



**UNIVERSITÀ
DEGLI STUDI
DI UDINE**
hic sunt futura

TECNOLOGIE WEB E LABORATORIO (a.a. 2022-2023)

MODELLO DI UN SITO WEB

Daniele Salvati

Dipartimento di Scienze Matematiche, Informatiche e Fisiche
Università degli Studi di Udine

Informazioni slide

- Il materiale contenuto in queste slide è riservato esclusivamente agli studenti del corso di **Tecnologie Web e Laboratorio** del Corso di Studio in **Internet of Things, Big Data, Machine Learning** dell'Università degli Studi di Udine.
- Non è consentita la diffusione del materiale contenuto in queste slide, ma solo l'utilizzo inerente la preparazione dell'esame del suddetto corso.

Sviluppo di un sito web

Componenti di un sito web

- Da un punto di vista astratto un sito web può essere considerato come costituito da tre componenti principali:
 - **Contenuto:** sono le risorse (es. informazioni, servizi) che il sito mette a disposizione degli utenti.
 - **Contenitore:** è la struttura atta a ospitare i contenuti che si intendono veicolare attraverso il sito e che conferisce ad essi una specifica organizzazione, una determinata forma di presentazione e allo stesso tempo definisce le regole di accesso e di interazione.
 - **Infrastruttura tecnica:** è l'ambiente tecnologico (piattaforma hardware e software) che ospita contenuto e contenitore.

Tipi di contenuto

- **Contenuto informativo:** il loro scopo è quello di **fornire informazioni** utilizzando in generale risorse multimediali. I siti web che offrono prevalentemente questo tipo di contenuti sono detti **siti informativi**.
- Esempi:
 - giornale on-line
 - sito istituzionale
 - previsioni del tempo
 - presentazione di una azienda e dei suoi prodotti

Tipi di contenuto (2)

- **Servizi applicativi:** sono applicazioni software specifiche messe a disposizione degli utenti mediante una interfaccia web che permettono di **espletare una procedura di lavoro** e di **eseguire un compito complesso** interagendo con il sito secondo un flusso operativo ordinato. I siti web che offrono prevalentemente questo tipo di contenuti sono detti **siti di servizio**.
- Esempi:
 - prenotare un posto a teatro
 - acquistare un biglietto aereo
 - espletare una procedura burocratica
 - acquistare un prodotto

Tipi di contenuto (3)

- **Servizi di comunità:** hanno lo scopo di permettere agli utenti di instaurare **relazioni sociali complesse** con altri utenti utilizzando il sito come tramite. I servizi di comunità riguardano tre aree: comunicazione (sincrona o asincrona), la condivisione di risorse, la collaborazione. I siti web che offrono prevalentemente questo tipo di contenuti sono detti **siti di comunità**.
- Esempi:
 - chat
 - forum
 - sistemi wiki
 - sistemi di notifica come i feed RSS

Tipi di contenuto (4)

- **Servizi di supporto:** mettono a disposizione degli utenti **servizi generali d'utilità**. Spesso i servizi di supporto sono forniti da applicazioni esterne, già disponibili nel momento in cui il sito web è progettato e realizzato.
- Esempi:
 - motori di ricerca
 - strumenti di registrazione
 - strumenti di personalizzazione
 - strumenti per l'accessibilità
 - ecc.

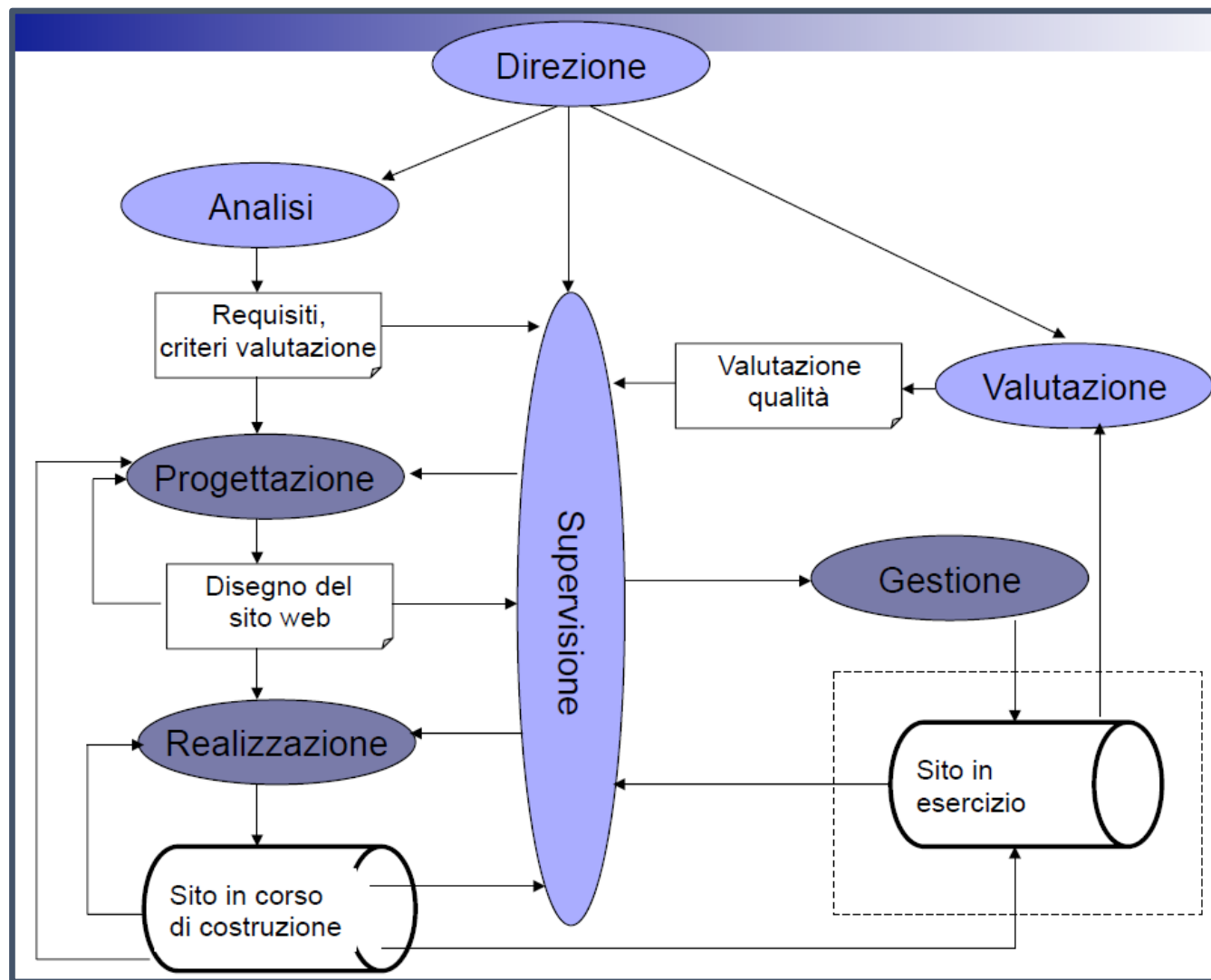
Ciclo di vita di un sito web

- Nella vita di un sito web si possono identificare due fasi principali:
 - **Costruzione del sito web**: il sito è progettato e realizzato.
 - **Vita operativa del sito web**: il sito è usato e aggiornato/modificato.
- Queste due fasi sono separate dal rilascio in esercizio o pubblicazione del sito web.
- Il **ciclo di vita di un sito web** è una rappresentazione dell'insieme di attività che ne caratterizzano la fase di costruzione e di vita operativa e della loro distribuzione nel tempo.

Funzioni del modello del ciclo di vita

- Le funzioni necessarie per costruire un sito web e per gestirne la vita operativa sono sette e sono classificate in due categorie principali:
 - **Funzioni di governo**
 - direzione
 - analisi
 - supervisione
 - valutazione
 - **Funzioni esecutive**
 - progettazione
 - realizzazione
 - gestione

Funzioni del modello del ciclo di vita (2)



Modello di progettazione

- La fasi principali del modello di progettazione sono:
 - **Fase-1:** analisi e specificazione
 - **Fase-2:** progettazione del **contenuto**
 - **Fase-3:** progettazione della **struttura ipertestuale**
 - **Fase-4:** progettazione della **presentazione**
 - **Fase-5:** realizzazione, collaudo e pubblicazione
 - **Fase-6:** gestione e valutazione
- Nota: le fasi non sono lineari. La progettazione è iterativa, incrementale, non-monotona (le decisioni prese possono essere riviste e modificate!). Specifiche e soluzioni co-evolvono durante la progettazione. Si realizzano prototipi intermedi.

Fase-1: analisi e specificazione

- **Categorie e profilo degli utenti:** comprendere in maniera profonda le caratteristiche e i bisogni/aspettative degli utenti.
- Attività:
 - definizione del target
 - analisi dei bisogni degli utenti
 - analisi delle aspettative degli utenti
- **Lista degli scopi e dei vincoli:** comprendere le intenzioni, gli scopi, e gli obiettivi del committente e i vincoli (le condizioni) da soddisfare.
- Attività:
 - definizione degli scopi
 - definizione dei vincoli di progetto

Fase-1: analisi e specificazione (2)

- **Risultati della analisi competitiva:** identificare, confrontare e valutare siti concorrenti
- Attività:
 - identificazione siti concorrenti
 - comparazione modelli
 - analisi competitiva
- **Personaggi e scenari d'uso:** elaborazione dei profili degli utenti per renderli strumenti di progetto e identificazione dei vincoli e dei requisiti d'uso (per ogni categoria d'utenza). I personaggi sono archetipi ipotetici di veri utenti.
- Attività:
 - definizione dei personaggi
 - definizione degli scenari d'uso

Fase-1: analisi e specificazione (3)

- **Documento di specifica:**
 - definizione dei requisiti funzionali e non funzionali del sito web
 - definizione dei criteri di valutazione e accettazione del prodotto finale
- **Specificazione:**
 - obiettivi del sito web
 - profilo degli utenti a cui esso si rivolge
 - insieme di scenari d'uso tipici (contesto applicativo)
 - posizionamento competitivo del sito rispetto ai siti concorrenti (tratti distintivi e i suoi punti di forza)
 - caratteristiche di massima dell'infrastruttura tecnologica

Fase-2: progettazione del contenuto

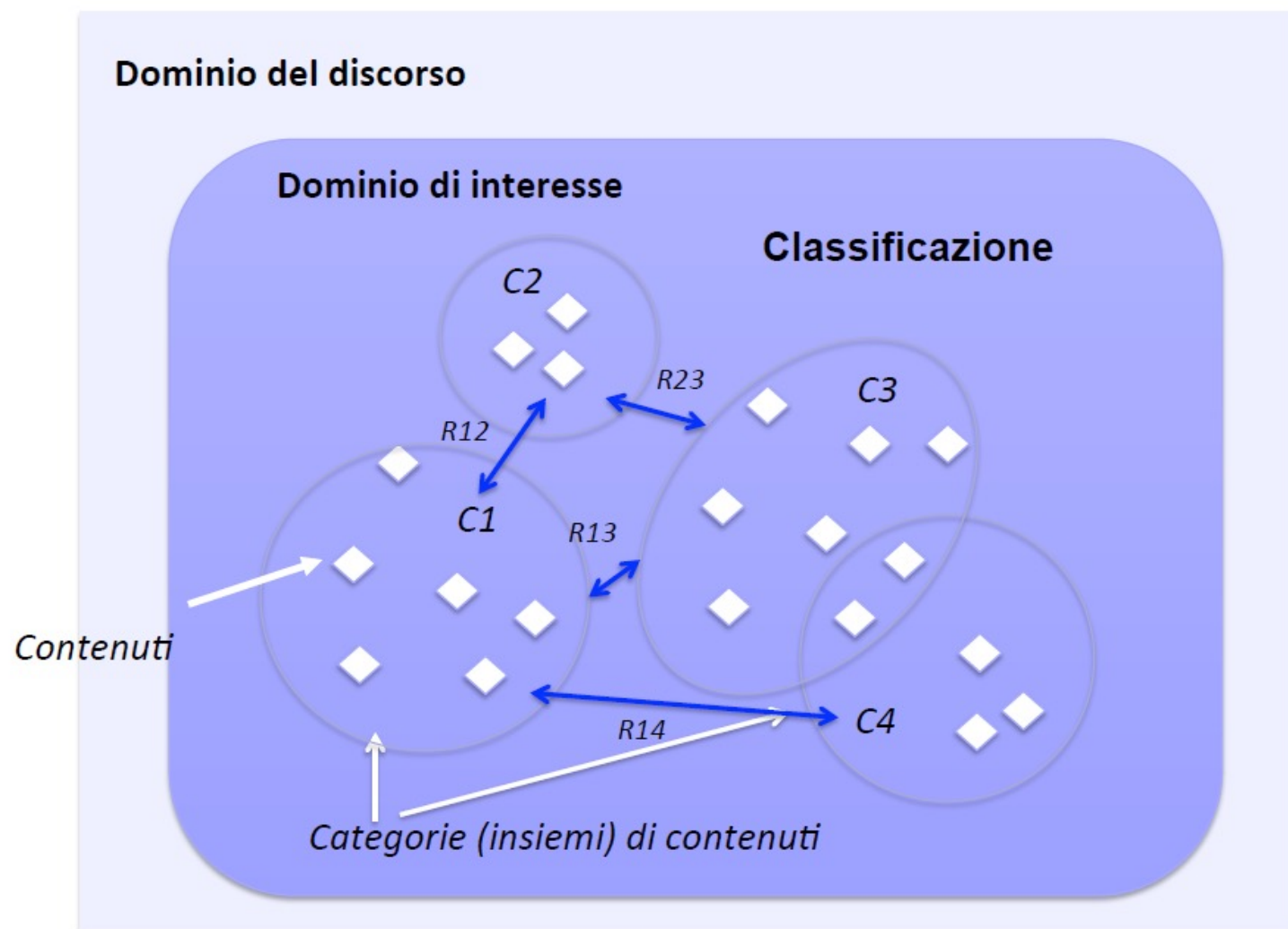
- **Lista dei contenuti e delle funzioni/servizi:** identificare l'ambito di interesse del sito web e i contenuti informativi e operativi (funzioni, servizi), nel rispetto dell'ambito applicativo e degli scopi, obiettivi e vincoli del progetto.
- Attività:
 - definizione del dominio del discorso
 - definizione del dominio delle unità di contenuto
- **Modello del contenuto:** classificare i contenuti in categorie, e identificare i legami logici e semantici tra categorie.
- Attività:
 - definizione della struttura logica
 - definizione della struttura semantica

Fase-2: progettazione del contenuto (2)

- **Criteri di classificazione del contenuto:**

- argomento, soggetto, tema
- attività
- ruolo o funzione
- tipologia d'utenza
- collocazione spaziale
- collocazione temporale
- uso
- ecc.

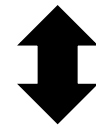
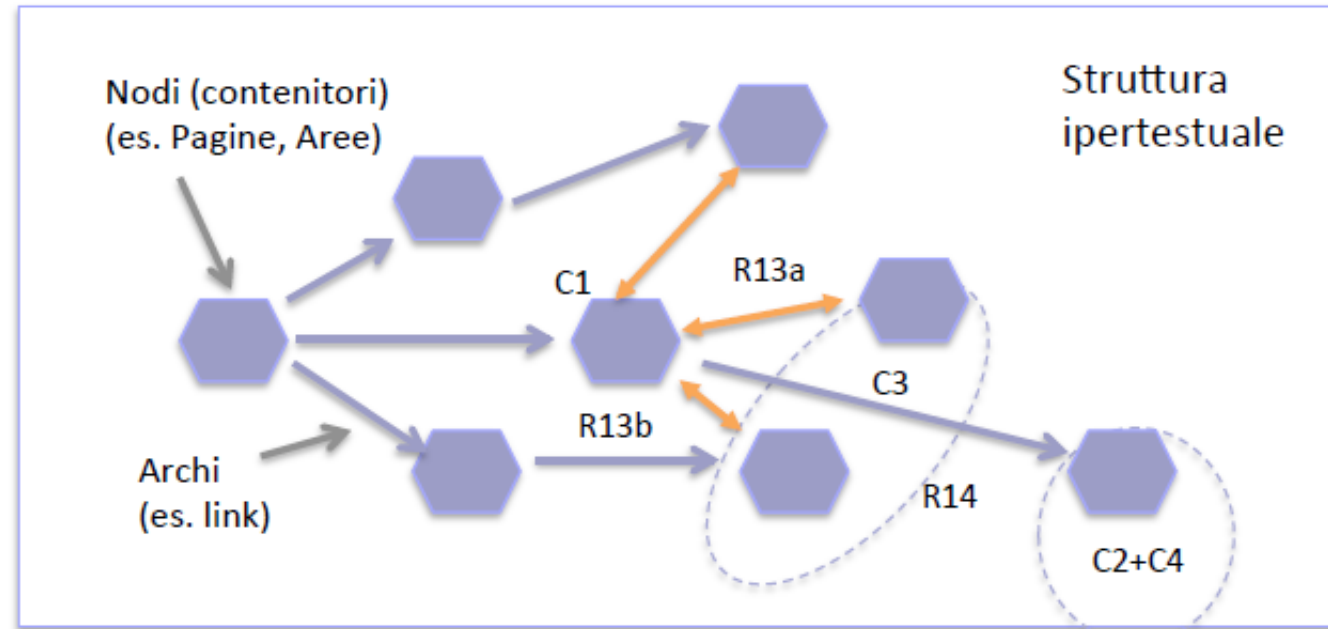
Fase-2: progettazione del contenuto (3)



Fase-3: progettazione della struttura ipertestuale

- **Struttura ipertestuale del sito web:** definire collegamenti e percorsi attraverso le unità di contenuto e prefigurare l'ordine di visita dei contenuti (informativi e operativi) da parte degli utenti.
- Attività:
 - definizione dei possibili percorsi di navigazione
- **Lista dei meccanismi di supporto alla navigazione e dei controlli:** definire quali informazioni sono fornite per supportare l'interazione e l'orientamento nello spazio informativo e operativo del sito (es. etichette, menù di navigazione, mappe, pulsanti, ecc.)
- Attività:
 - definizione dei sistemi di interazione (controlli)
 - definizione dei sistemi di orientamento (segnali)

Fase-3: progettazione della struttura ipertestuale (2)



dominio del discorso

Fase-4: progettazione della presentazione

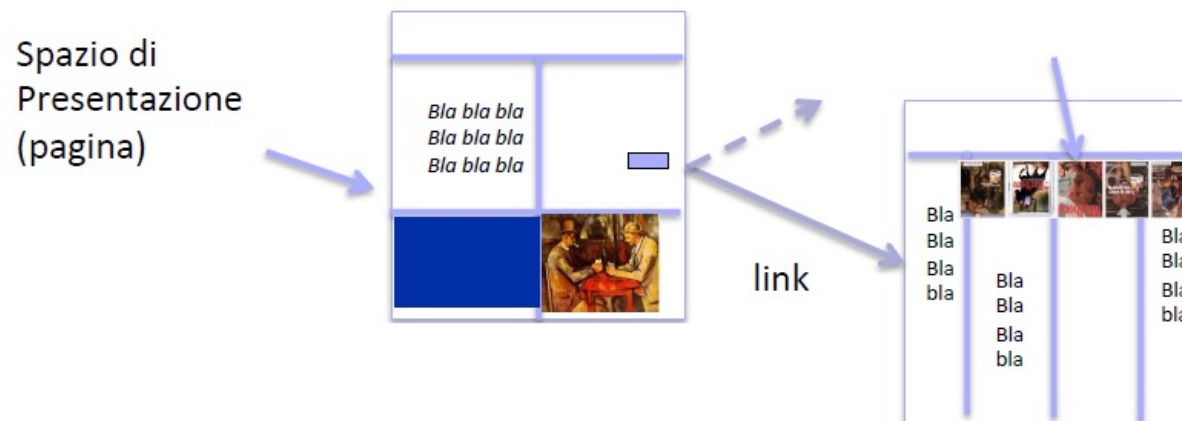
- **Caratteristiche dei dispositivi di erogazione e delle unità di presentazione:** individuare i dispositivi utilizzati per la fruizione del sito (PC desktop, tablet, cellulare, ecc.) e definire le unità di presentazione (es. finestre).
- Attività
 - scelta dei dispositivi di erogazione
 - scelta dell'unità di presentazione
- **Componenti multimediali:** rappresentare ciascun contenuto con il medium più appropriato rispetto alle specifiche di progetto.
- Attività:
 - scelta delle modalità di presentazione dei contenuti

Fase-4: progettazione della presentazione (2)

- **Struttura compositiva delle pagine:** decidere come strutturare lo spazio di presentazione (finestra, pagina) e quali componenti multimediali inserire e in quali posizioni.
- Attività:
 - definizione della struttura grafica (layout)
 - definizione della struttura compositiva delle pagine
- **Struttura fisica della interazione e della navigazione:** decidere come realizzare fisicamente i collegamenti ipertestuali e i meccanismi di supporto all'interazione e all'orientamento.
- Attività:
 - scelta delle modalità di presentazione del sistema di navigazione

Fase-4: progettazione della presentazione (3)

- **Caratteristiche sinestetiche delle pagine:** decidere le caratteristiche tipografiche dei testi, degli elementi grafici (filetti, cornici) usati per delimitare l'informazione, i colori e le combinazioni di colori da usare, e altre qualità visive, auditive, tattili e cinetiche.
- Attività:
 - scelta delle caratteristiche stilistiche ed espressive del sito



Fase-5: realizzazione, collaudo e pubblicazione

- **Realizzazione di un prototipo del sito web:** scegliere le tecnologie più appropriate per implementare le decisioni di progetto e realizzare il progetto con le tecnologie scelte (linguaggi di programmazione).
- Attività:
 - **realizzazione del progetto**
- **Esito delle prove di funzionamento e dei test:** garantire la piena corrispondenza del prodotto con i requisiti definiti in fase di analisi. Il sito web deve essere in grado di soddisfare i bisogni degli utenti, deve permettere di raggiungere gli obiettivi del committente e soddisfare i requisiti di qualità in uso.
- Attività:
 - **collaudo**

Fase-5: realizzazione, collaudo e pubblicazione (2)

- **Piano di pubblicazione:** capire quando pubblicare il sito, con quale procedura (es. interamente o attraverso rilasci successivi) e come farlo (es. a sorpresa o tramite annunci, anticipazioni o eventi).
- Attività:
 - redazione del piano di pubblicazione del sito web



Fase-6: gestione e valutazione

- **Modifiche al progetto originale o alla implementazione:** eseguire gli aggiornamenti resi necessari dall'evoluzione dei contenuti (es. informativi) e/o delle tecnologie.
- Attività:
 - aggiornamento dei contenuti
 - aggiornamento tecnologico
- **Interventi di manutenzione sistemistica, correttiva o evolutiva:** mantenere il sito e l'infrastruttura hardware e software correttamente funzionanti nel tempo, correggere errori e risolvere i problemi che si manifestano durante la vita operativa, migliorare la qualità in uso ed effettuare limitati interventi per rispondere a nuove esigenze.
- Attività:
 - manutenzione

Fase-6: gestione e valutazione (2)

- **Interventi di modifica o adattamento del progetto originale:** capire quale è il comportamento degli utenti mentre visitano o navigano tra le pagine del sito. Raccogliere commenti da parte degli utenti. Identificare eventuali problemi e proporre degli interventi da eseguire per migliorare l'esperienza di visita.
- **Web Analytics:** sistema di tracciamento dei visitatori di un sito web per finalità statistiche o per effettuare marketing mirato.
- Attività:
 - **analisi e valutazione della esperienza di visita**

Personalizzazione di un sito web

Personalizzazione

- **Personalizzazione:** **adattamento delle informazione** (es. del contenuto e della presentazione di un sito web) all'utente (o ad un gruppo di utenti).
- La personalizzazione è un tipo di **filtraggio**, in cui avviene una **selezione di informazioni e contenuti sulla base di specifici criteri**.
- La personalizzazione è un principio della persuasione (tecnologie persuasive).

Personalizzazione (2)

- La **personalizzazione e il filtraggio** sono nati per:
 - Dare una risposta al problema della **enorme quantità di informazione** che viene prodotta giornalmente sul web.
 - limitare la difficoltà crescente da parte del cittadino di identificare e scegliere la informazione rilevante o pertinente per un dato scopo o obiettivo (**carico informativo, iperscelta**).

Personalizzazione (3)

- La personalizzazione può tener conto di diverse informazioni:
 - I bisogni informativi, le intenzioni, gli interessi e le preferenze dell'utente (rilevabili dalla storia di interazione col sistema, dalle tracce lasciate dell'utente sul web e nel mondo reale).
 - Le relazioni e i legami sociali che l'utente intrattiene nel gruppo di appartenenza
 - La localizzazione geografica dell'utente (es. sistemi di *location awareness*).
 - Il contesto specifico in cui l'utente si trova in un dato momento.

Sistemi adattabili

- **Sistema adattabile:** l'utente prima di utilizzare il sistema, fornisce il proprio profilo attraverso una sessione di dialogo (es. compilando un questionario o scegliendo tra un insieme di profili predefiniti). Il sistema utilizza il profilo fornito dall'utente per personalizzare la presentazione o il contenuto.
- In questo caso **l'utente è consapevole** di fornire informazioni personali (**personalizzazione esplicita**).
- Nota: non è detto che le informazioni fornite dall'utente (es. preferenze, interessi) e quindi il profilo che ne deriva siano accurate e veritiere. Le informazioni fornite, inoltre, possono cambiare nel tempo, e questo richiede l'aggiornamento del profilo.

Sistemi adattativi

- **Sistema adattativo:** il sistema sfruttando le informazioni direttamente inferite dal comportamento dell'utente, adatta la presentazione o il contenuto in maniera automatica.
- In questo caso, **non è detto che l'utente sia consapevole** di fornire informazioni personali (**personalizzazione implicita**).
- Le informazioni sono sempre aggiornate (ossia riflettono il comportamento dell'utente).

Ottimizzazione di un sito web

Motori di ricerca

- I fornitori di pagine web che desiderano che le loro pagine siano riconosciute dai **motori di ricerca** devono porre attenzione a due attività fondamentali:
 - **Indicizzazione** delle pagine
 - **Ordinamento** dei risultati di ricerca nelle **SERP** (*Search Engine Results Page*)

Indicizzazione

- Una pagina web "esiste" se è presente nell' **indice** di un motore di ricerca. Se una persona vuole accedere ad una pagine che non è presente nell'indice lo può fare solo conoscendone l'indirizzo (es. l'URL).
- **Indice:** è una struttura dati (un database) che associa *parole (o frasi) chiave* ad *indirizzi* (es. URL) di pagine che contengono tali parole.
- Quando l'utente inserisce una o più parole chiave nel campo "cerca" di un motore (es. Google), il motore esplora l'indice e fornisce una lista di indirizzi URL di pagine che sono associate alle parole specificate e che idealmente sono in grado di soddisfare il bisogno informativo dell'utente.

Indicizzazione (2)

- **Indicizzazione:** attività svolta da un particolare tipo di software e finalizzata alla costruzione automatica dell'indice.
- Un **Web Spider** (o Web Crawler, o Web Bot) è un software che permette l'analisi della struttura ipertestuale del web visitando una pagina alla volta e, in maniera ricorsiva, le pagine ad essa collegate.
- Tale processo di scansione che avviene attraverso la catena di *hyperlinks* e che permette al motore di ricerca di scoprire nuovi contenuti è chiamata *crawling*.
- Nota: è possibile escludere una pagina dal processo di indicizzazione.

Indicizzazione (3)

- È quindi fondamentale conoscere come gli spiders indicizzano le pagine web.
- Alcuni problemi:
 - Come viene visitata la struttura ipertestuale del web?
 - Come vengono scelte le pagine da indicizzare?
 - Se non vengono considerate tutte le pagine esistenti, quali sono scelte?
 - Con quali criteri sono scelte?
 - Ci può essere il rischio di discriminazioni?
 - Come vengono analizzate le pagine?
 - Come sono scelte le keywords da usare nell'indice?

Indicizzazione (4)

- In una pagina web, le **parole chiavi** (*keywords*) sono un elemento fondamentale per l'indicizzazione.
- Esempi di keywords:
 - Le parole chiave indicate esplicitamente dall'autore della pagina nell'elemento <meta> di HTML.
 - Una o più parole che appaiono nell'elemento <title> di HTML.
 - Le prime N parole della pagina.
 - Tutte le parole della pagina.
 - Una o più parole che compaiono nelle intestazioni (elementi header H1,H2,H3 di HTML).
 - Termini che compaiono nei link.
 - Ecc..

Ordinamento

- **SERP** (*Search Engine Results Page*): la pagina che contiene i risultati di una ricerca. E' noto che gli utenti raramente prendono in considerazione i risultati successivi alle prime due o tre pagine (le prime 10-20 posizioni). **È importante quindi per un sito essere posizionato in "alto" nelle SERP del motore di ricerca usato.**
- L'obiettivo è quello di ordinare i risultati rispetto alla **pertinenza** del contenuto e rispetto al bisogno informativo espresso nella query dell'utente.

Ordinamento (2)

- **Algoritmo Page Rank:** l'ordinamento delle pagine dipende dalla importanza o popolarità del sito.
- L'importanza di una pagina $I(P)$ è misurata sulla base di:
 - il numero di link (voti) che puntano alla pagina considerata (*inlinks* o *backlinks*)
 - l'importanza delle pagine da cui partono i link

Ordinamento (3)

- Gli attuali motori di ricerca usano la **personalizzazione** (es. informazioni sul comportamento effettivo degli utenti: le keywords usate precedentemente, le pagine visualizzate, il tempo di permanenza sulle pagine, ecc.) per **modificare le SERP** (*Personalized Ranking Algorithms*).
- In molti casi i risultati ottenuti in maniera automatica sono trattati da operatori umani che rimuovono (manualmente) contenuti che considerano inappropriati, offensivi, ecc., rispetto ai termini di servizio, regolamenti governativi, richieste di terze parti (es. violazione di copyright).

Ordinamento (4)

- I motori di ricerca e molti servizi on line (es. Facebook, Twitter) svolgono una funzione di *gatekeeping* (**selezione e filtraggio dei contenuti**) analoga a quella esercitata dagli editori nei mass media tradizionali!
- Il *gatekeeping* è in generale un processo di selezione e trasformazione delle informazioni che quotidianamente raggiungono le persone, ed è il centro del ruolo dei media nella vita pubblica moderna. Questo processo determina non solo quali informazioni sono selezionate, ma anche quale sarà il contenuto e la natura dei messaggi e delle notizie.

Ottimizzazione per i motori di ricerca

- **Ottimizzazione per i motori di ricerca** (*Search Engine Optimization, SEO*): è un insieme di tecniche e procedure che possono essere usate per migliorare il posizionamento delle pagine di un sito nelle SERP.
- **SEO *on page***: interventi di ottimizzazione che riguardano direttamente le pagine interne del sito web (es. codice HTML, contenuti, struttura dei contenuti).
- **SEO *off page***: interventi di ottimizzazione che riguardano elementi esterni al sito web (es. percezione della popolarità, autorevolezza, affidabilità). Un aspetto importante del SEO off page riguarda, per esempio, le campagne di *link building* (si tratta di incrementare il numero di link in ingresso al sito web) e le attività di promozione sui social.

Web Analytics

- **Web Analytics:** insieme di metodi, tecniche e strumenti finalizzati alla **raccolta, misura, analisi e interpretazione** di dati relativi all'uso (visita) di un sito web con lo scopo finale di **ottimizzare l'esperienza degli utenti** rispetto a determinati scopi e obiettivi.
- Metodi di acquisizione dei dati:
 - Inserimento di codice JavaScript nelle pagine del sito. I dati acquisiti dal browser sono inviati ad un server remoto dove vengono analizzati. I risultati dell'analisi sono inviati all'utente sotto forma di report.
 - Utilizzo di cookie e dei log del server.

Web Analytics (2)

- Informazioni di interesse:
 - numero di visite (sessioni) in un determinato intervallo di tempo (al giorno, in un mese, ecc.)
 - pagina di arrivo
 - numero di pagine visualizzate
 - frequenza di visualizzazione di ciascuna pagina
 - tempo di permanenza su ciascuna pagina

Web Analytics (3)

- Informazioni di interesse:
 - azioni svolte dal visitatore della pagina
 - percorso di navigazione all'interno del sito
 - pagina di uscita
 - numero di visitatori che visitano una singola pagina e poi se ne vanno
 - numero di visitatori nuovi o che ritornano sul sito
 - la provenienza di visitatori (es. da social, diretta, ecc.)
 - il dispositivo usato per l'accesso

Bibliografia di approfondimento

- G. Brajnik, E. Toppano, *Creare siti web multimediali - Fondamenti per l'analisi e la progettazione*, Pearson, 2007.
- V. Roberto, M. Frailis, A. Gugliotta, P. Omero, *Introduzione alle Tecnologie WEB*, McGraw-Hill, 2004.