
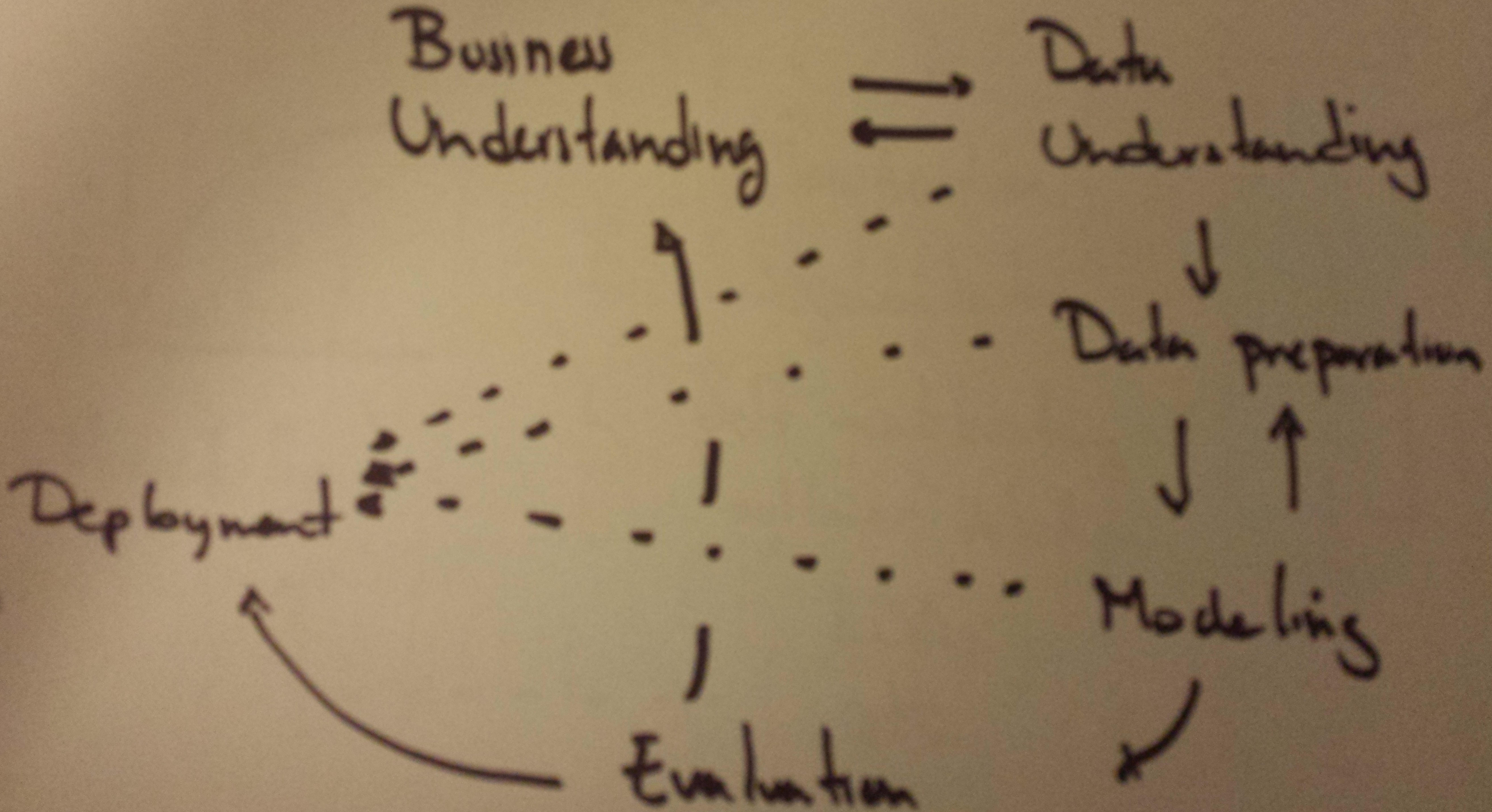


CRISP- PM




CRISP-DM


(Real)



¿Qué quiero hacer?



Predictión



Sensores

Logs

Web

Móvil

Recopilación
(Rápida)

Arquitectura
Web

Arquitectura
Móvil

Recopilación
(availability)

ETL

ETL

DW

BI

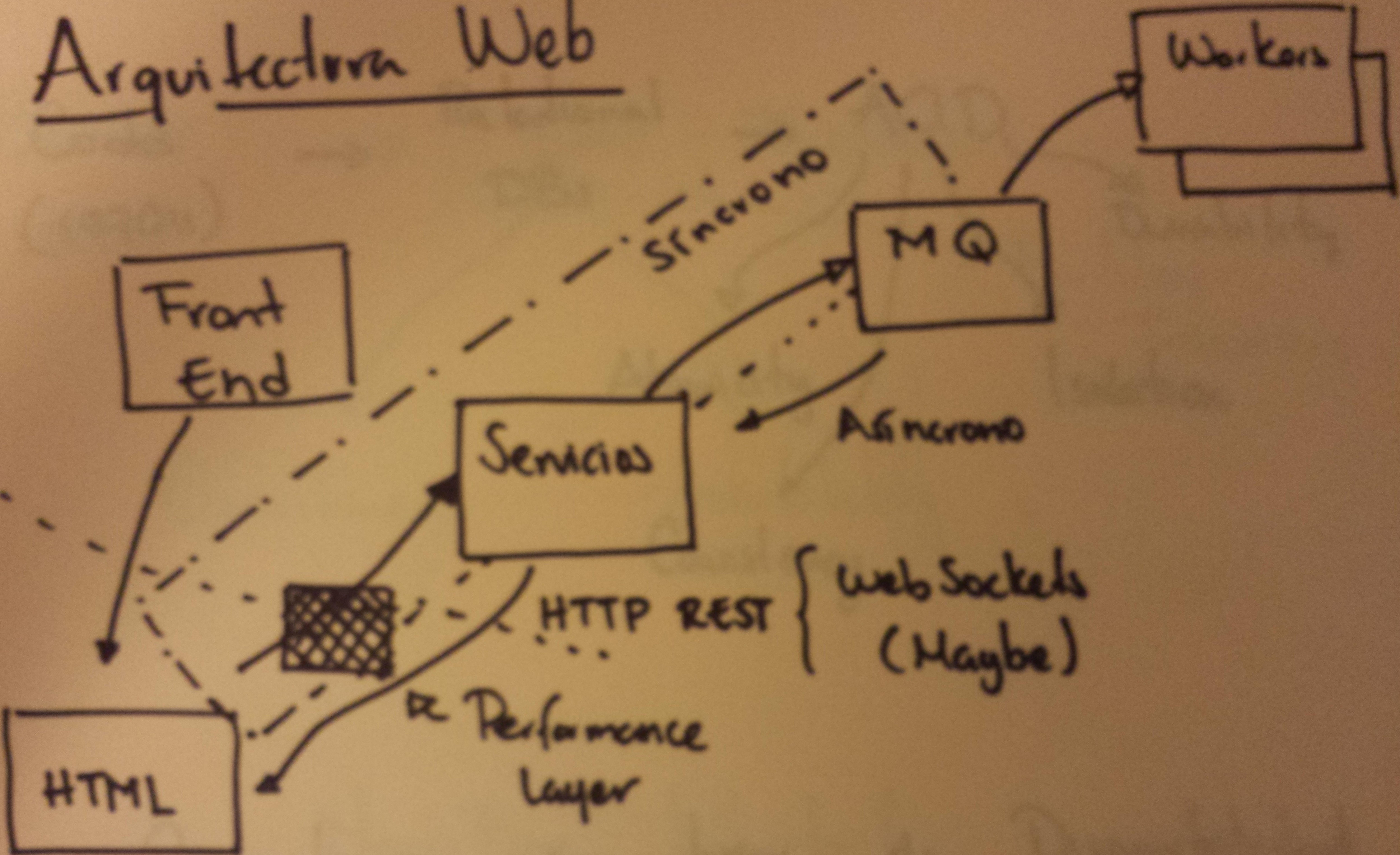
Análisis

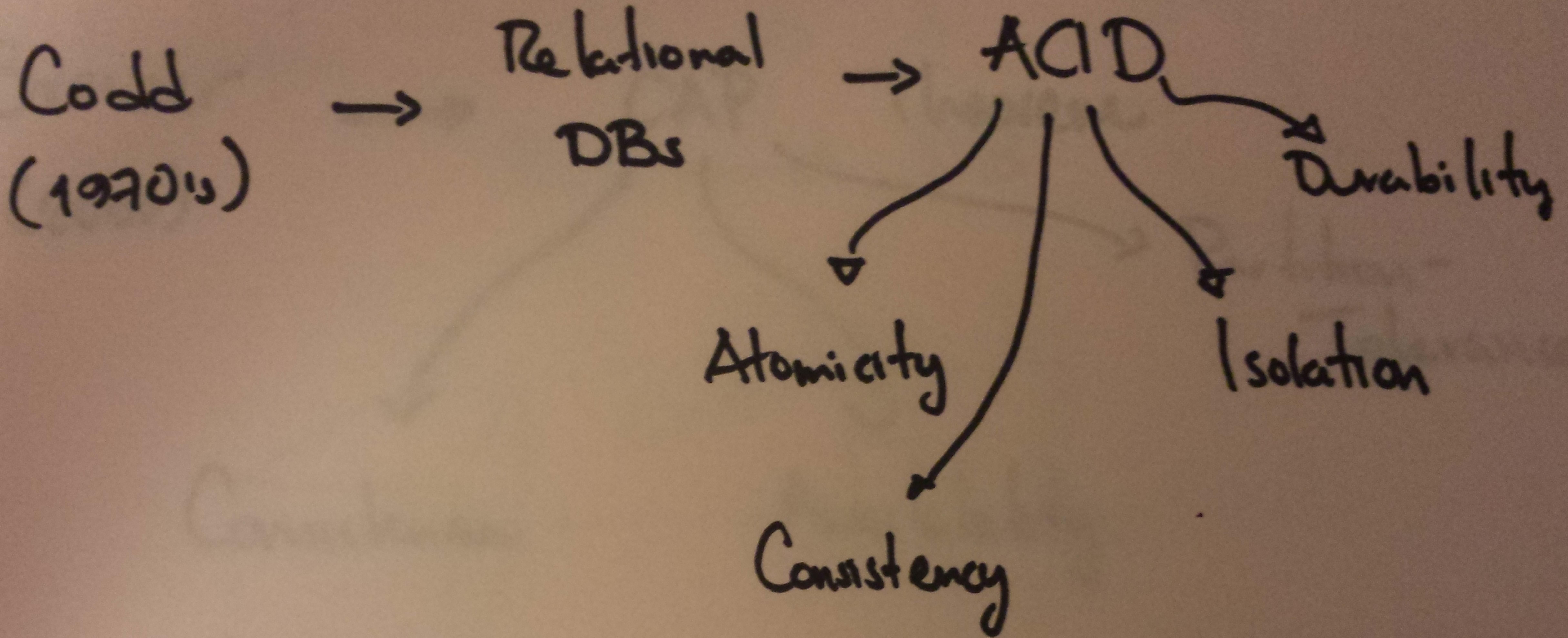
Visualización

Mediación

Arquitectura

Arquitectura Web





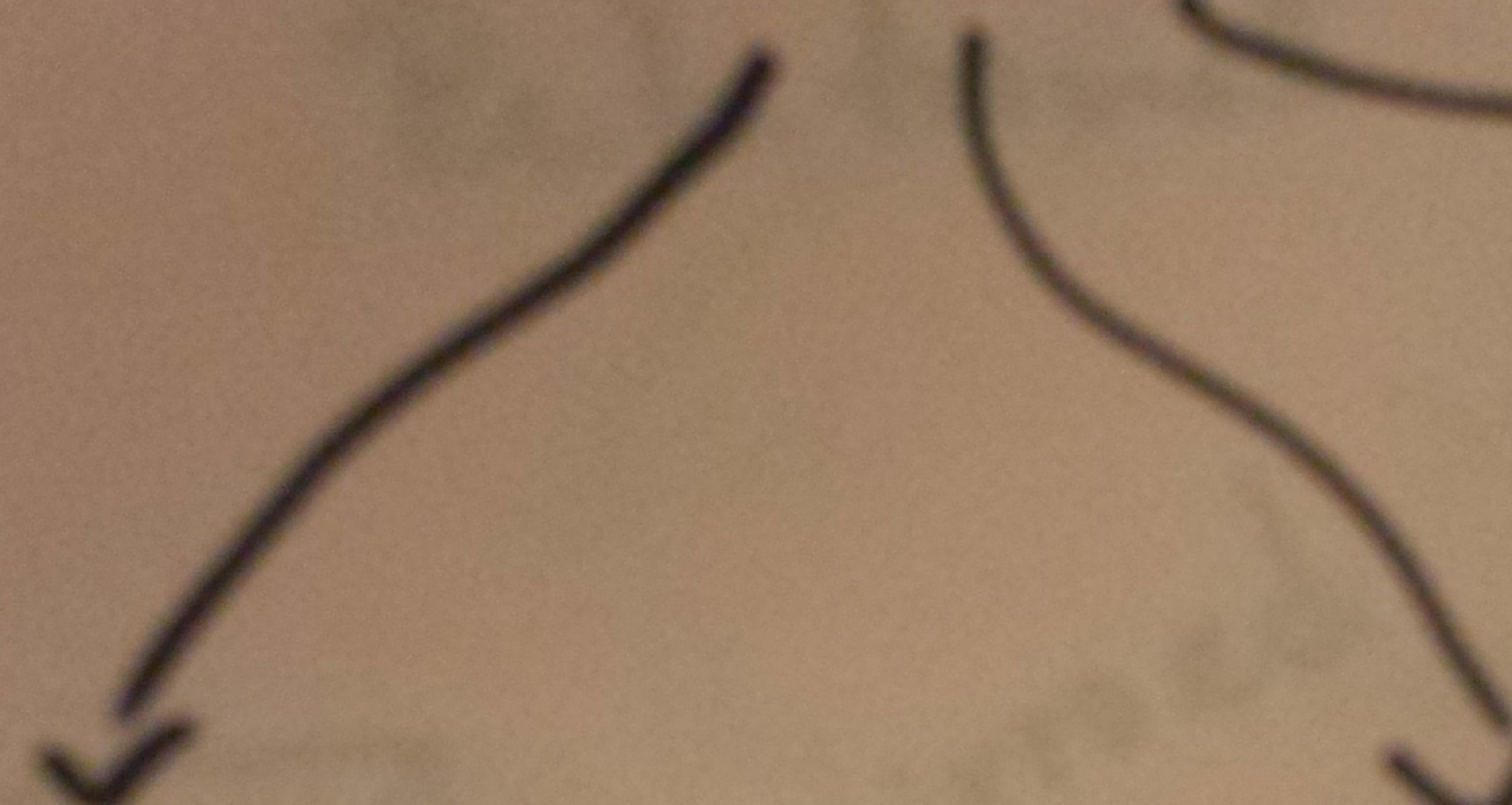
Consistencia en lugar de Disponibilidad

Brewer
(2000)



CAP

Theorem



Availability

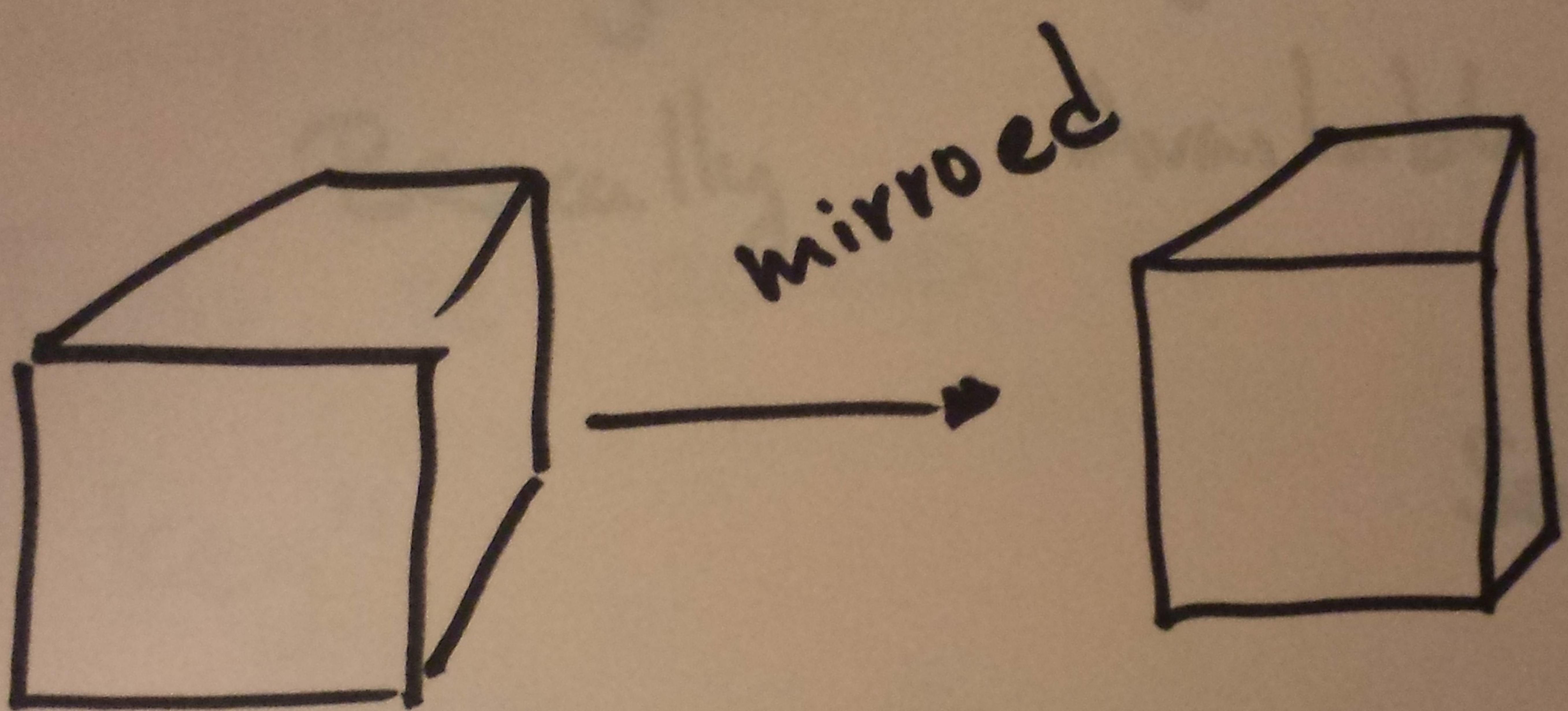
Partition-
Tolerance

Consistency

No Free Lunch!

Ejemplo:

Relational DB



A + PT !

C X

Proutchet
(2008)



B +

SE

Basically

Available

Soft State

ACID → Permissive

BASE → Optimistic

Eventually
Consistent

No relational

Key-value DB

- Big hash table
- 'Value' puede ser cualquier cosa
- Todo se accesa con el "key"
- Buen performance para escribir
- Fácil de replicar y distribuir
- NO HAY QUERIES

No relational

Document Store

- Guarda documentos
- Cada doc puede tener diferentes esquemas
- Permite buscar por "data" a diferencia de las key-value.

New SQL

Volt DB

MemSQL

In - Memory

Relational

Relational
y SQL

ACID

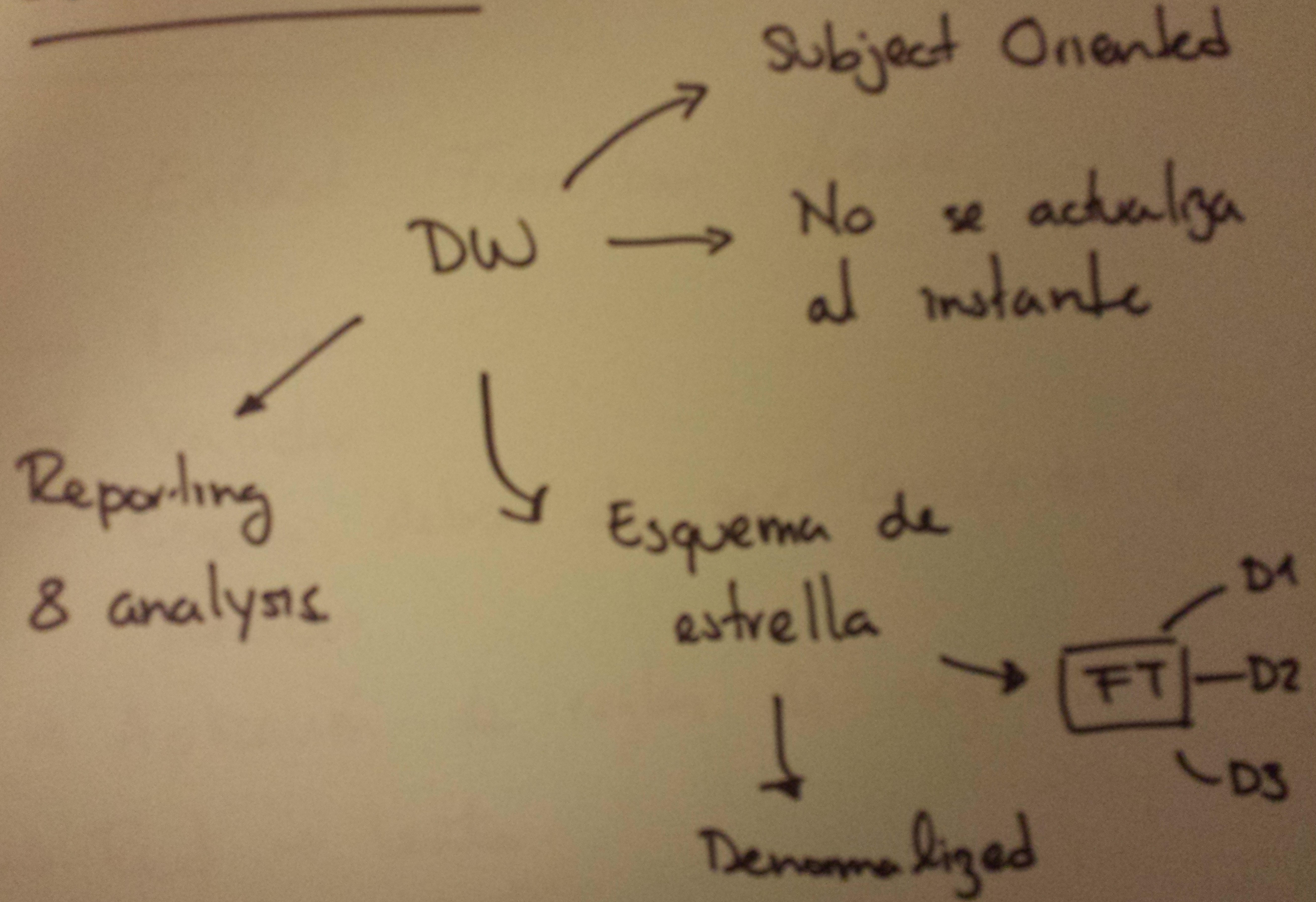
Compila a
C++

<http://kkowalc.eu/>

cassandra - vs - mongodb - vs

- couchdb - vs - redis

Data warehouse



ETL

Extract Transform Load

- Umpiar
- Validar
- Errores deben de ser corregidos o
borrados
- Aplicación de reglas de negocio
- Red → lenta
- Cambios ... UPS
- Several "Sources" → FAIL if one ...

What is Hadoop?