



INFORMATION ARCHITECTURE AND DATA MODELLING CYCLE

Bernard Suen
Center for Entrepreneurship
Chinese University of Hong Kong



Center for
Entrepreneurship

DATA MODELLING CYCLE

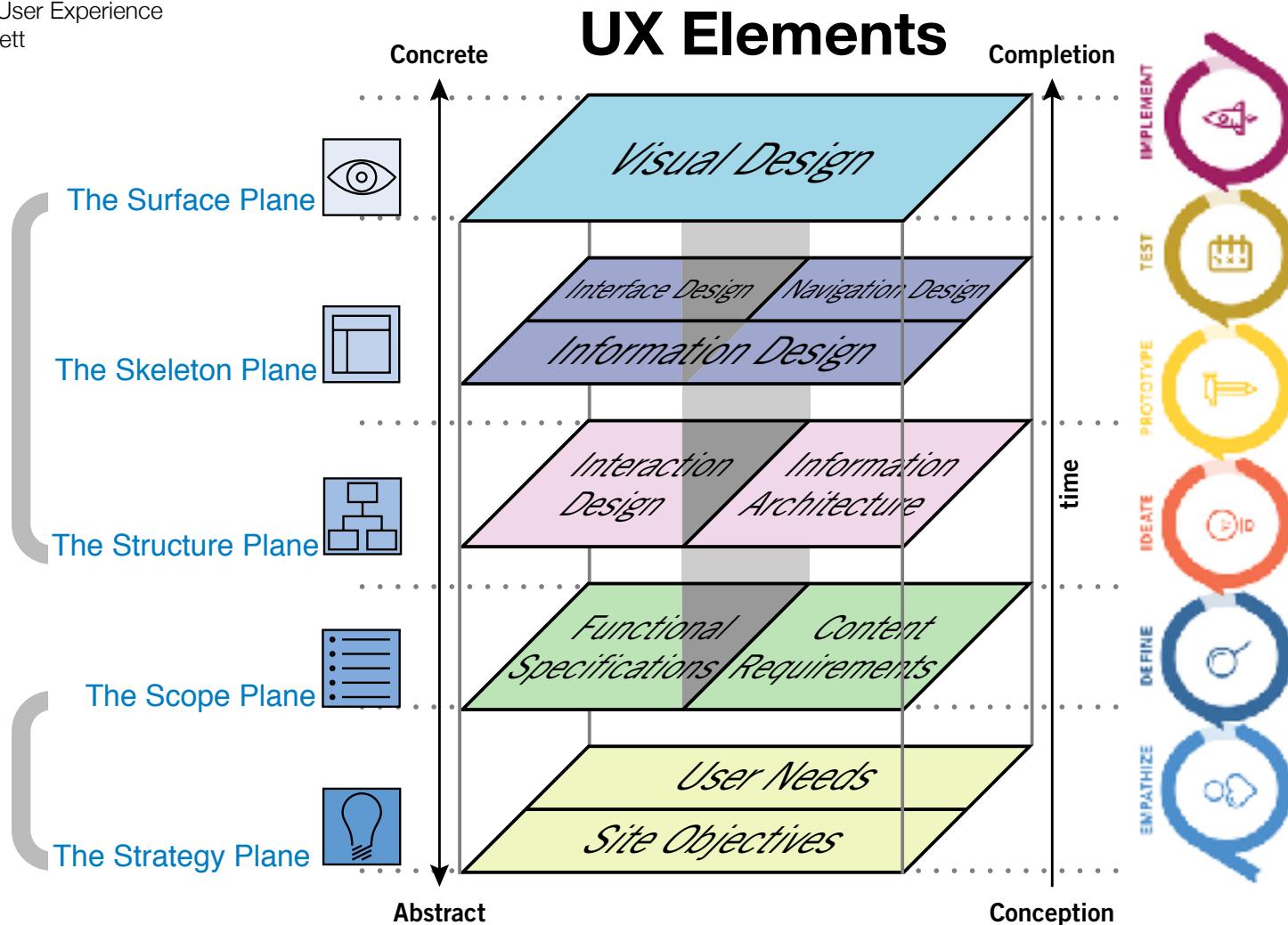
Source: Elements of User Experience
by Jesse James Garrett

Problem Solution

how and
how much

Problem Definition

who, what,
and why



HOME

LOBBY

ROOM

BOOKING

CHECK-IN

MEETUP

Source: MIT Mobile
Experience Lab

USER JOURNEY MAP

profile, activity, points

Booking
Data

Check-in
Data

Interact-
ion Data

Review
Data

Give
check
in point



profile, activity, points

Profile, activity, points
of other guests at the
same table. Table lights up
to show activity, interests shown
on table & lobby display map.
Profiles added to app network.

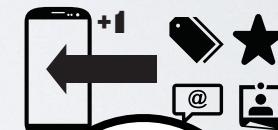
INTERACTIONS

INTER-
OUTSI-
ERNS



Data-
base

profile, activity,
points



Places
Media
Group
Actions.

| Design Thinking | Computational Thinking | Description |
|--------------------------|-------------------------------|--|
| Empathy | Decomposition | Collect and analyse <u>stories and data</u> to understand the stakeholders and discover their needs. |
| Definition | Patterns | Synthesise recurring <u>persona types, contexts, artefacts, and scenario</u> patterns to formulate problem (who, what, and why). |
| Ideation | Abstraction | Develop socio-cultural and technical systems to reshape user stories and data flow. |
| Prototyping | Algorithm | Build <u>experience prototype and computational models</u> to represent future scenarios for validation. |
| Testing & Implementation | Automation & Evaluation | Continuous testing, improvement and automation to evaluate <u>functional, emotional, social, economic and environmental</u> impacts (how and how much). |

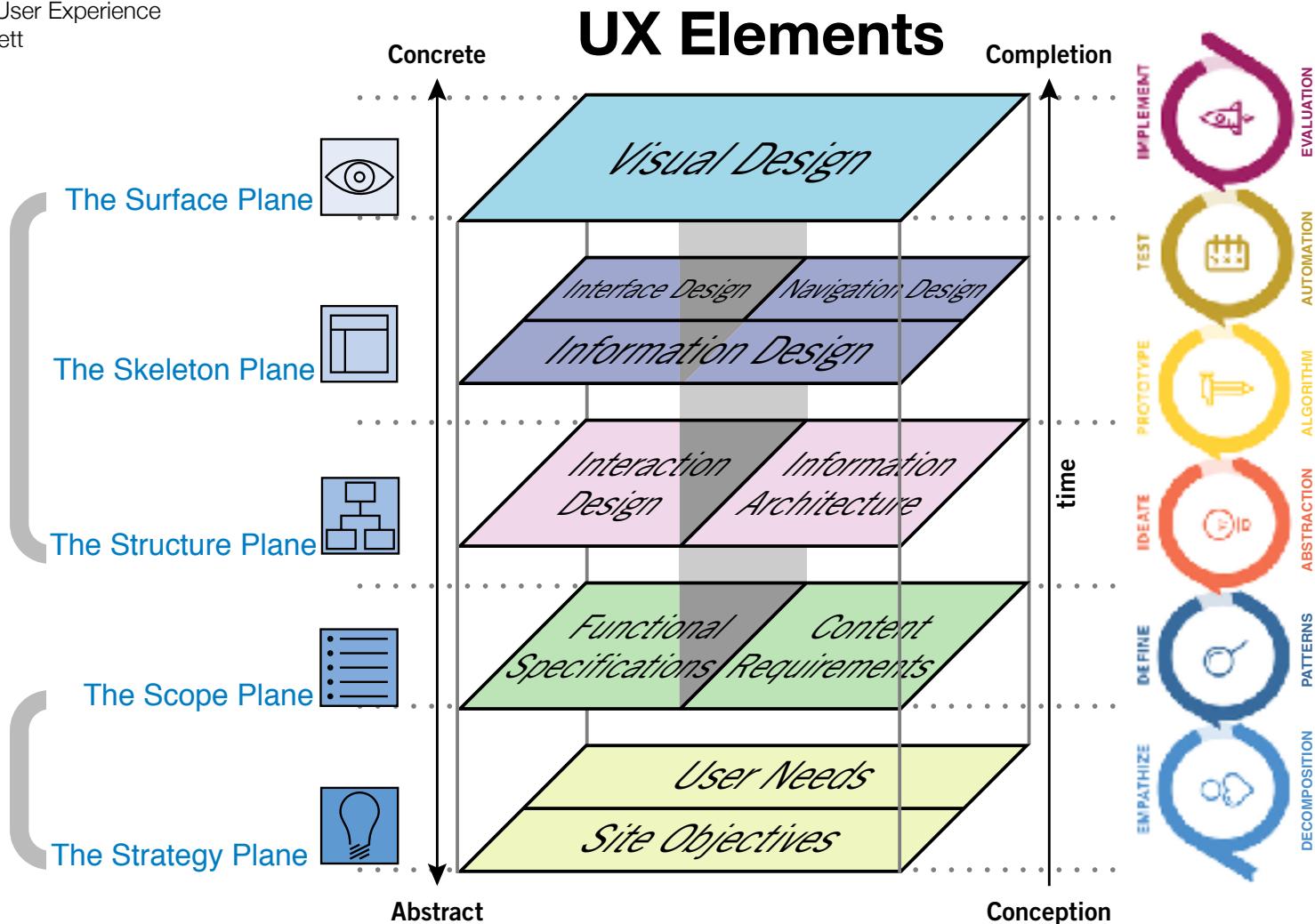
Source: Elements of User Experience
by Jesse James Garrett

Problem Solution

how and
how much

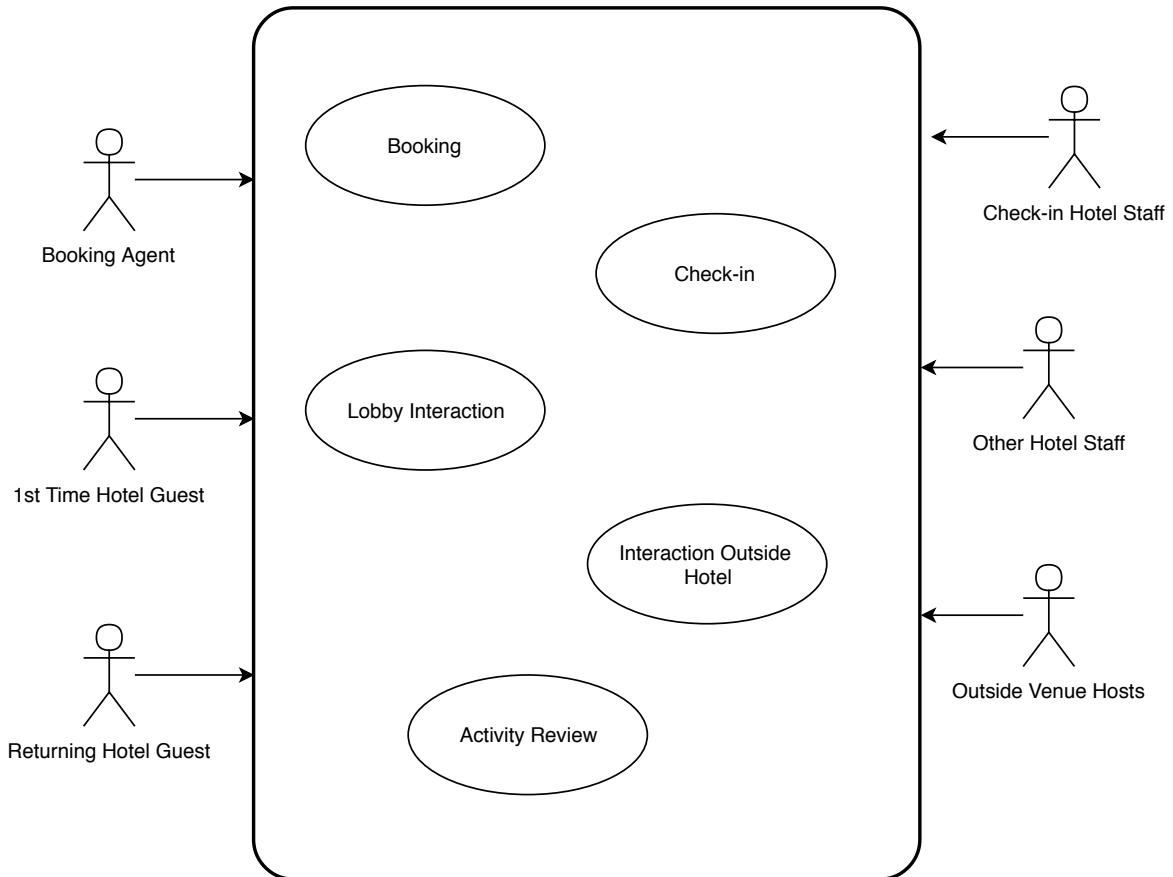
Problem Definition

who, what,
and why



Develop use cases from journey map.

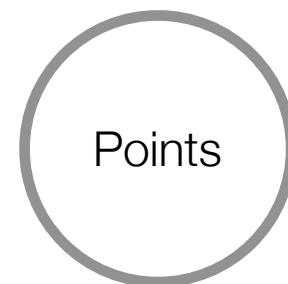
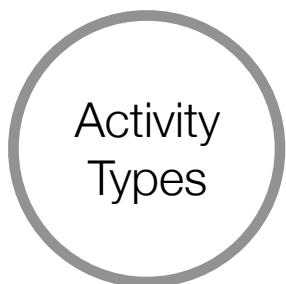
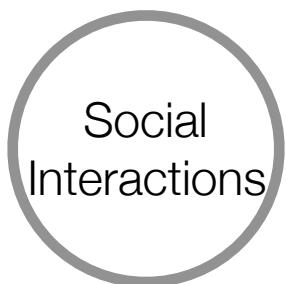
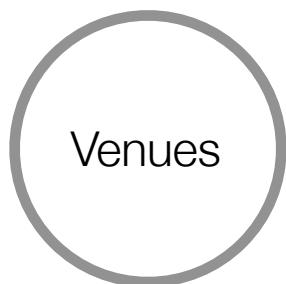
**To understand and construct our world
through data models.**



Use Cases of a Hotel Loyalty App

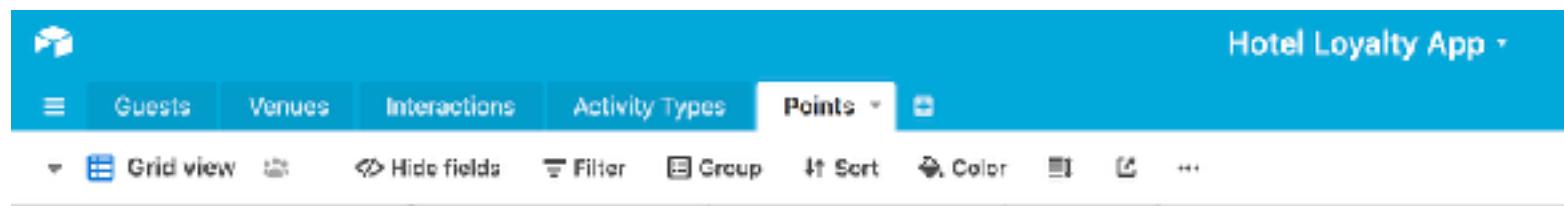
Derive data entities from use cases.

Data Entities to be Tracked

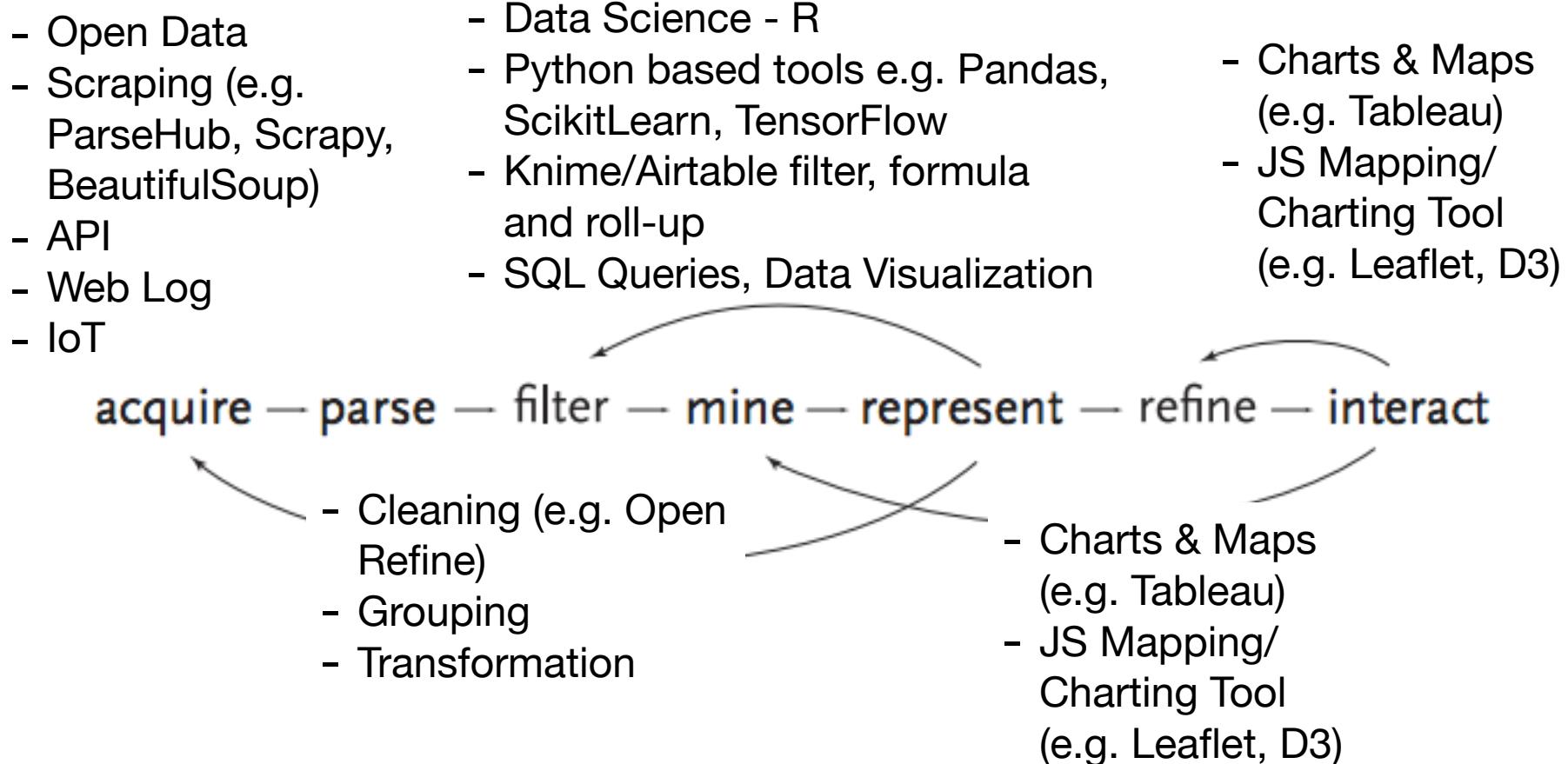


- e.g.
 - 1st time guests
 - Returning guests
- e.g.
 - Starbuck
 - Bar inside hotel
 - Boston Metropolitan Museum
- e.g.
 - Add friends
 - Take pictures
 - Share rides
 - Split bills
- e.g.
 - Museum tour
 - Bar hopping
 - Scenic picture taking
 - Business meeting
- e.g.
 - Revisit
 - Dine in hotel
 - Shop in hotel
 - Initiate contact
 - Give reviews

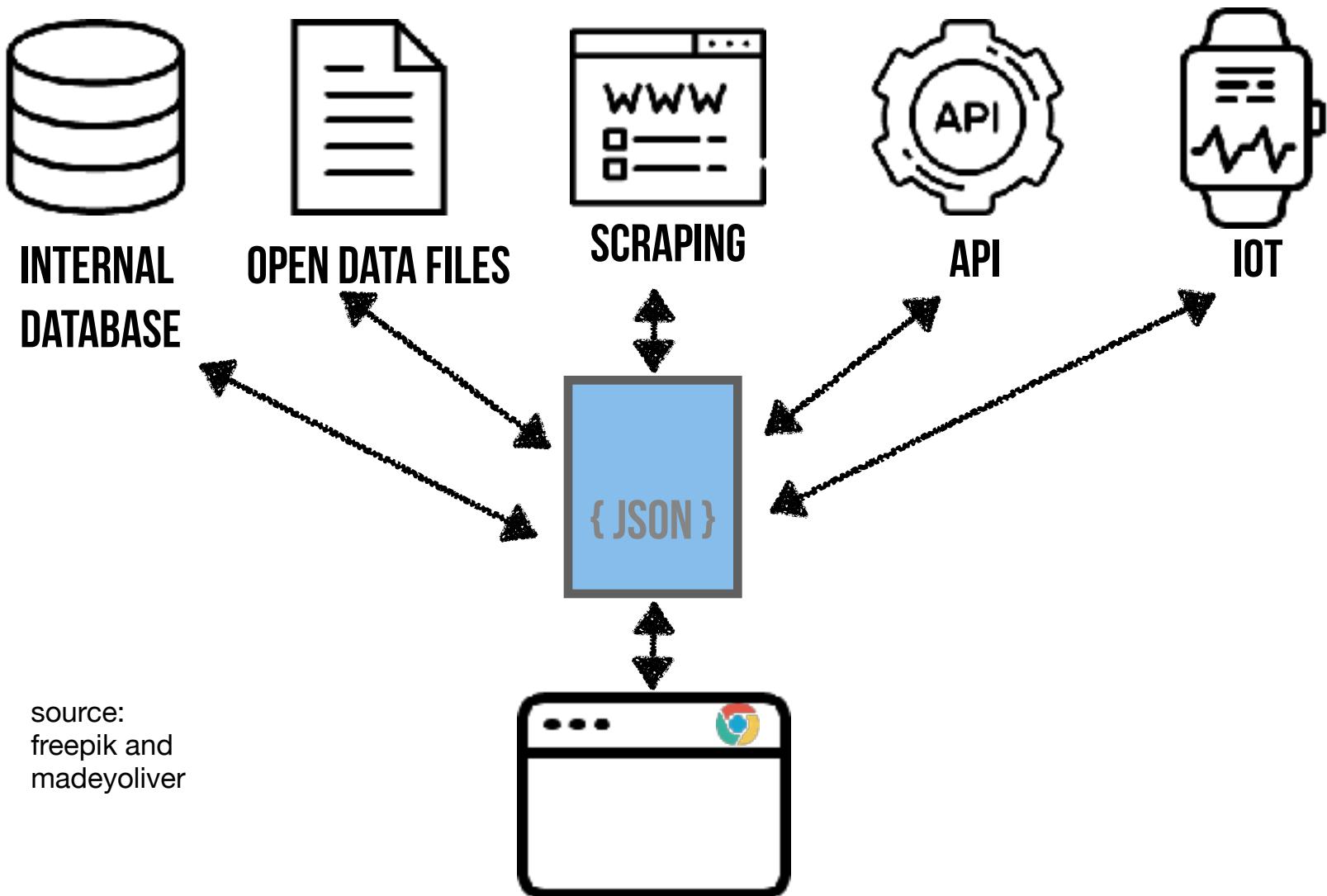
Tracking Data Entities



Where does the data come from?



source: Ben Fry's Visualizing Data



RISE OF THE SMART CITIES & OPEN DATA

香港智慧城市 Hong Kong Smart City 藍 Blueprint 圖



NYC OpenData[Home](#) [Data](#) [About](#) ▾ [Learn](#) ▾ [Alerts](#) [Contact Us](#) [Blog](#)

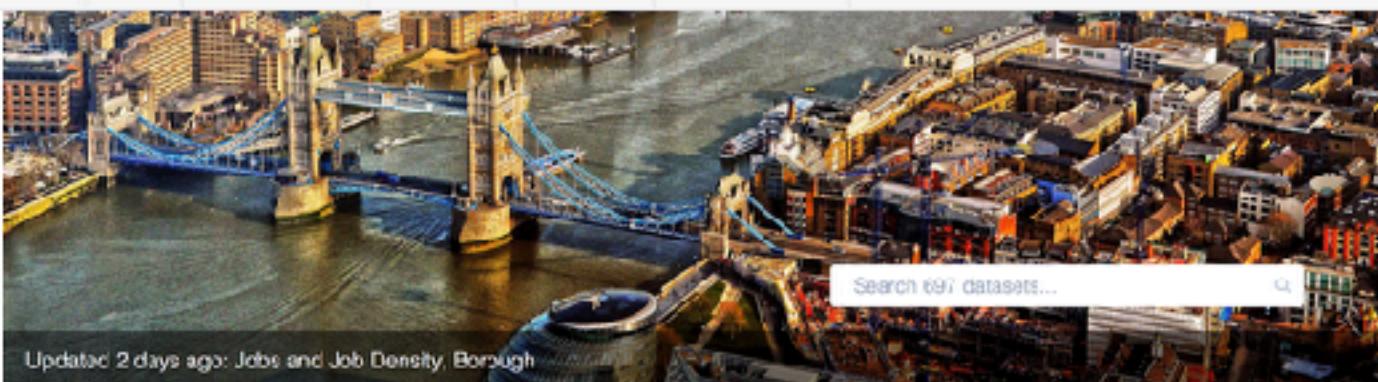
Open Data for All New Yorkers

Where can you find public Wi-Fi in your neighborhood? What kind of tree is in front of your office? Learn about where you live, work, eat, shop and play using NYC Open Data.

Search Open Data for things like 311, Buildings, Crime



LONDON DATASTORE

[Register an Account](#)[Login !\[\]\(3b451835b5cf44dc087a11f8c88642da_img.jpg\)](#)[Blog](#)[Data !\[\]\(380b8c89d31b6e4bc43715f362c2f817_img.jpg\)](#)[Apps & Analysis !\[\]\(b0ba827d2a487ec829cbd8e604d56e0e_img.jpg\)](#)[Developers !\[\]\(f42f75e81f61fd5f9f313cac09f9aa39_img.jpg\)](#)[Boroughs !\[\]\(c7143b06b3915be2311cf128bb2424aa_img.jpg\)](#)[City Data Strategy !\[\]\(eca9dca13a2688a2d5e75f35d4cc16aa_img.jpg\)](#)[More !\[\]\(6d4cdef06d803dcd63aa9aa8840a2ace_img.jpg\)](#) Search 697 datasets... 

Updated 2 days ago: Jobs and Job Density, Borough

JOBS AND
ECONOMY

TRANSPORT



ENVIRONMENT

COMMUNITY
SAFETY

HOUSING



COMMUNITIES



HEALTH

LONDON AS A
WORLD CITYGLA
PERFORMANCE

Click on a circle to see more...

Welcome to the Datastore

The London Datastore is a free and open data sharing portal where anyone can access data relating to the capital. Whether you're a citizen, business owner, researcher or developer, this site provides over 700 datasets to help you understand the city and discover solutions to London's problems. Please do not use this site to report an emergency - [call 999](#).



Hong Kong Smart City Blueprint

[Home](#) [Vision & Mission](#) [Development Plans](#) [Open e-Blueprint](#) [Download Blueprint](#) 繁體 AA

Open Data



[Home](#) > Open Data

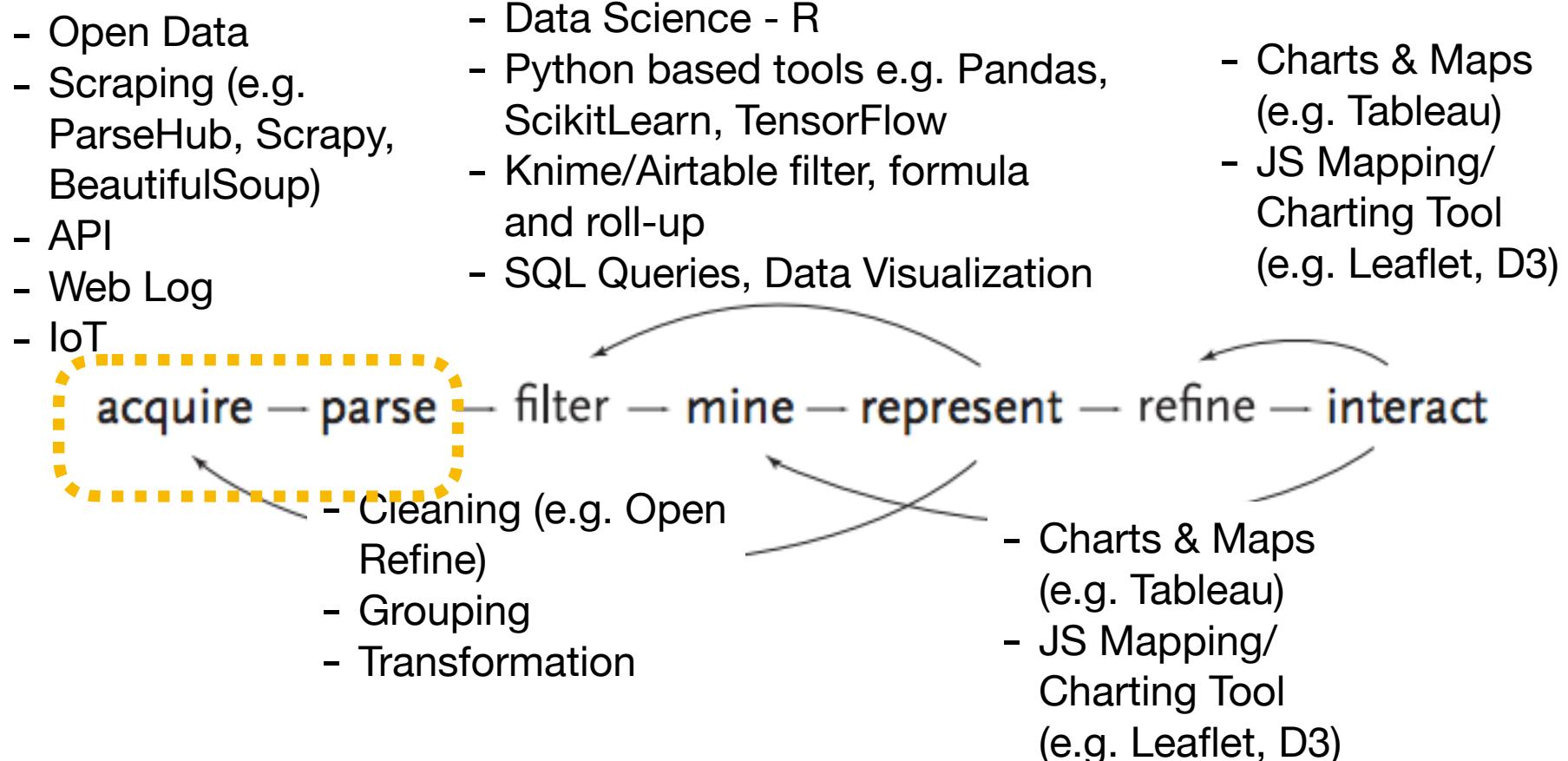
Smart Mobility

Smart Living

Smart Environment
Smart People

Smart Government
Smart Economy

Open Data



source: Ben Fry's Visualizing Data

THE REVAMPED PORTAL

HK

POPULAR DATA
GEOSPATIAL DATA
FEATURED DATASETSHow to The Site? **START HERE**Search Data e.g. population 

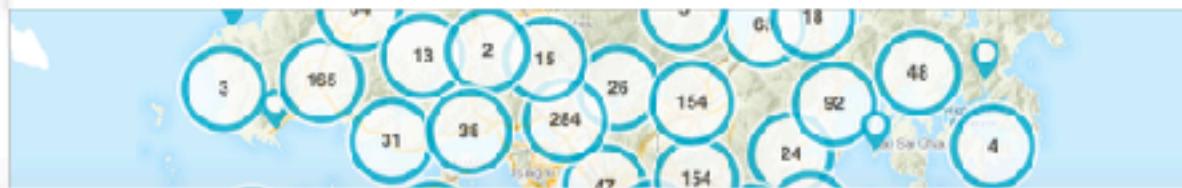
BROWSE DATASETS

City
ManagementClimate and
WeatherCommerce
and Industry

Development



Education

Employment
and Labour

CLASS EXERCISES

<https://www.luxonsoftware.com/converter/xmltocsv>

<https://json-csv.com/>



A web scraping tool that is easy to use

ParseHub is a free web scraping tool. With our advanced web scraper, extracting data is as easy as clicking the data you need.

Download our free app



Open a website

Download our [desktop app](#). Choose a site to scrape data from.



Click to select data

Get data from multiple pages. Interact with AJAX, forms, dropdowns, etc.



Download results

Access data via JSON, Excel and [API](#). Data is collected by your servers.



about:parsehubupdate

about:parsehubupdate

ParseHub

Double tap ⌘ to view all ParseHub hotkeys.

Connecting to the ParseHub server.

If this persists for more than a few seconds, update your ParseHub extension and restart Firefox.

DEVMode v2014 DEVMode Web tools

A preview of your data will appear here

API Tutorials Contact

Visuals enabled

Import data from any website... + https://www.parsehub.com/landing

parsehub

New Project

My Projects

Tutorials

Welcome!

Beginner Tutorials

1. Start Here: Create your first project
2. Extract text from a web page
3. Extract data from many pages (pagination)
4. Run project & download Excel & JSON data
5. Use the REST API

Advanced Tutorials

1. Collect data on a schedule
2. Enter text into a search box
3. Get data from behind a log-in
4. Infinite scrolling pages
5. Enter URLs for ParseHub to crawl

API Docs

Use our REST API and webhooks.

[View API >](#)

More Tutorials

Browse 80+ tutorials, videos & examples.

[View Tutorials >](#)

Contact Us

Need help with your project? Reach out.

[Contact >](#)

Product Demo

Get a custom product walk-through.

[Book a demo](#)

CSVExcel JSON

API Tutorials Contact

Visuals enabled



Welcome!

Beginner Tutorials

1. Start Here: Create your first project
2. Extract text from a web page
3. Extract data from many pages (pagination)
4. Run project & download Excel & JSON data
5. Use the REST API

Advanced Tutorials

1. Collect data on a schedule
2. Enter text into a search box
3. Get data from behind a log-in
4. Infinite scrolling pages
5. Enter URLs for ParseHub to crawl



API Docs

Use our REST API and
web-tools.

[View API](#) >

More Tutorials

Browse 50+ tutorials, videos &
examples.

[View Tutorials](#) >

Contact Us

Need help with your project?
Please our:

[Contact](#) >

Product Demo

Get a custom product
walk-through.

[Schedule](#) >

Sample is disabled when working on interactive templates. Use [Test Run](#) instead.



<http://www.dianping.com/hongkong/food>



A free, open source,
powerful tool for working
with messy data



[Home](#)
[Community](#)
[Documentation](#)
[Download](#)
[Contact Us](#)
[Blog](#)

Enhanced with Java profiler



Welcome!

OpenRefine (formerly Google Refine) is a powerful tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data.

OpenRefine is available in English, Chinese, Spanish, French, Russian, Portuguese (Brazil), German, Japanese, Italian, Hungarian, Hebrew, Filipino, Cebuano, Tagalog

OpenRefine is supported by:

Google News Initiative

Introduction to OpenRefine

1. Explore Data

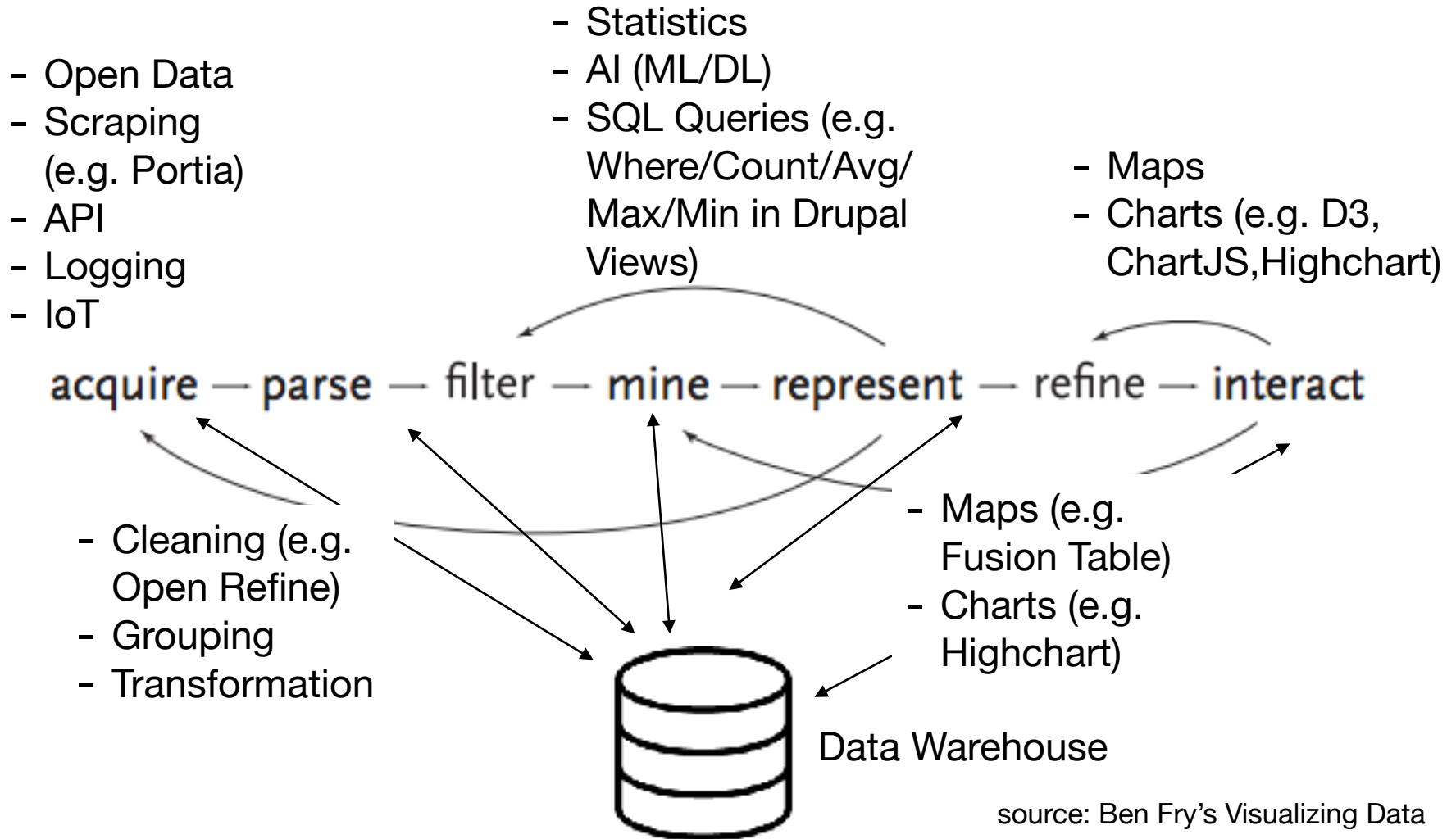
OpenRefine can help you explore large data sets with ease. You can find out more about this functionality by watching the video below and going through these articles.

The screenshot shows the OpenRefine interface with a data grid. The columns are labeled: 'Remove All', 'Merge', 'Delete', 'Current ID', 'Type of Identifier', 'Date of Record', 'Start Date', 'End Date', and 'Text value'. The data grid contains several rows of company names, such as 'GOOGLE INC.', 'GOOGLE LLC', 'GOOGLE HOLDINGS INC.', 'GOOGLE INCORPORATED', 'GOOGLE INCORPORATED', 'GOOGLE INCORPORATED', 'GOOGLE INCORPORATED', and 'GOOGLE INCORPORATED'. A video player at the bottom indicates the video is at 1:00 of 3:00.

Google refine

<http://d3-media.blogspot.hk/2013/11/how-to-refine-your-data.html>

HOW DOES THE DATA VISUALIZATION CYCLE WORK WITH DATABASE MANAGEMENT?



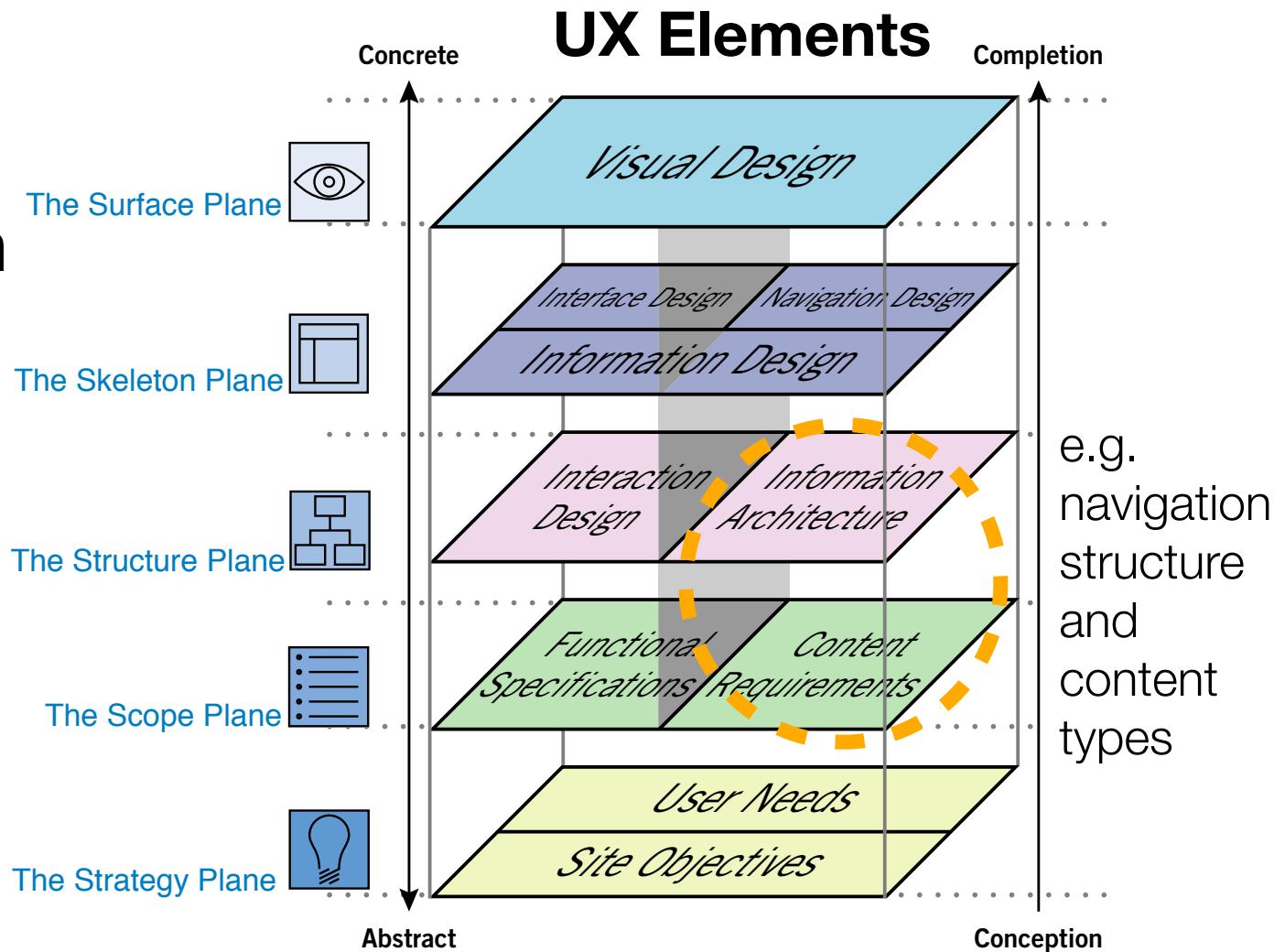
source: Ben Fry's Visualizing Data

FROM DATABASE MANAGEMENT TO CONTENT MANAGEMENT AND INFORMATION ARCHITECTURE

Can you see
a visualization
process
here?

Where does
data fit?

Source: Elements of User Experience
by Jesse James Garrett



Empathise
身同感受

Define
界定问题

Ideate
創意發想

Prototype
開展原型

**Test &
Implement**
測試執行

THE 5 STEP DESIGN THINKING JOURNEY

Source: Stanford D.School

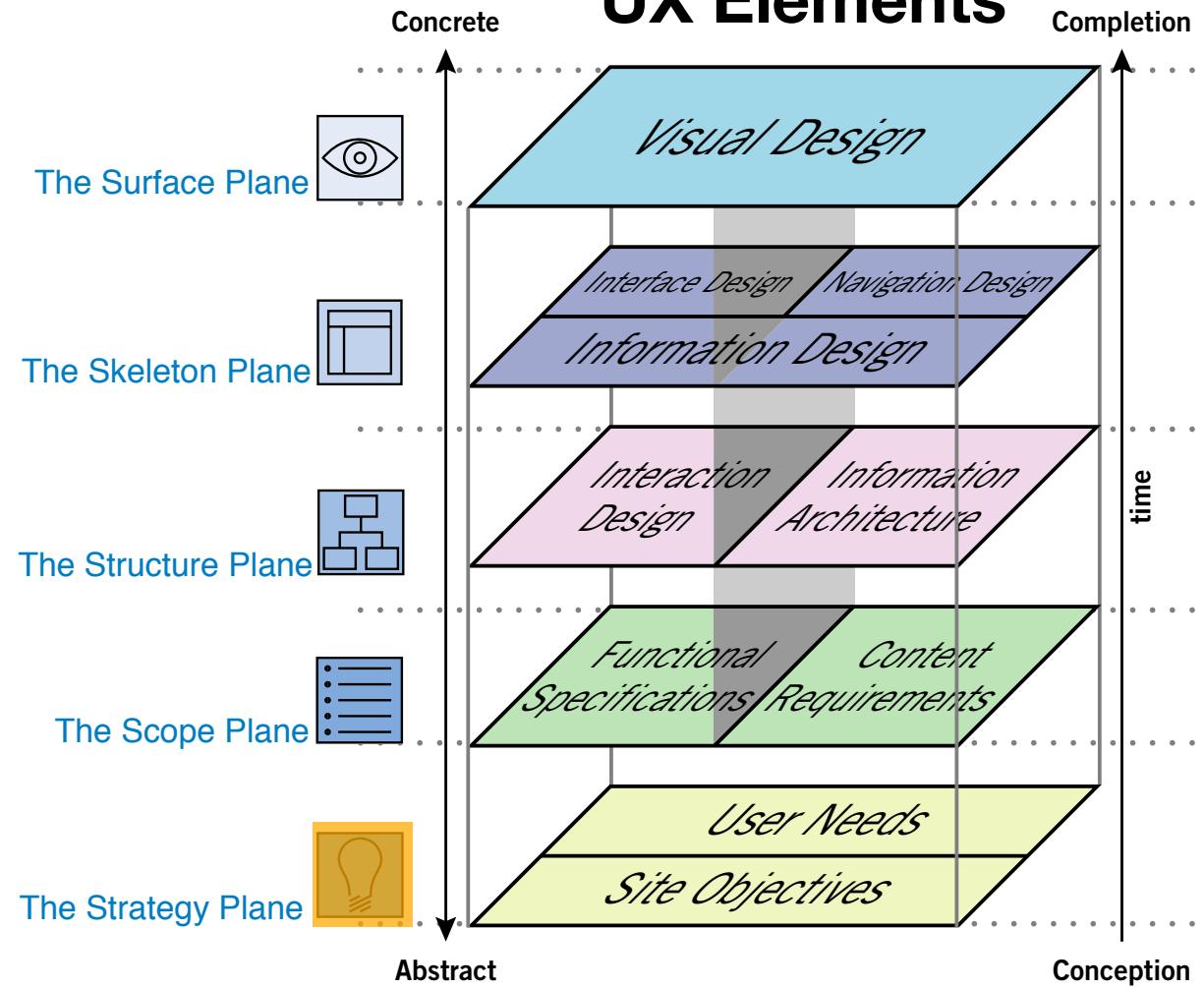


EXPERIENCE = JOURNEY

e.g. The Journey of a New CUHK Applicant

EXAMPLE:
THE JOURNEY OF A NEW CUHK APPLICANT

UX Elements



Source: Elements of User Experience
by Jesse James Garrett



About CUHK



Welcome Message

I am delighted that you are considering The Chinese University of Hong Kong for your undergraduate studies.

Founded in 1963 with a mission "to combine tradition with modernity, and to bring together China and the West", the University has, since its founding, developed into an academic institution of international standing, well-known for its excellence in both teaching and research.

Welcome Message

[Explore CUHK](#)
[Explore CUHK](#)
[Distinguished Scholars](#)
[Facts & Figures](#)
[Virtual Campus Tour](#)

Student Life

[Campus & Accommodation](#)
[Student Exchange Programme](#)
[HCARE Programme](#)
[Student Support](#)
[Voices of Students](#)

Colleges

[A Unique College System](#)
[College Assignment](#)

CUHK Through the Lens

Empathise
身同感受

Define
界定问题

Ideate
創意發想

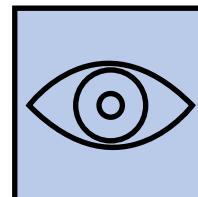
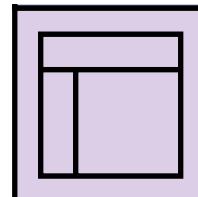
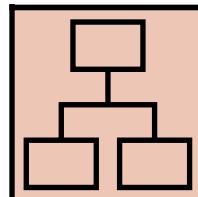
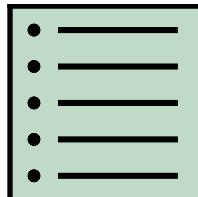
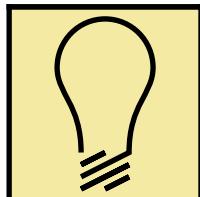
Prototype
開展原型

**Test &
Implement**
測試執行

THE 5 STEP DESIGN THINKING JOURNEY

Source: Stanford D.School

| SITE OBJECTIVES | | TARGET USERS | | USER NEEDS |
|-----------------|--|----------------|--|--|
| Primary goal | Become top tier university in the world. | Primary User | Top local students and world-class academics | Apply for degree enrollment and job vacancies. |
| Additional goal | Increase research output. | Secondary User | Chinese students interested in the west | Interested to come to CUHK to study. |
| Additional goal | Build strong alumni network. | Secondary User | Foreign students interested in PRC | Will consider CUHK for full-degree and exchange program. |



STRATEGY PLANE

Empathise
身同感受

Define
界定问题

Ideate
創意發想

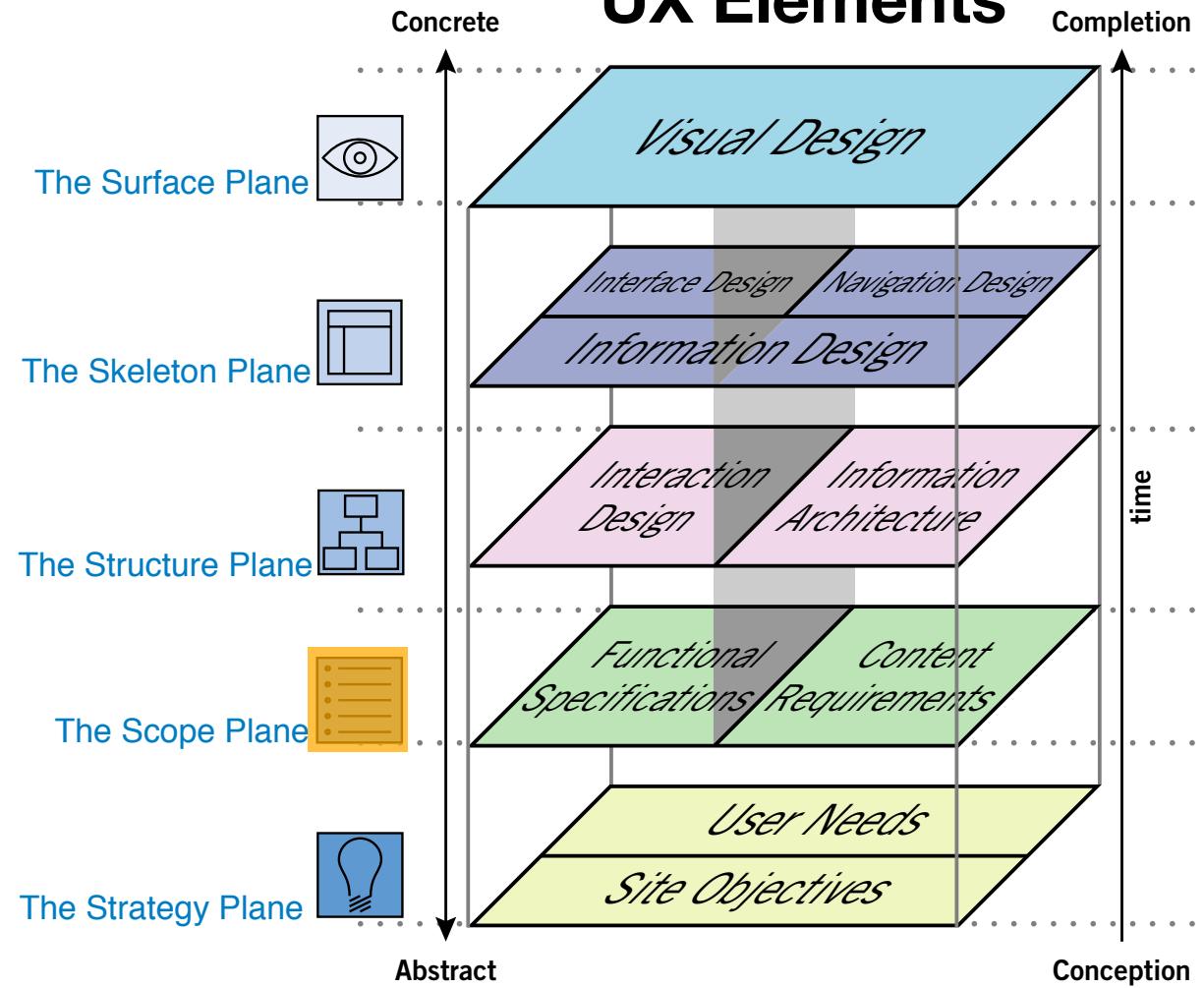
Prototype
開展原型

**Test &
Implement**
測試執行

THE 5 STEP DESIGN THINKING JOURNEY

Source: Stanford D.School

UX Elements



Source: Elements of User Experience
by Jesse James Garrett

THE JOURNEY AS A STORY

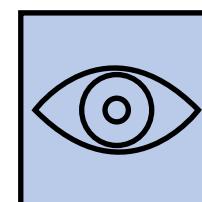
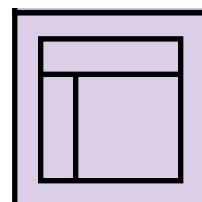
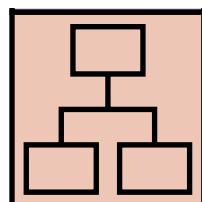
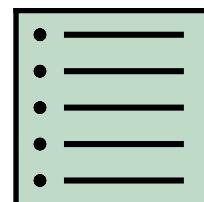
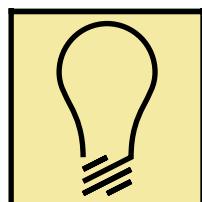
ELEMENTS

1. 人 (人物) PERSONA
2. 景 (場景) CONTEXT
3. 物 (物件) ARTIFACTS
4. 用 (使用) SCENARIOS

STRUCTURE

1. 起 (人物背景) BEGINNING
2. 承 (問題) CHALLENGE
3. 轉 (發展方案) RESPONSE
4. 合 (完成狀況) END

資料來源---劇本導引：
資訊時代產品與服務設計新法
作者余德彰，林文綺，王介丘



SCOPE PLANE

Journey Map (歷程图)

功用: 有助收集和分类用户的故事。

有助描绘用户、器物或系统的故事，以人景物用的故事元素、起承转合的故事结构，更好地理解目标对象的需求，如何在故事中浮现，并寻找设计机会。例如用户日誌，用户的生活地图，或產品如何在空间中移动的流程（从製造商到商店货架，再转送到用户的手上）。

参考资料来源: Designthinker Academy

| | 起 | 承 | 转 | 合 |
|----|---|---|---|---|
| 人物 | | | | |
| 场景 | | | | |
| 文物 | | | | |
| 情况 | | | | |

| | 起 | 承 | 转 | 合 | |
|----|---|--|--|---|-----------|
| 人物 | - Local - PRC/ International | Curious about the school and program | Getting serious and anxious | Excitement | |
| 场景 | online promotion | school recruitment | CUHK | online/email | |
| 文物 | website | brochure/ presentation | - campus map - hk map - application form | letter of acceptance | |
| 情况 | online research Persona wants to filter program | staff roadshow Persona meets professor and CUHK staff | school visit Check out the school | application Finalize options Accept offer | admission |

例子

TYPES OF PERSONA (e.g. STUDENT & STAFF)

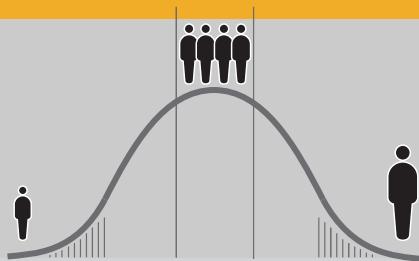
Persona (人物轮廓)

功用: 有助分类不同人物的轮廓。

人物轮廓包括个人档案（年龄、性别、教育程度、家庭关係、职业、收入等），生活方式偏好，购物和媒体使用习惯，人生目标和生活日程。工作小组必须為人物轮廓的类型数目达成共识，并且為每种类型命名，方便大家沟通。Persona 也可算在设计思维中很重要的pattern. 不同的persona有不同需要。

参考资料来源: Interaction Design Foundation

| | |
|--|---------|
|  | 生活方式偏好: |
| 名字: | 购物习惯: |
| 年龄: | 媒体使用习惯: |
| 性别: | 人生目标: |
| 教育程度: | 生活日程: |
| 家庭关係: | |
| 职业: | |



OBSERVATION AND UNDERSTANDING



LOCAL APPLICANTS



FOREIGN APPLICANTS

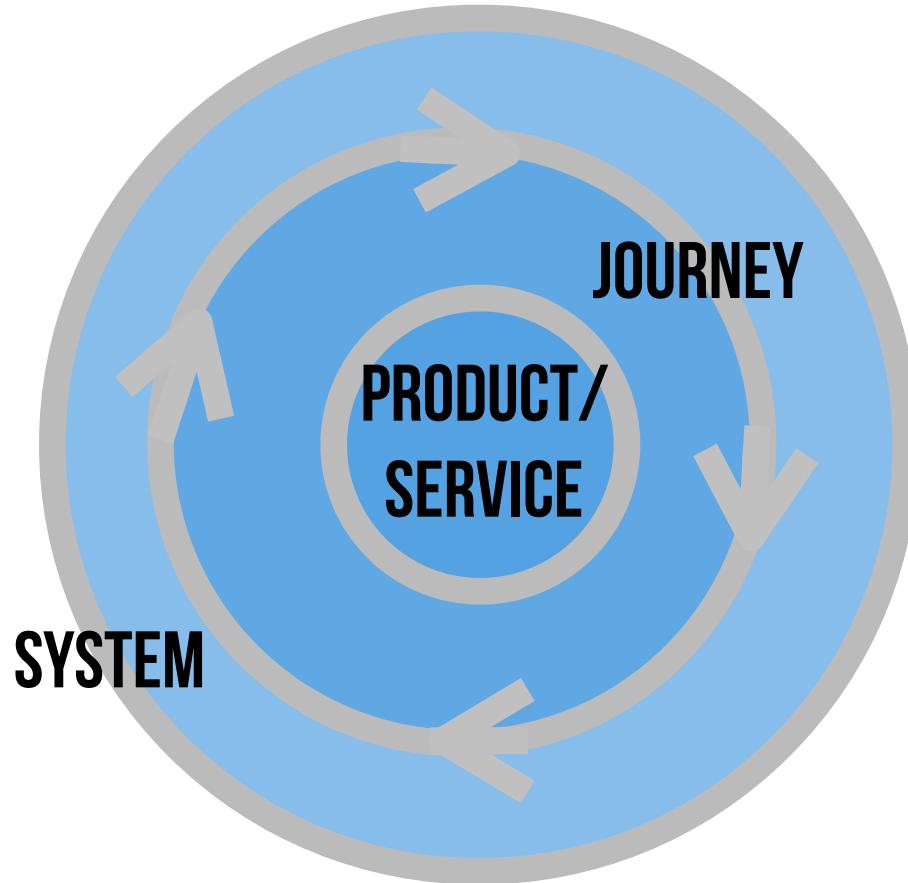


MAINLAND APPLICANTS

PROBLEM STATEMENT:

How can CUHK's website help new student and staff applicants **pursue their career objectives** over the alternatives out there while at the same time **communicate** the school's brand to continue attracting the best?

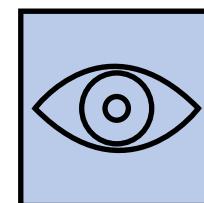
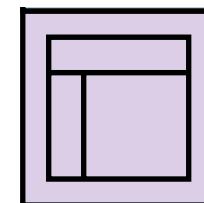
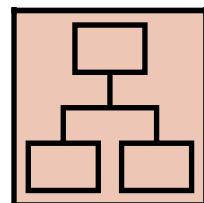
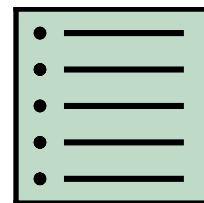
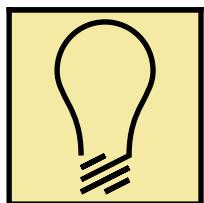
THE BIGGER PICTURE



- Product/Service: the offering (e.g. enrollment at CUHK) which is designed to address a need (e.g. world-class education).
- Journey: the process between a start and end point through which the user can have one's needs satisfied by accepting the offering (e.g. from gathering school info to final admission).
- System: the technical and social systems (e.g. school/faculty administration/research/teaching/student life) existing in an environment where competitive alternatives (e.g. other schools/programs) exist to address the same need. The question is: what are the compelling reasons for the persona to choose one offering over the others?

Content Production Plan

| Level | Section | Description | Status | Deadline | Handled by |
|-------|------------|--|-------------|----------|------------|
| 1.0 | About Us | An overview of the organization | Current | NA | Peter |
| 1.1 | History | Historical background & timeline | In progress | 15/10/11 | Mary |
| 1.2 | Contact Us | Telephone no., address, and contact | New | 22/10/11 | John |
| 2.0 | Product | Listing of current products and services | In progress | 25/10/11 | Henry |



SCOPE PLANE

Empathise
身同感受

Define
界定问题

Ideate
創意發想

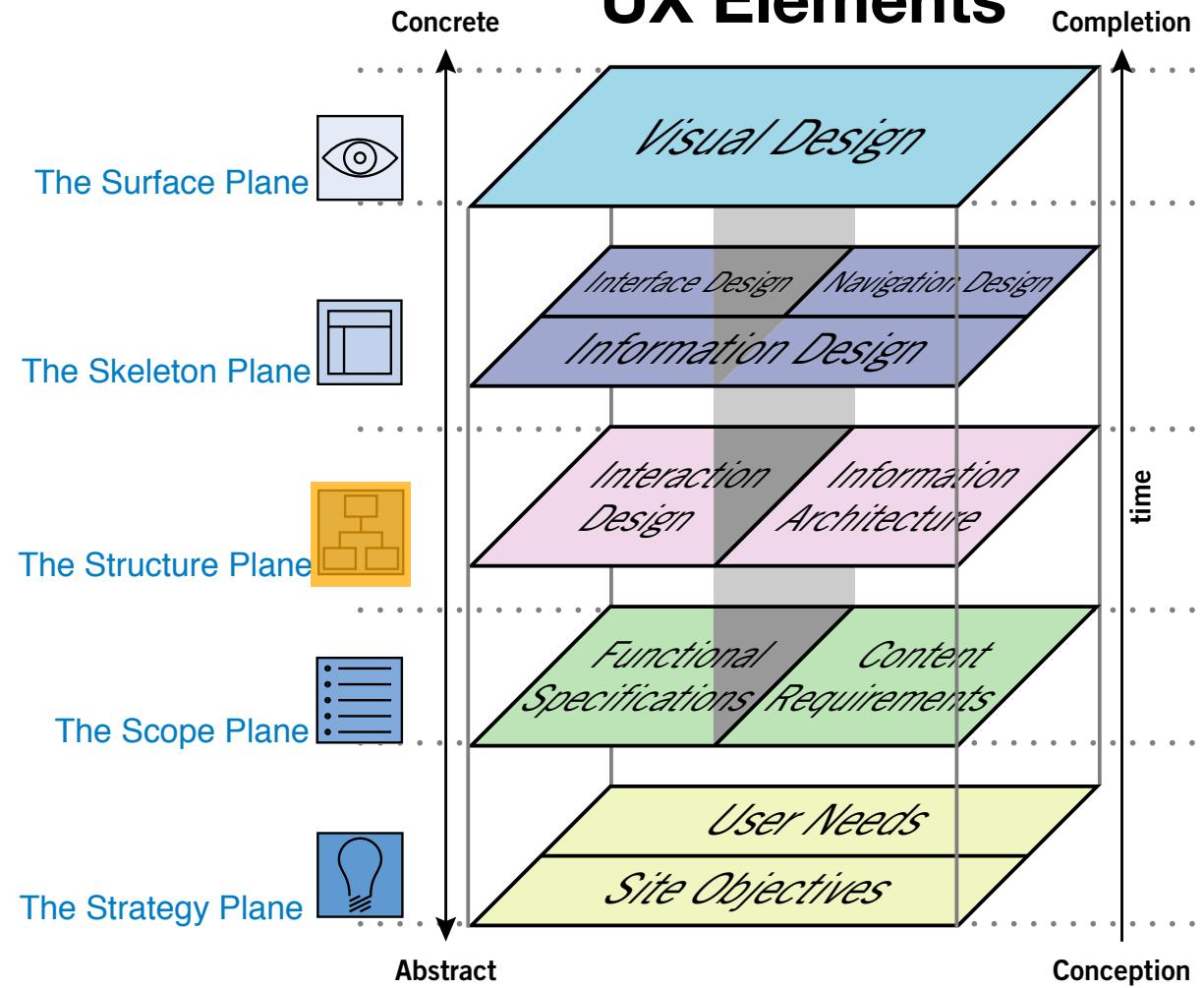
Prototype
開展原型

**Test &
Implement**
測試執行

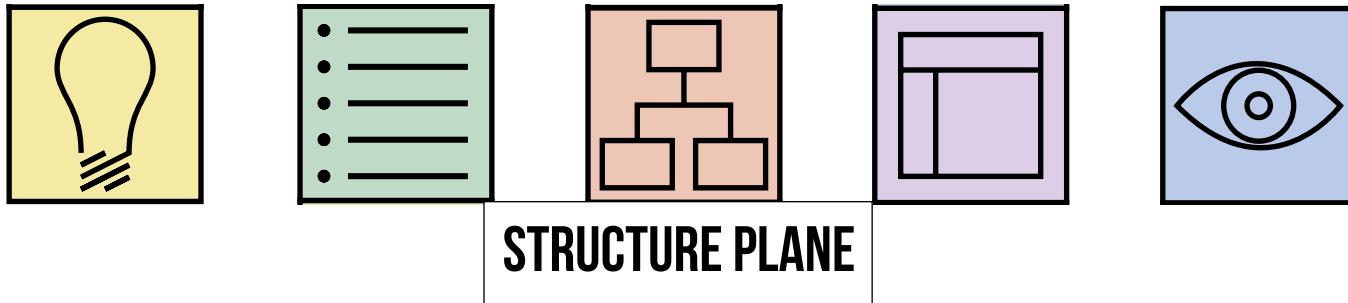
THE 5 STEP DESIGN THINKING JOURNEY

Source: Stanford D.School

UX Elements



Source: Elements of User Experience
by Jesse James Garrett



Source: Elements of User Experience
by Jesse James Garrett

Information architecture combines the art and science of structuring, organizing, labeling, navigating, and searching information space.

How do you organize this?



Content Organization Scheme

Logical Grouping

- **Alphabet**
- **Time**
- **Place**
- **Hierarchy**
- **Category**

Source: Richard Wurman, “Information Anxiety”

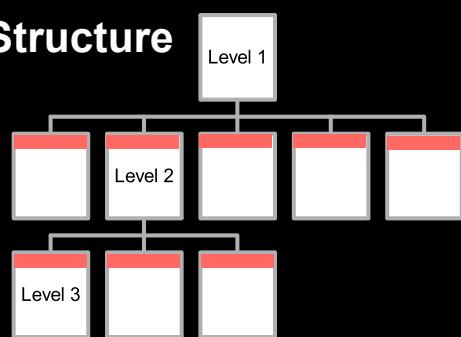
Visual Grouping

- **Linear**
- **Parallel**
- **Web**
- **Matrix**
- **Hierarchical**
- **Spatial Zoom**
- **Overlay**

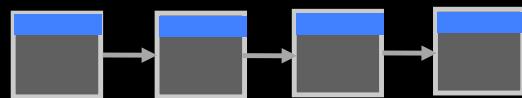
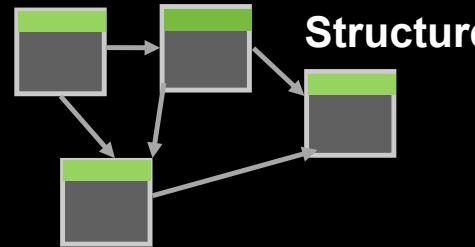
Source: Clement Mok, “Designing Business”

Information Architecture

Hierarchical
Structure



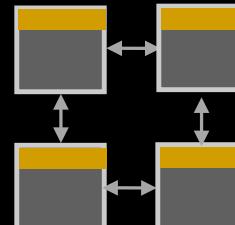
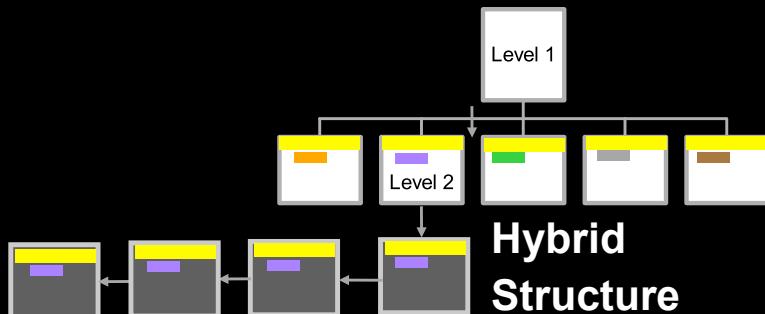
Organic
Structure



Sequential
Structure

Level 1

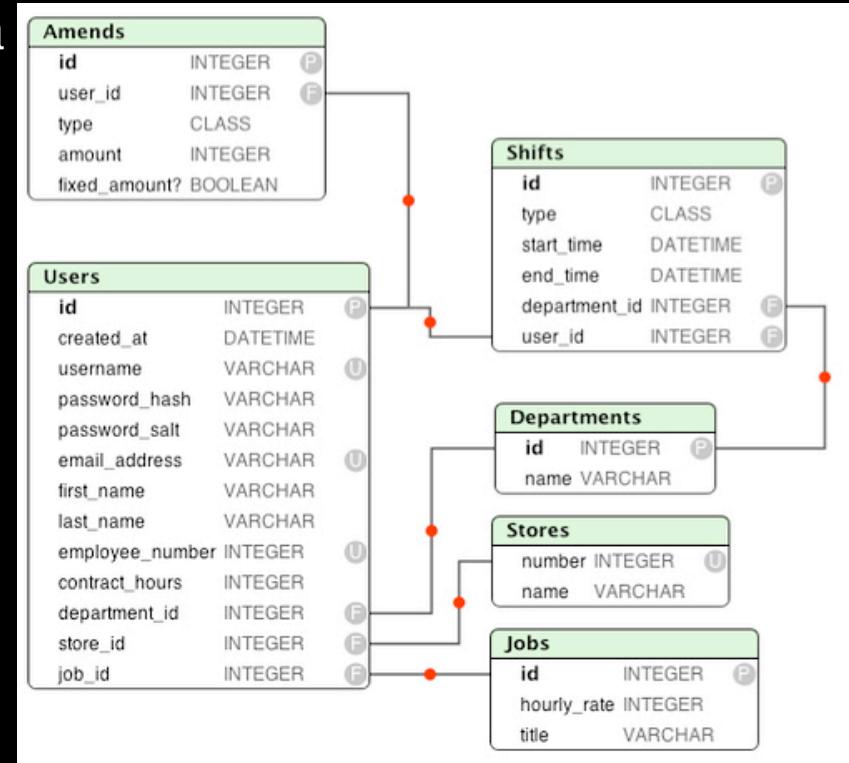
Hybrid
Structure



Matrix
Structure

Database can be defined as a collection of tables and table can be defined as collection of records. Each record is an unique entry of an object or event, which is made up of attributes:

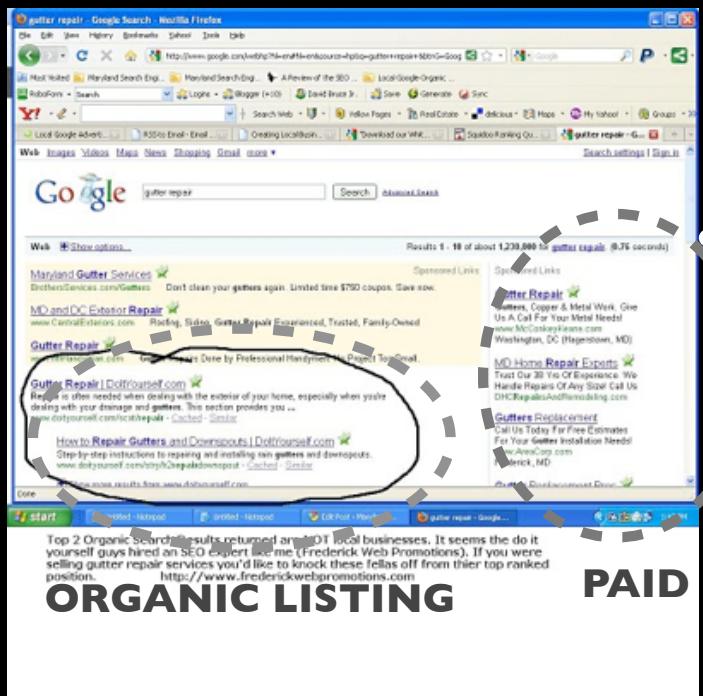
- ✓ Most important attribute is a key identifier.
- ✓ There are primary key and secondary key.
- ✓ Information between tables are related by keys.
- ✓ Data about data is called meta-data (e.g. Thesaurus, glossary, controlled vocabulary which can cross-referenced the database)



Source:Flickr

4

SEARCH INTERFACE



PAID LISTING



6

RESULT LISTINGS

CRAWLER/SPIDER/WEB BOT

1



Meta-Data Defined
QUERY ENGINE SEARCH INDEX



3

INDEXER BUILD UP WEB PAGE INDEX DATABASE



2

CRAW THE WEB

Question: Can you give me an example of what cannot be easily searched by web spider/crawler?

What is Web Analytics?

Web-based Tool Used for Analyzing Traffic to Your Site:

- What pages have been visited? How many times?
- Where your site visitors come from? Who are they?
- When do they visit? How frequent?
- How long do they stay? What do they buy?
- Where do they stop ? How frequent ?

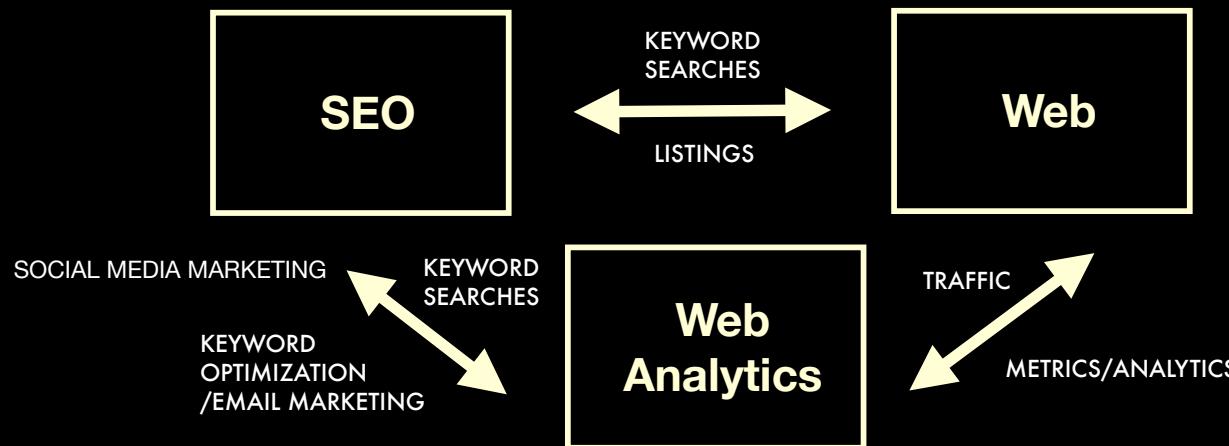
Organic and Paid Listings

- Organic (Natural) Listing (Content Driven - Page Rank Algorithm)
 - Tags (Meta, Title)
 - Body content
 - Links to you (negotiate partnership, build numerous websites, blogs linking to yourself)
- Paid Listing (Pay Per Click:view-->click-->transaction)
 - Sponsored ads
 - Conversion rates (browser->member->customer)

Why Web Analytics?

- Feedback to improve web design
- Know whether the content is enticing
- Has the user experience been captured?
- Who are the real users?
- What do they really care?
- How to make them come back for more?
 - conversion ratio (visitor-->member-->customer)

SEO and Web Analytics



- Web analytics tell you whether your web pages attract traffic (eye balls) or not.
- With feedbacks provided by web analytics, you make changes to your web pages, site structure, title, meta-tag, key word, and link strategies.

| Keyword | Search Popularity | Relevance | Competition | Landing Page/Section |
|---------|-------------------|-----------|-------------|----------------------|
| | | | | |
| | | | | |
| | | | | |

Source: Search Engine Optimization An Hour a Day by
Jennifer Grappone and Gradiva Couzin



1. Google Trends - <https://trends.google.com/trends/?geo=US>
2. Google Correlate - <https://www.google.com/trends/correlate/>
3. Keyword Shiffer - <https://keywordshitter.com/>
4. Keyword Everywhere - <https://keywordseverywhere.com/>
5. Answer the Public - <https://answerthepublic.com/>

WHY DO WE NEED CONTENT MANAGEMENT SYSTEM?

Data entities and taxonomy help build the information architecture of the website while the menu items, buttons, and links set the structure for navigation and interaction.

SEO relies on the architecture to build the right content (with proper keywords found in the content Title, Body, Links, and URL alias) to increase web traffic and user engagement!

Web analytics provides the feedback to track the results to see if the content and design is working or not.

END OF MORNING SESSION