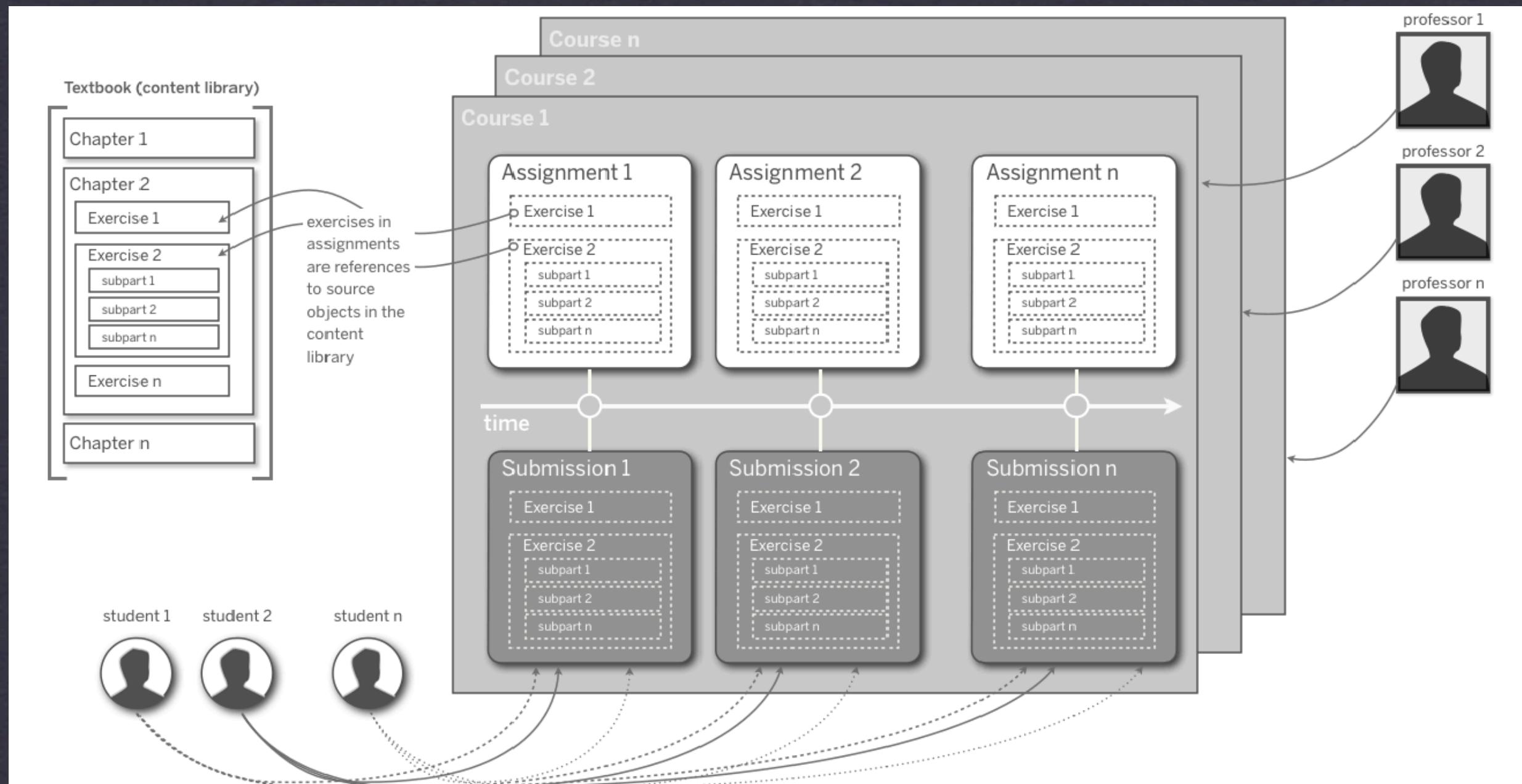
*metadata*



## Hierarchy & Sets

This baroque example includes a simple hierarchy illustrating the general case, and then a modified set display to demonstrate a specific case.

# Data Model Example



## Illustration

An illustration can be helpful to show relationships between sibling nodes in the structure.

# How do you design a data model?

# Data modeling challenges

- **Controlled Vocabulary**

a restricted list of names for the objects in the system

- **Taxonomy**

a hierarchical structure that defines the inheritance relationships in the system

- **Metadata**

the object attributes used to distinguish taxonomic siblings

- **Information Architecture**

the design of the vocabulary, taxonomy and metadata

# Controlled vocabulary captures meaning

The Whorfian hypothesis is about the question of whether words influence thought.

Data modeling is agnostic with respect to this question, but deterministic on the point that words influence communication.

The controlled vocabulary should capture the terms used by system experts to communicate precisely.



Edward Sapir/Benjamin Whorf



Different words for different states of rice

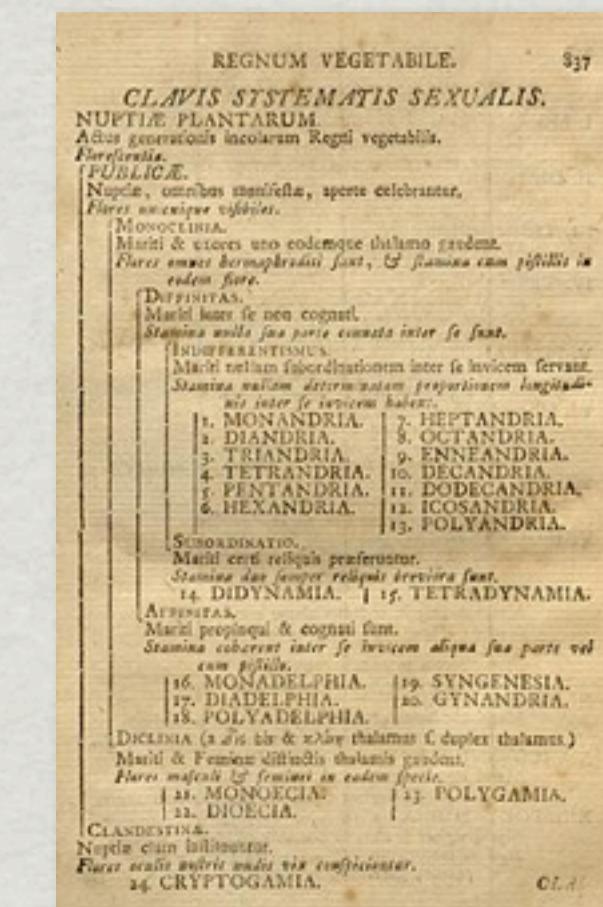
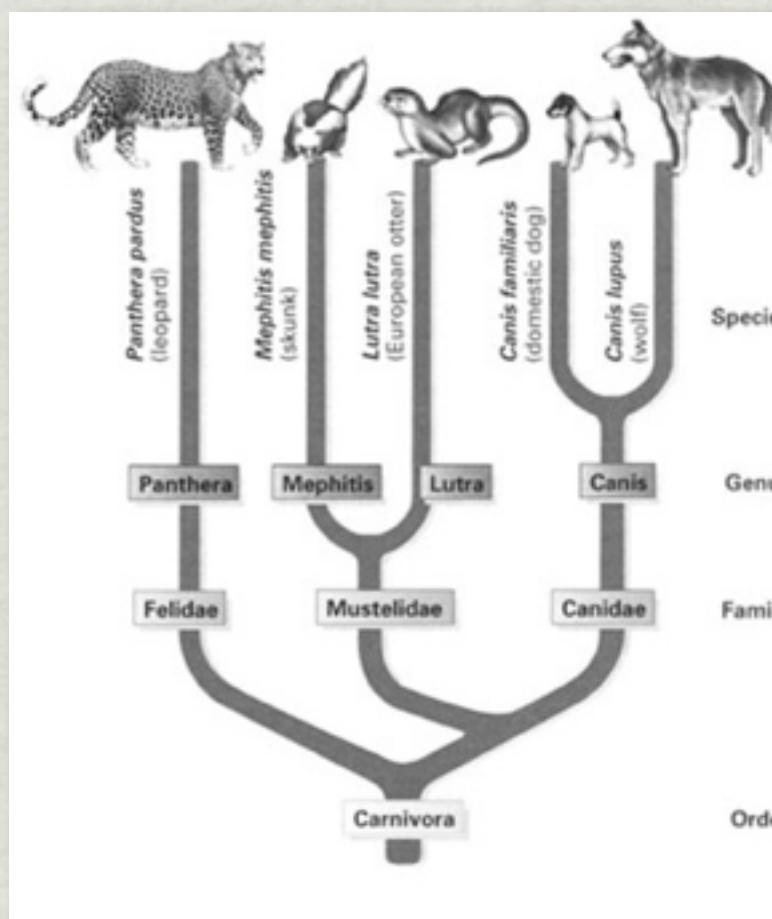


# Taxonomy is arbitrary

Originally, taxonomy described a method for classifying the natural world.

In the context of a data model, it is the rule set by which objects are placed into a hierarchical structure based upon a controlled vocabulary.

Regardless of how carefully and objectively these rules are constructed, never doubt that they are by definition arbitrary.



Carl Linnaeus

# Metadata separates siblings

## data model

Activity

Equipment Type

Equipment Subtype

Brand

Model

*metadata*

## instance

- Camping & Hiking
  - Backpacking Stoves
  - Liquid Fuel Stoves

- MSR

- WhisperLite Shaker Jet

- *item #*: 708999
- *price*: \$79.95
- *weight*: 11oz

- EGK EX

- *item #*: 702201
- *price*: \$149.95
- *weight*: 13oz



# Information architecture is the art

Information architecture is the design of the data model

## Data Storage

Activity

Equipment Type

Equipment Subtype

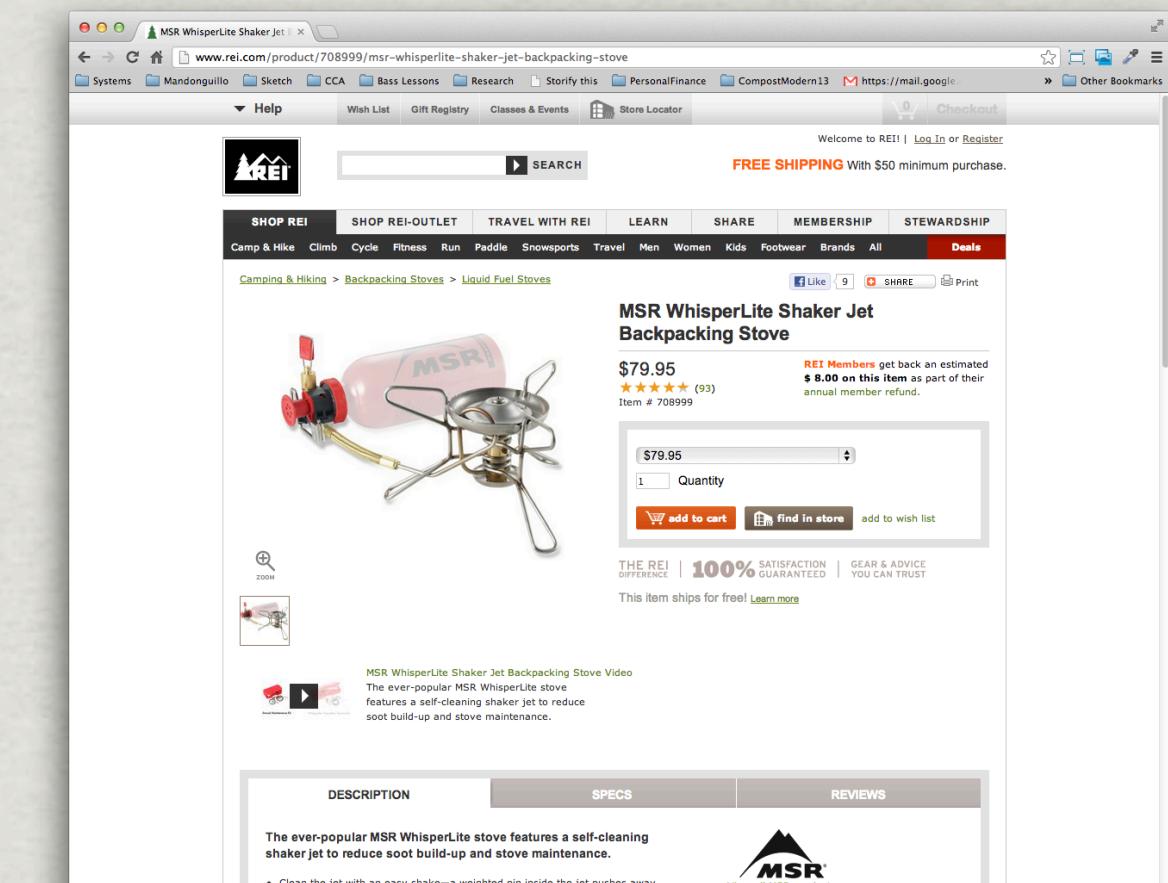
Brand

Model

## System Navigation

- Camping & Hiking
  - Backpacking Stoves
    - Liquid Fuel Stoves
      - MSR
    - WhisperLite Shaker Jet

## Information Display



*Information architecture's concern is a data model that is functional across cases*

**fin**