



Corso di ***Progettazione*** ***di sistemi multimediali***

prof. Pierluigi Feliciati

a.a.2011/12

Modulo 2

le risorse digitali aggregate:

Collezioni digitali, ipertesti, banche dati, sistemi informativi

Le risorse digitali aggregate

le risorse **devono essere associate in insiemi/collezioni che siano organizzati / associati, per avere la garanzia di essere correttamente contestualizzate .**

Semplificando, i sistemi di risorse (collezioni digitali) possono essere:

- **Associati tramite dati esterni** (metadati strutturali)
- **Basarsi su collegamenti ipertestuali** (siti e applicazioni web, che papprofondiremo più avanti - modulo 3)
- **Essere strutturati tramite Data Base**, quindi rispondere a una logica e sintassi definita, che può essere anche molto complessa e che comunque è la più rigida dei tre casi

Le risorse digitali aggregate

Collezioni digitali

Nei casi di **digitalizzazione** di materiali analogici, ma anche nella raccolta sistematica di risorse nate digitali si creano le cosiddette **collezioni digitali**, cioè aggregazioni ex novo di oggetti digitali associati sulla base di criteri diversi: associazione originaria (serie di oggetti simili, collezioni artistiche, pagine di un libro, cataloghi di prodotti etc.), oppure per **tema**, **cronologia**, **obiettivi formativi** o **didattici**, **esposizioni**, **dimostrazioni**, etc.

Le risorse digitali aggregate

Collezioni digitali

Si capirà fin da ora come siano importanti le informazioni su chi ha effettuato l'aggregazione, sulla base di che criteri, quando, per conto di chi, adottando che tecniche, etc.

Le collezioni meritano di essere valorizzate anche oltre l'effimero, specie se finalizzate ad esser **pubblicate sul web**. Devono insomma essere identificate, recuperate, scelte, indicizzate come collezioni, non solo attraverso le singole risorse che le compongono!

Le risorse digitali aggregate

Collezioni digitali

Parleremo più avanti delle tecniche di supporto per tali funzioni, ovvero l'associazione a ciascuna delle nostre risorse di **metadati descrittivi**.

Fin da ora è importante che si tenga conto che le collezioni sono prodotte, pubblicate, promosse, gestite, conservate come entità unitarie.

Quindi è necessario associare dei metadati descrittivi anche a questo **livello di granularità**.

Le risorse digitali aggregate

Descrizione delle collezioni

esistono attualmente:

lo schema per la descrizione delle collezioni del
Research Support Library Programme (RSLP)

<http://www.ukoln.ac.uk/metadata/rslp/>

il *Dublin Core Collection Description Application Profile*

<http://dublincore.org/groups/collections/collection-application-profile/>

Inoltre, il progetto MICHAEL ha costituito un inventario europeo di collezioni culturali e scientifiche digitali basato sugli standard Dublin Core e le specifiche MINERVA.

Cfr. <http://www.michael-culture.it/> e <http://www.michael-culture.eu/>

Le risorse digitali aggregate

Descrizione delle collezioni

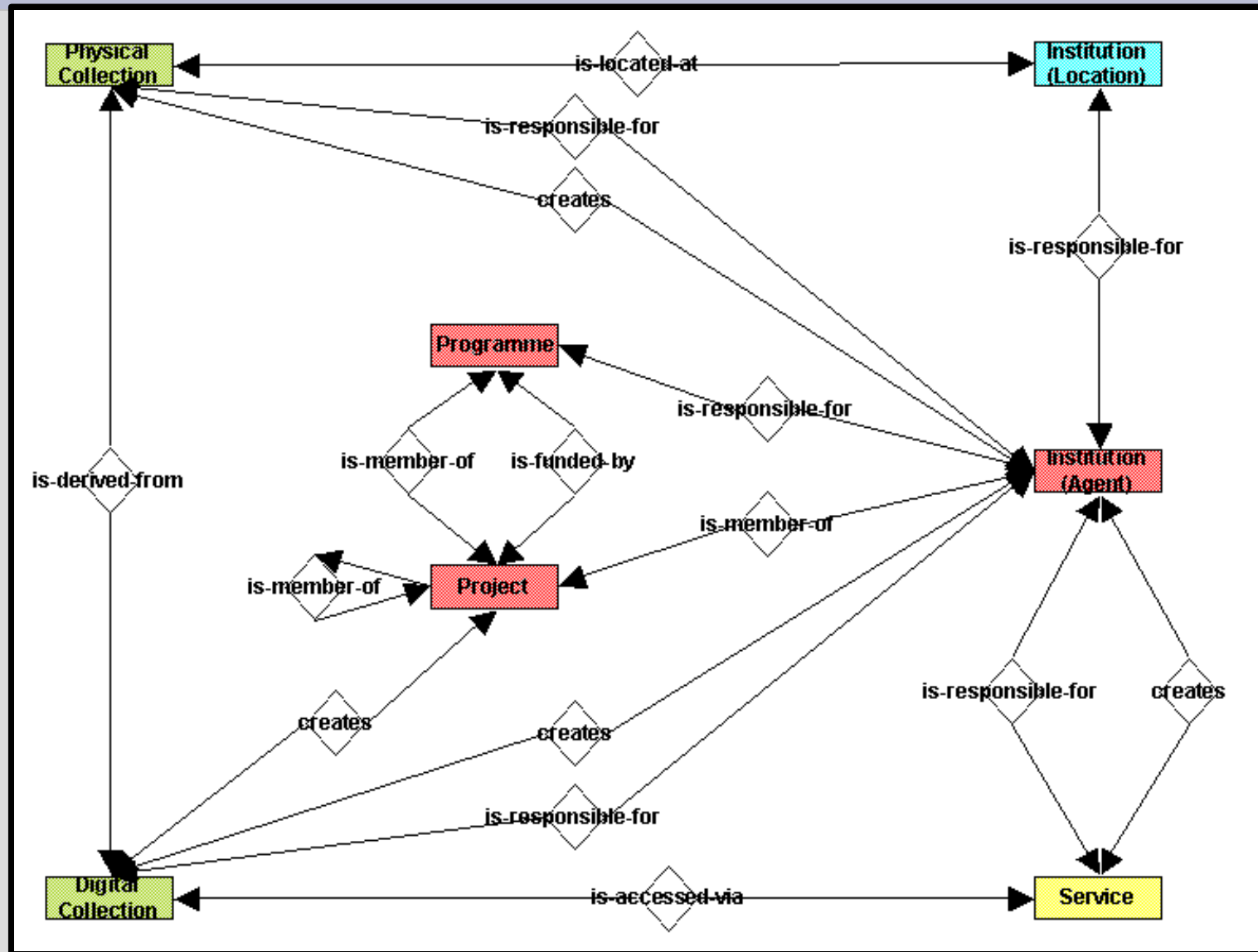
Per descrivere le collezioni sui MICHAEL, ad esempio, si raccolgono informazioni su 5 entità distinte, in relazione tra loro:

- Collezione digitale
- Eventuale collezione/i fisica
- Organizzazione/i coinvolt/e
- Progetto/i all'origine della collezione
- Servizio/i che mettono la collezione a disposizione degli utenti



La descrizione delle collezioni

(sistema relazionale)



Le risorse digitali aggregate: *metadati strutturali*

I libri hanno un dorso e un frontespizio, le opere d'arte in un museo hanno etichette e pannelli esplicativi, le medicine in una farmacia hanno informazioni sul principio attivo e codici che distinguono quelle per cui necessita ricetta da quelle da banco... ecco cosa sono i metadati.

Informazioni sui dati, che non fanno parte del loro contenuto ma svolgono l'essenziale ruolo di renderli riconoscibili, selezionabili e gestibili (se opportuno) nel tempo.

Le risorse digitali aggregate: *metadati strutturali*

I metadati strutturali descrivono le **relazioni logiche o fisiche che collegano le parti di un oggetto composito**.

Un libro “fisico”, ad esempio, fa parte delle offerte con sconto, e a sua volta può consistere in una sequenza di capitoli, che sono composti da pagine...

Il processo di digitalizzazione/aggregazione di risorse può generare un certo numero di risorse digitali distinte, ad esempio una immagine per ogni pagina, ma il fatto che queste risorse formino una sequenza e che quella sequenza costituisca un oggetto composito o faccia parte di una struttura complessa multilivellare è evidentemente essenziale per il loro uso e la loro interpretazione.

Le risorse digitali aggregate: *metadati strutturali*

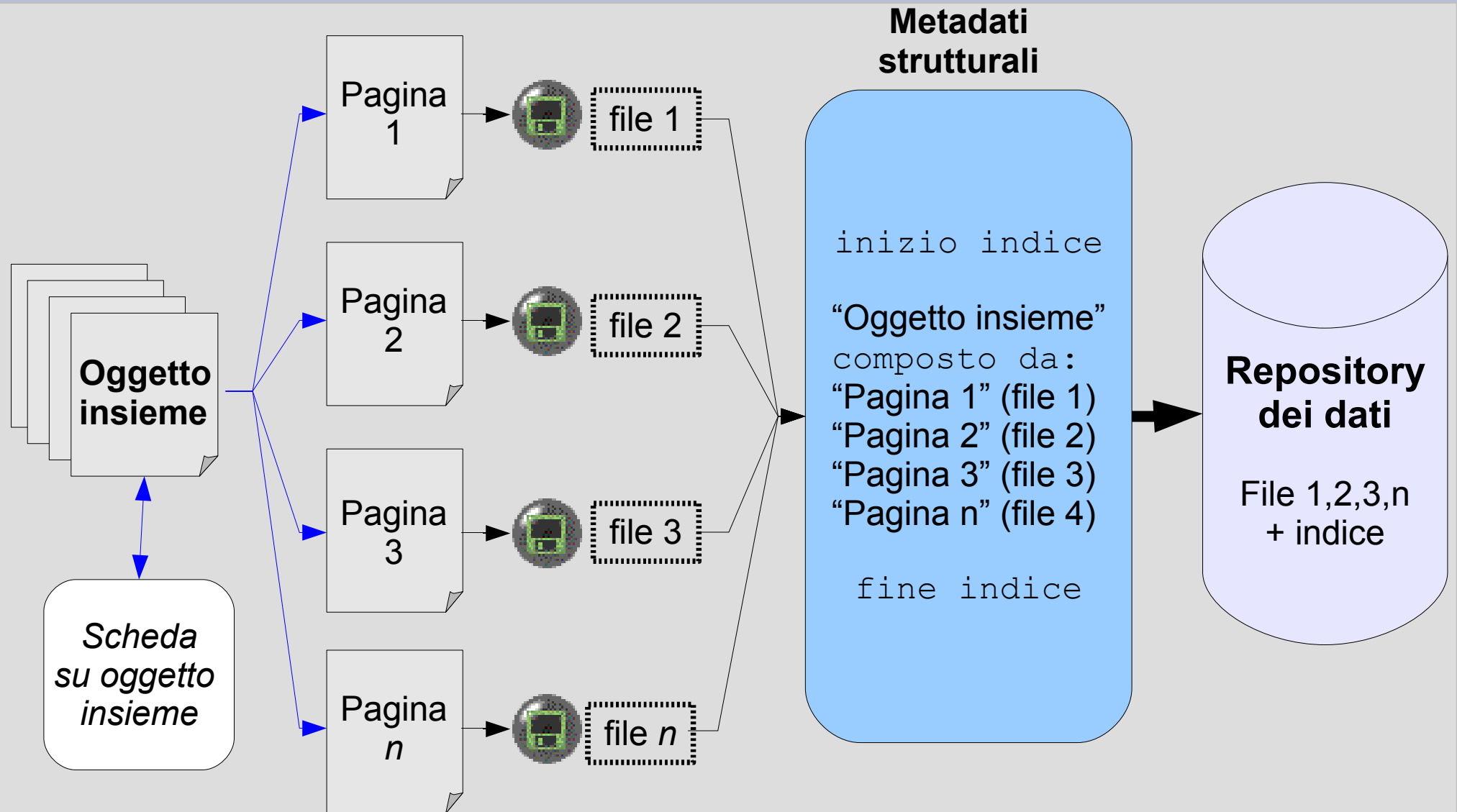
I metadati strutturali **collegano insomma le varie componenti delle risorse** per garantirne un'adeguata e completa fruizione, che spesso avviene attraverso la mappatura di schemi di metadati diversi.

Inoltre, forniscono dati di **identificazione e localizzazione** della risorsa, come il codice identificativo, l'indirizzo di ogni file sul server, l'archivio digitale di appartenenza e il suo indirizzo Internet.

Ne sono esempio:

1. XML (ne parleremo...)
2. i profili applicativi MAG e METS (ne parleremo...)

Metadati strutturali



Le risorse digitali aggregate: **ipertesto**

Con **ipertesto** intendo **una scrittura non sequenziale**.

La scrittura normale è sequenziale per due ragioni: è nata dal linguaggio e dall'oratoria, che devono essere sequenziali e i libri sono comodi da leggere solo in sequenza.

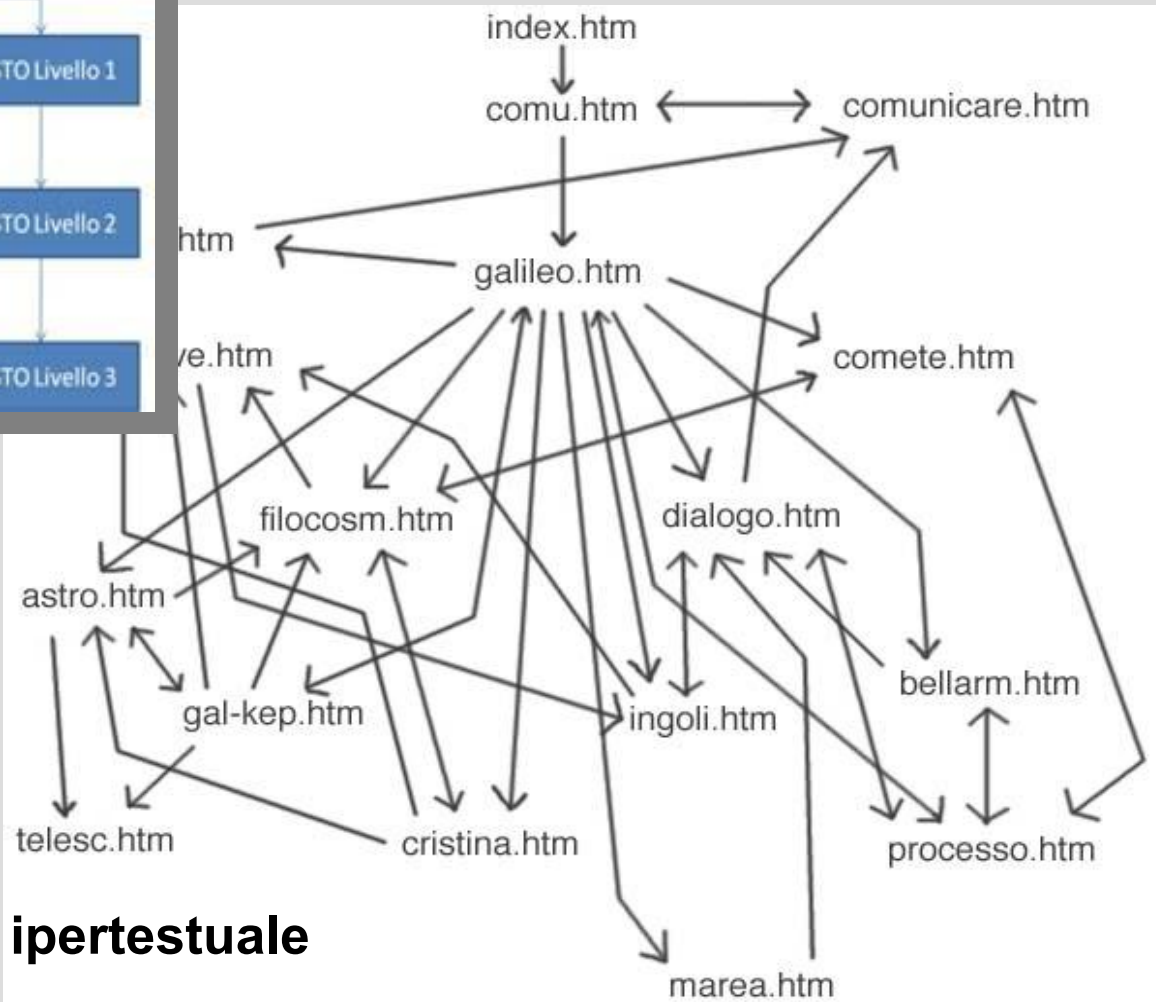
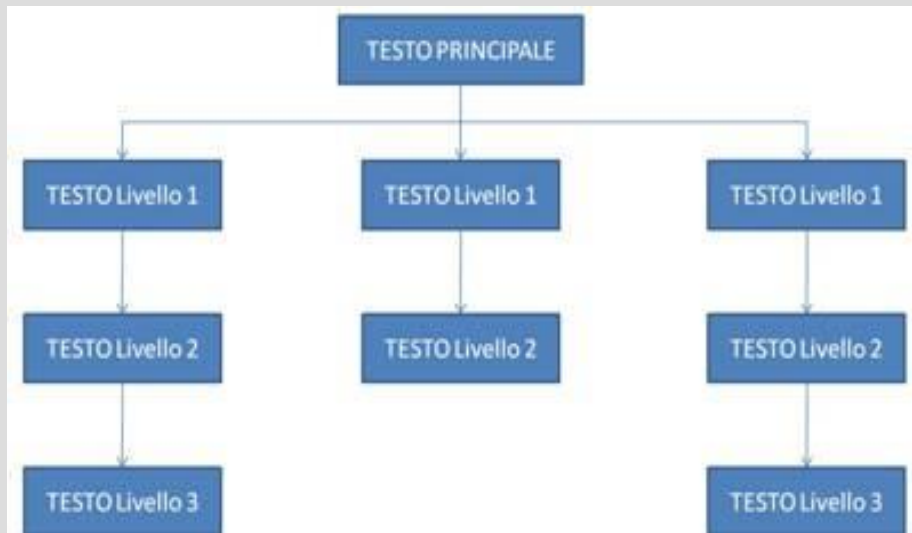
Ma **le strutture delle idee non sono sequenziali**. Hanno legami in ogni direzione, E quando scriviamo, cerchiamo sempre di creare legami non sequenziali. (...) Molti scrittori hanno tentato di liberarsi dalla sequenza.

Io penso che **stia sorgendo una nuova era**. La memoria del computer e lo schermo significano che non siamo più obbligati ad avere le cose in sequenza; sono possibili strutture del tutto arbitrarie e credo che una volta che le avremo provate abbastanza molti si accorgeranno di quanto siano auspicabili.

(Ted Nelson, *Dream machines*, 1974)

Le risorse digitali aggregate: **ipertesto**

Connessione gerarchica



Connessione ipertestuale

Le risorse digitali aggregate: **ipertesto**

- Di ipertesto - sia in teoria che applicato nella progettazione di un ambiente web - parleremo con dettaglio nel modulo 3
- Per ora vi basti sapere che le connessioni tra oggetti informativi in una rete ipertestuale sono **sempre unidirezionali** e che avvengono definendo al loro interno dei **punti/aree di ancoraggio**, cui si associano **informazioni sulla destinazione** della connessione (link): nome, posizione nel file system, etc.

Le risorse digitali aggregate:

data base

Un **database** è una **raccolta di dati permanenti**, gestita da un elaboratore e suddivisa in:

Metadati o schema: definizioni che ci dicono quali regole valgono sui dati, quali valori possono essere validi (vincoli di integrità), come i dati sono strutturati e collegati tra loro. Lo schema può cambiare nel tempo, non ha alcun vincolo con i programmi che accedono al database e deve essere definito prima dei dati.

Dati: rappresentazione (astrazione) dei fatti, ma può anche includere previsioni, supposizioni ed ipotesi conformi ai Metadati (o schema).

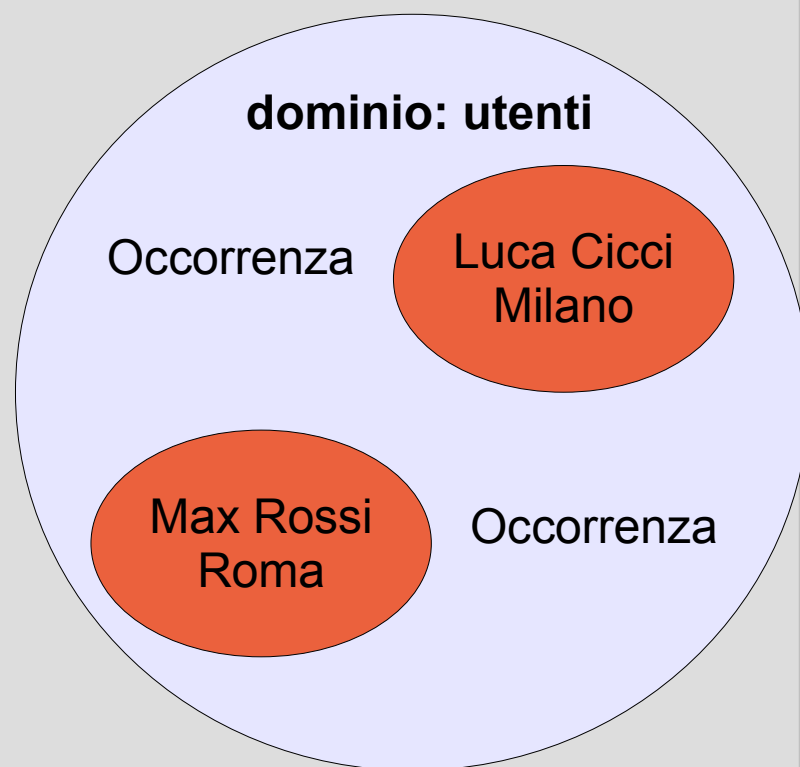
I dati sono organizzati in **insiemi omogenei** (stessa struttura) e sono collegabili con altri insiemi. Il **modello dei dati** è un insieme di astrazioni che descrivono le associazioni tra i dati.

Le risorse digitali aggregate:

data base

Una rappresentazione testuale di cosa significa informazione strutturata è quella di un file di testo che segue la sintassi XML (ne parleremo ancora):

```
<?xml version="1.0" encoding="UTF-8"?>
<utenti>
  <utente>
    <nome>Luca</nome>
    <cognome>Cicci</cognome>
    <indirizzo>Milano</indirizzo>
  </utente>
  <utente>
    <nome>Max</nome>
    <cognome>Rossi</cognome>
    <indirizzo>Roma</indirizzo>
  </utente>
</utenti>
```



Le risorse digitali aggregate

DBMS, Sql

Data Base Management System. Sistema centralizzato (programmi coordinati) o distribuito (rete) che permette di memorizzare, modificare ed estrarre informazioni da un database. **Garantisce l'integrità dei dati** (unica raccolta di dati anziché copie distinte scoordinate che potrebbero causare duplicazioni, ridondanze). **Organizza le informazioni** del database secondo la struttura di un database gerarchico, di un database di rete o di un database relazionale. **Garantisce l'accesso** alle informazioni solo a persone autorizzate, p.e. tramite una password. Un DBMS oltre a stabilire schemi organizzativi e di controllo, rende le informazioni accessibili agli utenti, tramite query. Esempi di DBMS: Oracle, Microsoft SQLServer, Basis, MySQL, PostgreSQL, InterBase.

Structured Query Language, linguaggio di interrogazione strutturato. SQL, oltre a creare tabelle bidimensionali permette la formulazione di query (interrogazioni) sulle relazioni (tabelle). SQL e' il linguaggio piu' utilizzato per definire database.

Le risorse digitali aggregate

Tipi di database

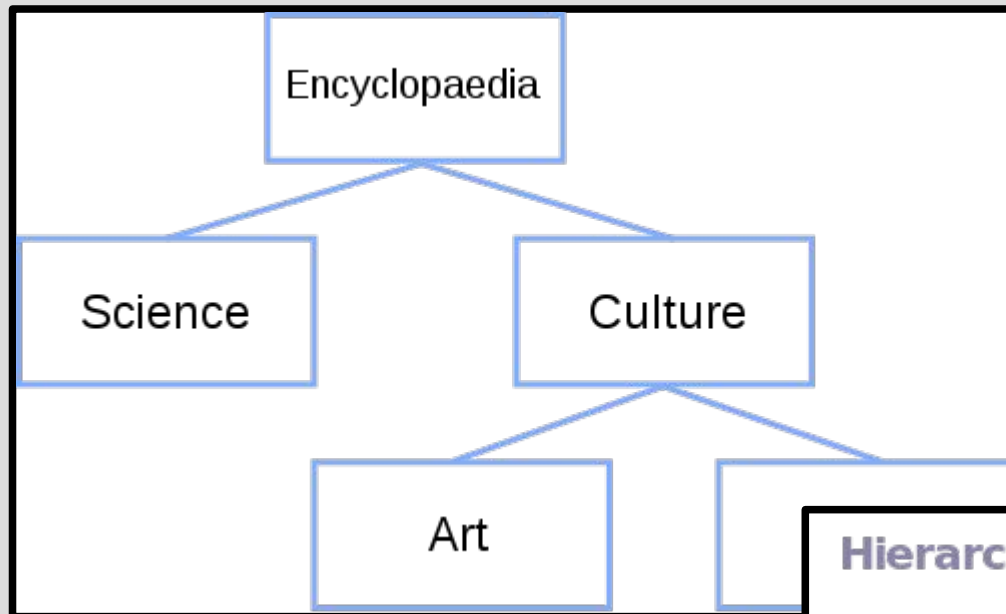
Le basi di dati possono avere **varie strutture**, tipicamente, in ordine cronologico:

1. **gerarchica** (rappresentabile tramite un albero - anni sessanta),
2. **reticolare** (rappresentabile tramite un grafo - anni sessanta),
3. **relazionale** (attualmente **il più diffuso**, rappresentabile mediante tabelle e relazioni tra esse - anni settanta),
4. **ad oggetti** (estensione alle basi di dati del paradigma "Object Oriented", tipico della programmazione a oggetti - anni ottanta),
5. **semantica** (rappresentabile con un grafo relazionale - inizio anni duemila).

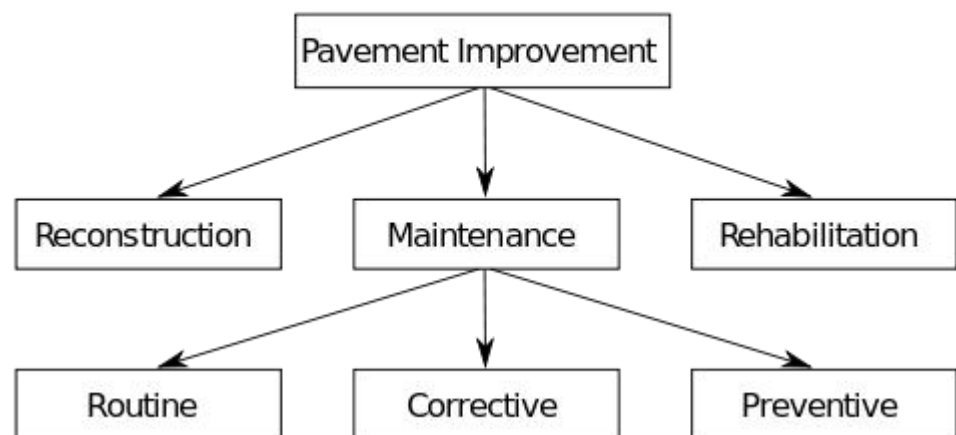
Un requisito importante di una buona base dati consiste nel *non duplicare inutilmente le informazioni in essa contenute*: questo è reso possibile dai gestori di database relazionali, che consentono di salvare i dati in tabelle che possono essere collegate.

Le risorse digitali aggregate

Database gerarchico



Hierarchical Model



Le risorse digitali aggregate

data base relazionale

L'assunto fondamentale del modello relazionale è che **tutti i dati sono rappresentati come relazioni**; i dati sono manipolati con gli operatori dell'algebra relazionale.

Il modello relazionale consente al progettista di database di creare una **rappresentazione consistente e logica dell'informazione**.

La **consistenza** viene ottenuta inserendo nel progetto del database appropriati **vincoli**, normalmente chiamati **schema logico**.

La struttura base del modello relazionale è il **dominio** o **tipo di dato**, definito come l'insieme dei valori che può assumere un determinato attributo.

Un **attributo** è una coppia ordinata di "nome di attributo" e "nome di tipo", mentre un valore di attributo è un valore specifico valido per quel tipo di dato. Una **relazione** consiste di una testata e di un corpo, dove la testata è un insieme di attributi e il corpo è un insieme di n valori possibili. La **tabella** è la rappresentazione grafica normalmente accettata per rappresentare la relazione.

Le risorse digitali aggregate

data base relazionale

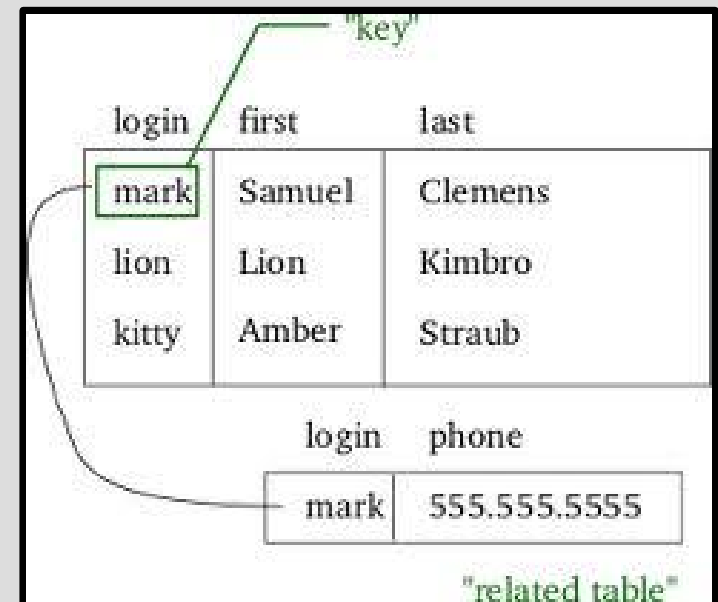
Relational Model

Activity Code	Activity Name
23	Patching
24	Overlay
25	Crack Sealing

Key = 24

Activity Code	Date	Route No.
24	01/12/01	I-95
24	02/08/01	I-66

Date	Activity Code	Route No.
01/12/01	24	I-95
01/15/01	23	I-495
02/08/01	24	I-66

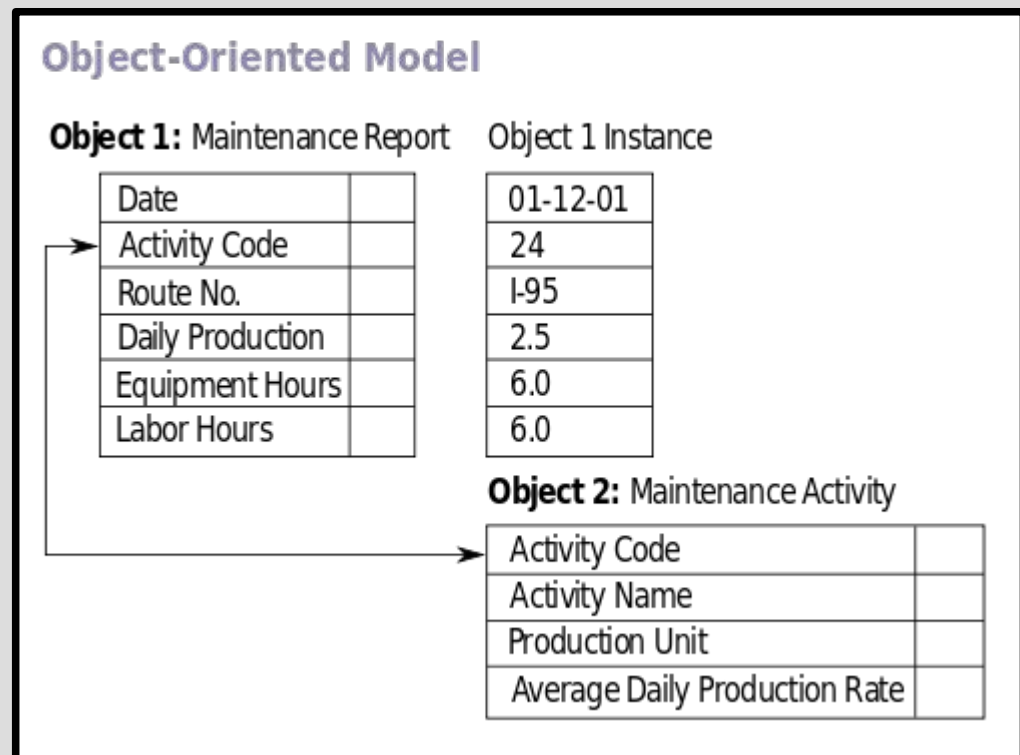


Le risorse digitali aggregate

data base ad oggetti

Le basi di dati a oggetti sono situati in un mercato di nicchia rispetto al modello relazionale (RDBMS). Vennero presi in considerazione sin dagli anni ottanta e novanta ma ebbero uno scarso impatto commerciale, e vennero usati solo in alcuni campi. I campi d'applicazione sono stati: ingegneria, basi di dati spaziali, telecomunicazioni, fisica delle alte energie e biologia molecolare.

i dati sono oggetti/istanze,
gli schemi sono separati,
tabelle/report



Le risorse digitali aggregate

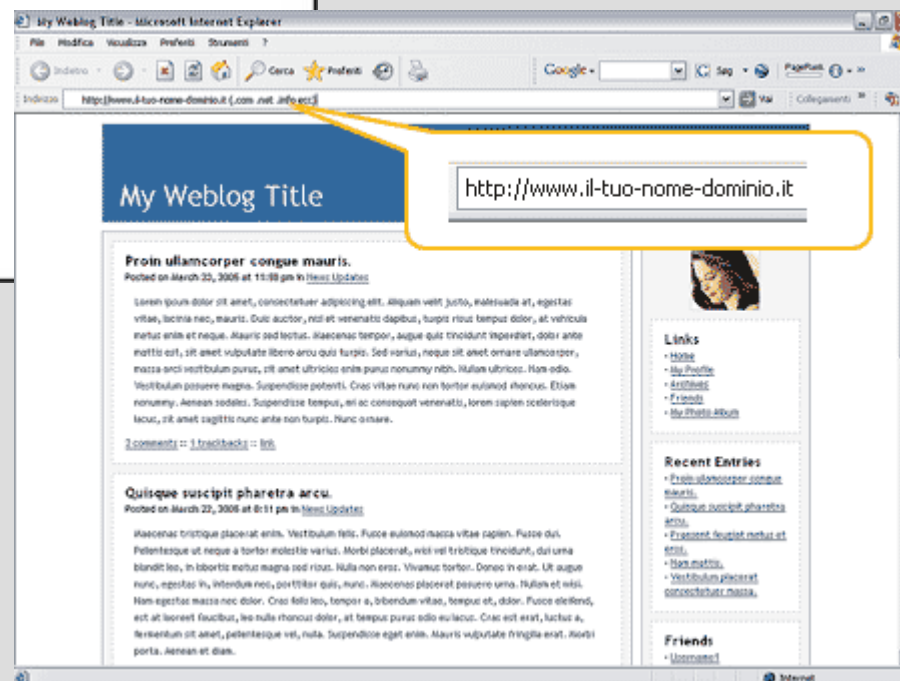
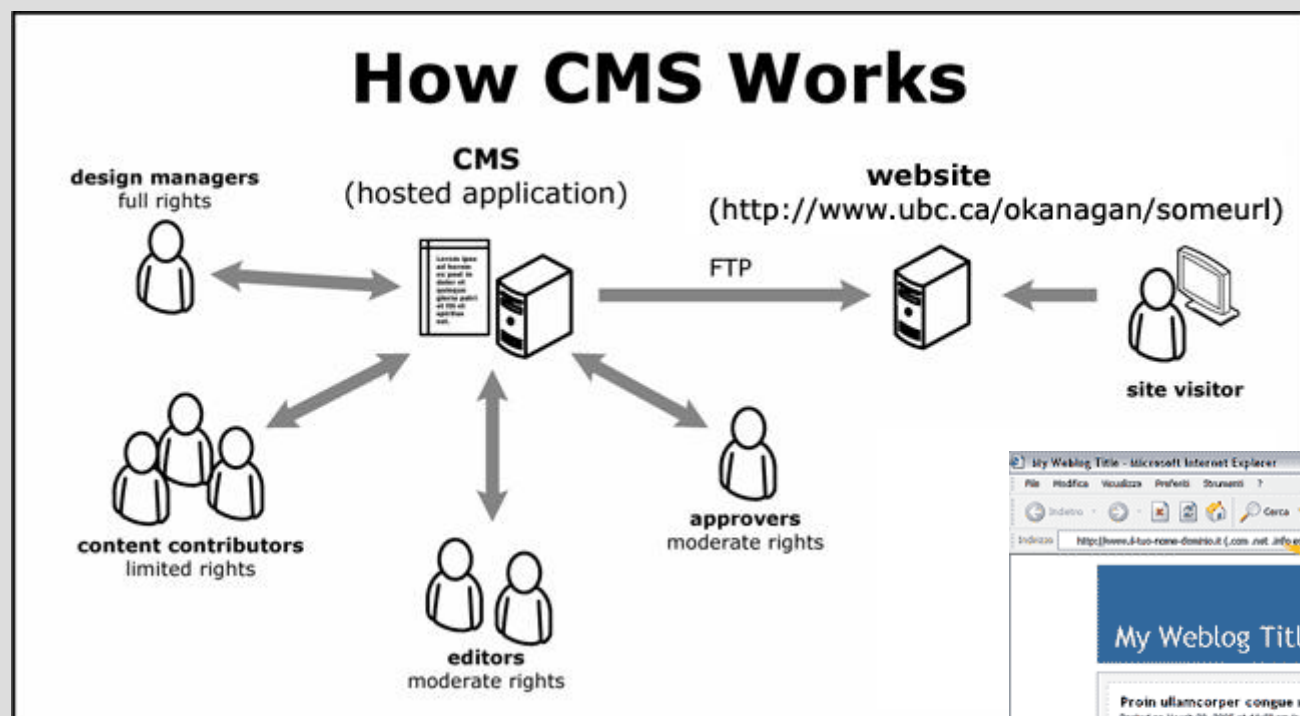
CMS, blog

Blog e forum sono particolari **Content Management System (CMS)**, implementati da **database** (realizzati ad esempio con DBMS come MySQL). I database vengono interrogati e gestiti generalmente attraverso istruzioni di programmazione. L'insieme dei file che contengono le istruzioni per operare con la base di dati e per impaginare gli interventi, costituisce il pacchetto di blogging installato nel server web.

Blog: contrazione di Web e Log, commenti (“tracce”) sul web. Particolare pagina web strutturata come un diario che raccoglie testi aperti ai commenti dei lettori. Gli utenti possono scrivere nelle pagine-blog pensieri, opinioni, poesie, esperienze, oppure inserire dei link su specifici argomenti, nonché immagini. Ogni nuovo intervento inviato dall'utente viene inserito in cima alla pagina web, riportandone la data e l'ora dell'inserimento. L'aspetto (layout) delle pagine-blog può essere modificato utilizzando specifici **modelli grafici (Template)**.

Le risorse digitali aggregate

CMS e blog



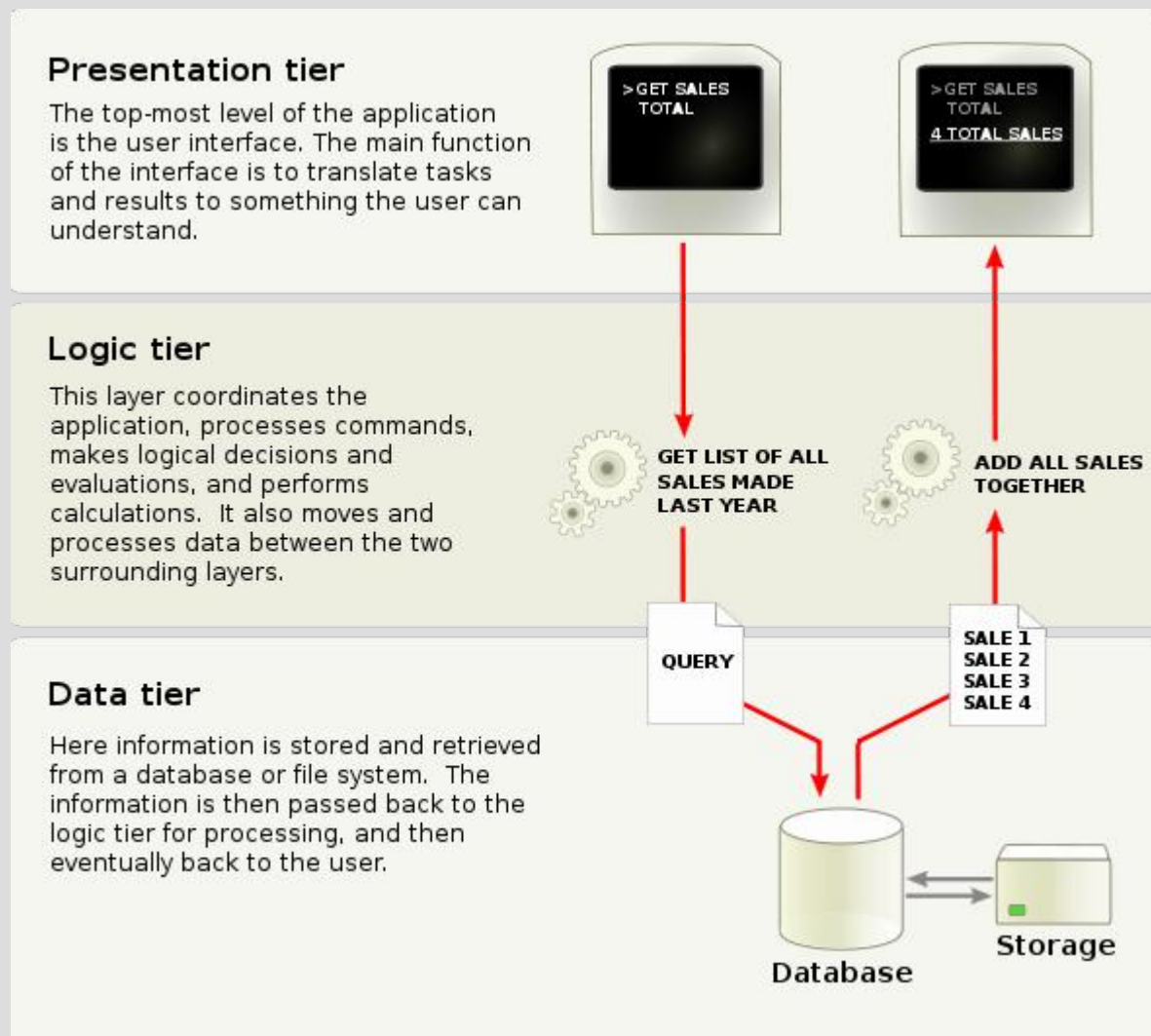
Le risorse digitali aggregate

In genere , per le **banche dati** complesse e in rete, si adottano dei **modelli di architettura**, sempre basati sulla distinzione dei ruoli tra client e server, detti multi-tier, ovvero multi-livello.

Ad esempio, se si distinguono come tre livelli distinti la **presentazione** (es. carrello, pagina di risultati di ricerca..), **livello applicativo** (*business logic*, ovvero di elaborazione dei dati) e **livello dati** (il DBMS vero e proprio) si parla di **architettura 3-tier**.

Uno dei vantaggi è consentire a qualsiasi dei tre livelli di **essere aggiornati o sostituiti** indipendentemente l'uno dall'altro.

Le risorse digitali aggregate



Le risorse digitali aggregate

Detto questo, possiamo distinguere, semplificando e solo a scopo didattico, **due tipi sostanziali di banche dati**, rispetto alla organizzazione delle macchine su cui sono conservati i dati e dove si trova il software gestionale:

- con **dati solo organici** al DBMS
- con **dati indipendenti** dal DBMS

In generale, un **database distribuito** è un database che si trova sotto il controllo di un DBMS nel quale gli archivi *non sono memorizzati sullo stesso computer*.

Il database può essere dislocato in più computer situati nello stesso luogo, oppure distribuito in una rete di computer connessi tra loro e basarsi anche su più di un DBMS.

Le risorse digitali aggregate

- **con dati organici al DBMS**

in questo caso la struttura, il formato, le maschere di inserimento, le query e le modalità di aggregazione, ordinamento e output (anche Web) sono interamente gestite dal DBMS.

Questo se semplifica la fase di progettazione iniziale (“chiavi in mano”) può creare delle limitazioni e difficoltà per la portabilità e l’interoperabilità dei dati.

Le risorse digitali aggregate

- **con dati indipendenti dal DBMS**

in questo caso

- la struttura dei dati sono basati su formati leggeri e standard (XML prima di tutto)
- le maschere di inserimento possono essere costruite con software di editing anche molto semplici e gratuiti
- le query e le modalità di aggregazione, ordinamento, sicurezza, conservazione sono gestite dal DBMS
- l'output e l'interoperabilità possono essere indipendenti dal DBMS, "tornando" ai dati originari

Le risorse digitali aggregate

i sistemi informativi

attualmente si cerca di avviare, nell'approccio alle risorse digitali, progetti che comprendano in un unico package le funzioni di descrizione, digitalizzazione, aggregazione, amministrazione, gestione, conservazione e comunicazione: i SI

“Il sistema informativo è costituito dall'insieme delle informazioni utilizzate, prodotte e trasformate da un'azienda durante l'esecuzione dei processi aziendali, dalle modalità in cui esse sono gestite e dalle risorse, sia umane, sia tecnologiche, coinvolte.”
(Wikipedia)

Le risorse digitali aggregate

sistemi informativi e sistemi informatici

Nelle aziende il sistema informativo ha il compito di:

- **raccogliere** i dati;
- **conservare** i dati raccolti, archiviandoli;
- **elaborare** i dati, trasformandoli in informazioni;
- **distribuire** l'informazione.

Per fare questo il sistema informativo si avvale di tecnologie informatiche e prende anche il nome di **sistema informatico**.

Almeno a livello concettuale, però, il sistema informativo non implica di per sé l'uso dell'informatica (pensate ai registri di entrata/uscita, ai protocolli...)

Le risorse digitali aggregate

Se dal punto di vista dei formati e dei metadati valgono tutte le cose dette finora, vale la pena di accennare qualcosa sulle possibili **architetture dei sistemi informativi**.

Sinteticamente, questi possono essere **distribuiti o centralizzati**, ma i modelli sono in realtà ben più complessi, anche in relazione alla “dimensione”, cioè alla complessità e quantità dei dati e al contesto organizzativo.

Le risorse digitali aggregate

i **sistemi centralizzati** raccolgono i database (anche se magari non le funzioni gestionali e quelle di output Web) in un sistema centrale, alimentato tramite connessione di rete da n terminali di immissione periferici.

Non cambia molto se questa immissione avviene tramite poli intermedi, anche strutturati a più livelli.

Le risorse digitali aggregate

i pro e i contro dei **sistemi centralizzati**

- **i pro**: facilitano la omogeneità formale e la validazione dei dati e la creazione di liste d'autorità, evitando ripetizioni e sovrapposizioni (pensiamo al modello SBN)
- **i contro**: l'immissione dei dati è spesso lenta e faticosa, e non è sempre agevole valorizzare i contesti locali

Le risorse digitali aggregate

i sistemi distribuiti

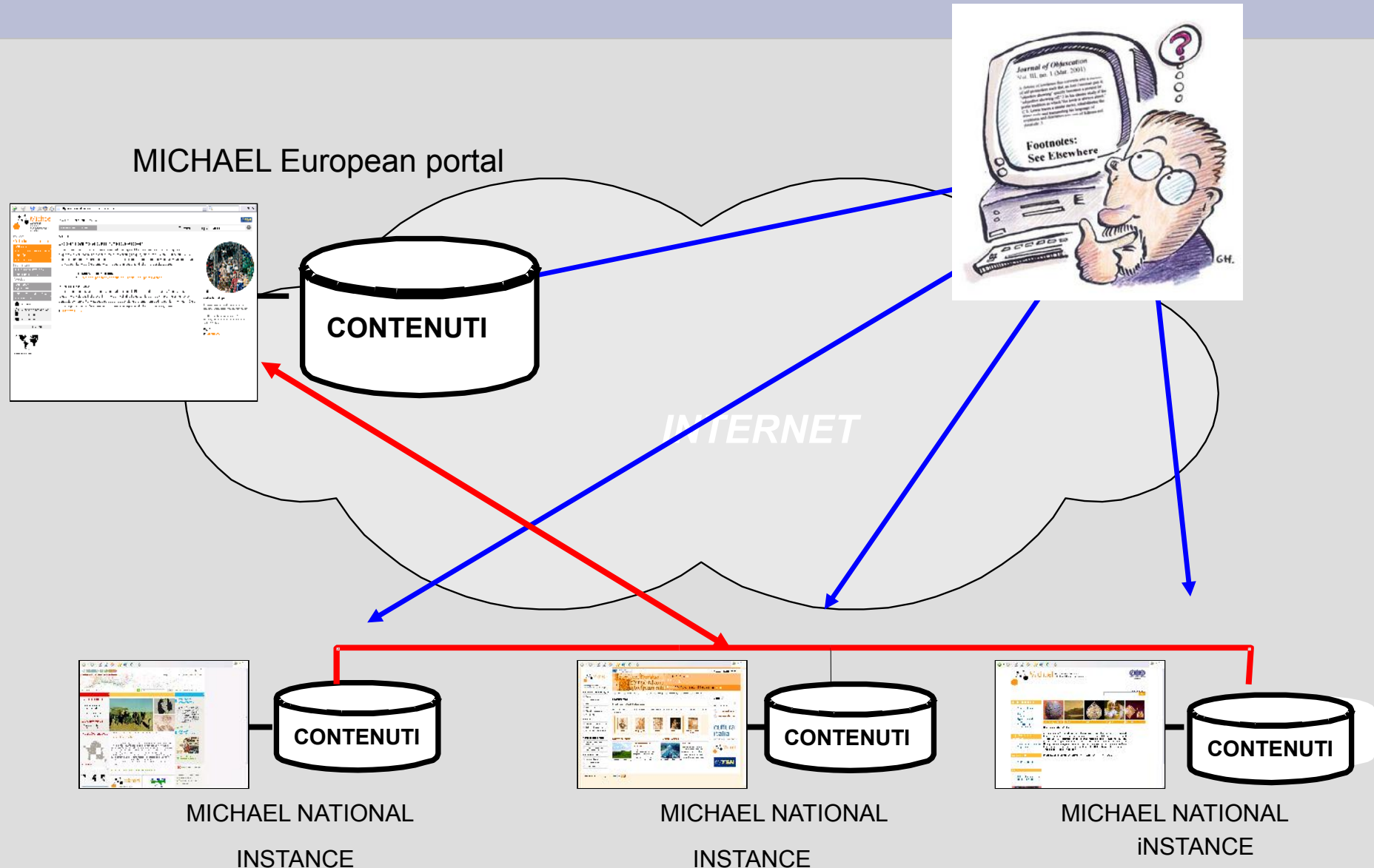
si basano sulla costruzione di n sistemi locali autonomi, che riversano tutti o parte dei loro dati su un sistema centrale deputato a funzioni particolari, tra cui quella Web. I sistemi locali sono autonomi ma nel loro rapporto col “centro” sono allineati attraverso procedure automatiche (ad es. OASIS) oppure no.

Le risorse digitali aggregate

i pro e i contro dei **sistemi distribuiti**

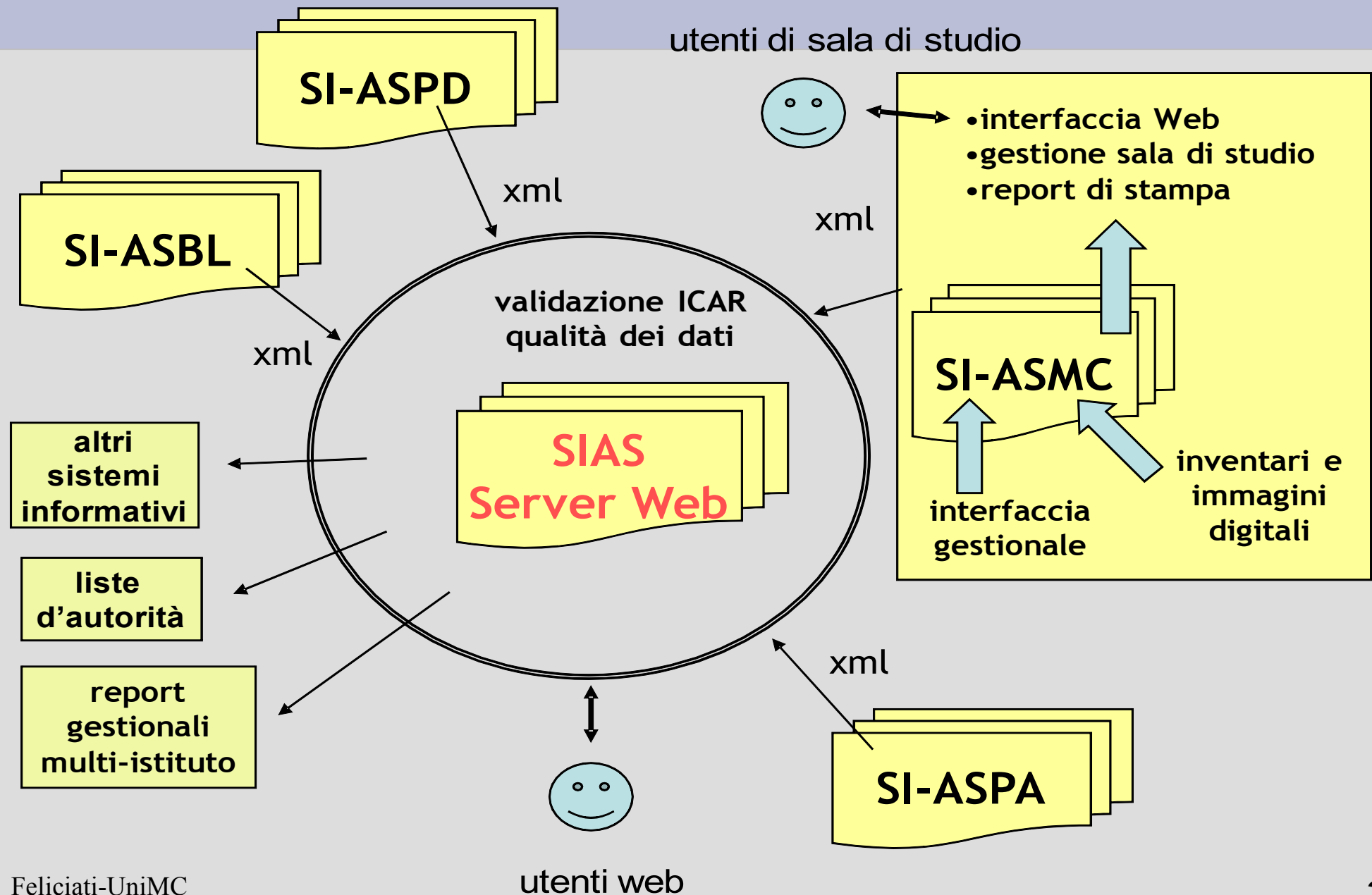
- **i pro**: valorizzano i contesti locali permettendo personalizzazioni e applicazioni specifiche; velocizzano immissione e gestione dei dati
- **i contro**: le procedure di omogeneizzazione e validazione dei dati sono più delicate

Esempi di Sistemi distribuiti (18 +1 portali MICHAEL)



Esempi di sistemi distribuiti

(Sistema Informativo Archivi di Stato)



Esempi di Sistemi distribuiti

(Il sistema pubblico di cooperazione e connettività)

