



Università degli Studi di Parma
Dipartimento di Scienze Matematiche,
Fisiche e Informatiche
Corso di Laurea in Informatica

Sistemi Informativi

La risorsa Informazione

Giulio Destri

Dr. Ing. Giulio Destri, Ph.D.

**Professore a contratto di Sistemi Informativi
@Università di Parma dal 2003**

**Digital Transformation Advisor, Innovation Manager,
Business Coach, Trainer @LINDA**

**Esaminatore ISO27021 e UNI11506-11621 BA (EPBA)
@Intertek**

Membro commissione UNI/CT 526 @UNINFO

Blogger @6MEMES di MAPS

**Certificazioni: ISO27001LA , ISO9001LA, ISO27021, ITILv3 e
v4, COBIT-2019, SCRUM Master, EPBA, NLP Coach, NLP AMP**

<https://www.linkedin.com/in/giuliodestri>

<http://www.giuliodestri.it/articoli.shtml>

giulio.destri@unipr.it

twitter.com/GiulioDestri

Scopo del modulo

Definire

**I concetti base
dell'informazione
e delle sue rappresentazioni digitali**

Argomenti

- La risorsa informazione
- La piramide DIKW
- Rappresentazione dell'informazione
- I flussi informativi entro l'azienda
- XML: l'esperanto elettronico
- Dati e informazione entro l'azienda
- La comunicazione



La risorsa informazione

La risorsa informazione

- Il lavoro del sistema informativo ha come oggetto l'**informazione**
- L'informazione ha caratteristiche particolari
- Che contribuiscono a rendere il sistema informativo diverso dai tradizionali settori tecnici di un'azienda

La risorsa informazione - 2

- L'informazione è la principale risorsa scambiata, selezionata ed elaborata nelle attività gestionali di coordinamento e controllo (E. Ciborra)
- Comunque, un qualunque compito nell'ambito di un'organizzazione, operativo o no, ha un contenuto gestionale e, in quanto tale, elabora informazione (P. Maggiolini)

La risorsa informazione - 3

- L'informazione è una risorsa immateriale (o intangibile)
- Costituisce la radice di ogni altra risorsa organizzativa immateriale come
 - **Conoscenza**
 - **Esperienza individuale**
 - **Esperienza organizzativa**

La risorsa informazione - 4

- L'informazione non viene distrutta dall'uso (non-depletable)
- Permette la creazione di nuova conoscenza (self-generating)
- Non è facilmente misurabile o divisibile o appropriabile
- Può essere soggetta a obsolescenza

La risorsa informazione - 5

- Processi gestionali efficienti sono in grado di instaurare circoli virtuosi di
 - generazione di conoscenza
 - e arricchimento dell'informazione disponibile
- Tali circoli, in linea teorica, generano un aumento delle prestazioni dei processi gestionali e dell'organizzazione in toto



La piramide DIKW

Informazione ed altro

- Dati
- Informazione
- Conoscenza
- Consapevolezza
- Saggezza

(da G. Bellinger, N. Shedroff ed altri)

La piramide DIKW



Dati

- Sono materiale informativo grezzo, non (ancora) elaborato dal ricevente
- I dati possono essere scoperti, ricercati, raccolti e prodotti
- Sono la materia prima che abbiamo a disposizione o produciamo per costruire i nostri processi comunicativi
- Esempio: l'insieme dei valori dei dati di accesso ad un determinato portale Web

Informazione (1/2)

- Viene costruita dai dati elaborati cognitivamente
- cioè trasformati in un qualche schema concettuale successivamente manipolabile e usabile per altri usi cognitivi

Informazione (2/2)

- L'informazione conferisce un significato ai dati, grazie al fatto che
 - li pone in una relazione reciproca
 - e li organizza secondo dei pattern.
- Trasformare dati in informazioni significa
 - organizzarli in una forma comprensibile,
 - presentarli in modo appropriato
 - e comunicare il contesto attorno ad essi
- Esempio: il risultato dell'analisi dei dati di accesso al sito Web

Conoscenza (1/3)

- E' informazione applicata,
- come un senso comune, o non comune, che "sa" quando e come usarla
- E' attraverso l'esperienza che acquisiamo conoscenza
- E' grazie alle esperienze che facciamo - siano esse positive o negative - che arriviamo a comprendere le cose

Conoscenza (2/3)

- La conoscenza viene comunicata
 - sviluppando interazioni stimolanti, con gli altri o con le cose,
 - che rivelano i percorsi nascosti e i significati dell'informazione
 - in modo che possano essere appresi dagli altri.

Conoscenza (3/3)

- La conoscenza è fondamentalmente un livello di comunicazione partecipatorio
- Dovrebbe rappresentare sempre l'obiettivo a cui tendere poiché consente di veicolare i messaggi più significativi
- Esempio: azioni di marketing svolte sulla base delle informazioni tratte dall'analisi dei dati di accesso al sito Web

Saggezza (1/3)

- Verità “eterna” distillata dalla conoscenza
- L'informazione costituisce lo stimolo di un'esperienza, mentre la saggezza può derivare dalla comprensione del messaggio che acquisiamo attraverso l'esperienza.
- La saggezza è il livello di comprensione più indefinito e più intimo.

Saggezza (2/3)

- La saggezza è una sorta di "meta-conoscenza" di processi e relazioni che viene acquisita attraverso l'esperienza.
- E' il risultato di contemplazione, valutazione, retrospezione e interpretazione - tutti processi estremamente personali.
- Non è possibile creare la saggezza allo stesso modo di come creiamo i dati e le informazioni, e non possiamo dividerla con gli altri come invece avviene per la conoscenza.
- Possiamo fornire la conoscenza atta a creare la saggezza dentro le persone

Saggezza (1/3)

- E' soltanto possibile creare esperienze che siano in grado di offrire opportunità e descrivere dei processi.
- Esempio: regole di azione e di uso dello strumento Web estratte dalla conoscenza guadagnata dall'esperienza

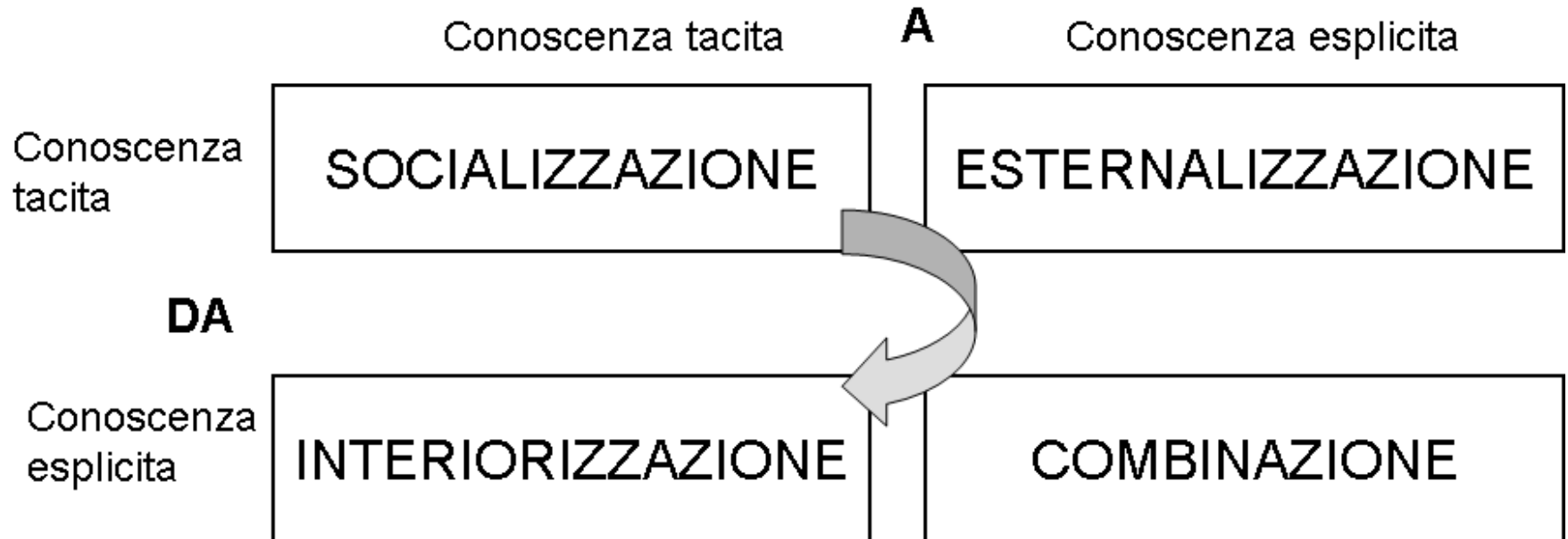
Informazione ed altro

- “La conoscenza consiste di fatti, verità, credenze, prospettive e concetti, giudizi e aspettative, metodi e saper fare”
(K.M. Wiig)

Trasmissione della conoscenza

- *Socializzazione* (da conoscenza tacita a tacita, in persone diverse che interagiscono)
- *Esteriorizzazione o esternalizzazione* (da conoscenza tacita a esplicita, tramite formalizzazione)
- *Combinazione* (da conoscenza esplicita a esplicita, tramite elaborazioni o semplici trasferimenti)
- *Interiorizzazione* (da conoscenza esplicita a tacita, tramite apprendimento e assimilazione)

La spirale della conoscenza



Nonaka e Takeuchi (1995)



La rappresentazione dell'Informazione

Rappresentazione dell'informazione

- Per esistere **nel mondo fisico**, l'informazione deve essere rappresentata **in modo fisico**
- L'informazione può essere rappresentata come **variazioni di grandezze fisiche entro opportuni supporti fisici**
- Esempi: colori su carta, livelli di tensione entro circuiti elettronici

L'Informazione nella nostra mente

- L'informazione poi può essere decifrata e entrare nel nostro cervello attraverso i nostri **organi di senso** (vista, udito, tatto, odorato, gusto ecc...).
- Qui esiste nella forma delle nostre **rappresentazioni mentali**, che sono costituite, nel mondo fisico, dalla **configurazione istantanea dei neuroni** nel nostro cervello e dagli **stati elettrochimici** delle loro connessioni.

Informazione e supporti fisici

- Per potere essere immagazzinata e trasmessa l'informazione necessita di supporti fisici
- Supporti di immagazzinamento
 - Mezzi cartacei
 - Mezzi informatici
- Supporti di trasmissione
 - Posta
 - Fax
 - Sistemi digitali (EDI, Internet...)

La codifica

- Per rappresentare l'informazione non bastano le semplici grandezze fisiche.
- Serve anche una opportuna ***codifica***,
- espressa attraverso un appropriato insieme di *regole*,
- che faccia *corrispondere* una certa *variazione di grandezza fisica* ad un *preciso componente di informazione*.

Meta-Informazione

- E' formata dalle **regole** che associano le **variazioni di grandezze** fisiche come il colore o il suono ai **simboli elementari** che compongono l'informazione
- Esempi
 - I fonemi delle lettere che formano le parole
 - Le forme delle lettere scritte
 - I colori rosso per l'alt e verde per l'avanti nel semaforo

Dati per rappresentare l'informazione

- I dati divengono sottocomponenti di informazione
- I dati possono essere rappresentati in forma digitale, attraverso opportune codifiche
- Standard pubblici (es. ASCII, MP3, TIFF, ebXML)
- Rappresentazioni proprietarie (.doc, .xls, tracciati proprietari...)

L'unità digitale dell'informazione

- L'elemento base fondamentale è il bit
- Un bit esprime due valori possibili
- Un insieme di n bit esprime 2^n valori possibili
 - Byte (8 bit) = 256 valori [0-255]
 - Short (16 bit) = 65536 valori [0-65535]
 - Int (32 bit) = più di 4 miliardi di valori!
- Le codifiche associano simboli diversi a ciascuno dei valori

Il testo digitale

- La rappresentazione di testo più diffusa è il codice **ASCII** (American Standard Code for Information Interchange)
- Associa ai valori da 0 a 255 rappresentati dai **byte**
 - le lettere (maiuscole e minuscole) dell'alfabeto latino internazionale,
 - le lettere accentate,
 - le cifre da 0 a 9,
 - i segni di interpunzione
 - caratteri di controllo
 - Ecc...

Il testo digitale: codici ASCII estesi

- Il codice ASCII è in realtà un insieme di vari codici diversi
- E' univoca solo l'assegnazione dei caratteri per i codici con il numero inferiore a 127
- lo standard ISO 8859 prevede 15 diverse versioni di ASCII esteso (128-255) comprese quelle per gli alfabeti diversi dal latino
- Più le rappresentazioni proprietarie

Il testo digitale: UniCode

- E' l'estensione dell'ASCII a 2 e 4 byte
- Unicode Consortium e ISO 10646
- UTF-8, UTF-16 e UTF-32

Il testo formattato: HTML

- HyperText Markup Language
- Pagine HTML: file di testo ASCII con incorporati comandi di formattazione (tag)
- Alcuni comandi comprendono i link
- Consente l'inclusione di ogni tipo di file o direttamente o attraverso link

Il testo formattato: HTML - 2

- La sintassi generale di un tag (anche detto comando, marcatore, elemento) in (X)HTML è:
`<tag attr1="val1" attr2="val2" ...>`
Testo a cui si applica il tag
`</tag>`
- `tag` è il nome del comando che vogliamo dare.
- `attr1`, `attr2`, ... si chiamano *attributi* e se presenti specificano l'azione del tag. `val1`, `val2`, ... sono i loro *valori*, scritti tra virgolette.
- Esempio: `<input type="password">` assegna il valore `password` all'attributo `type`.

La pagina HTML

- Lo schema di un documento HTML è:

```
<!DOCTYPE html>
```

```
<html ... >
```

```
  <head>
```

```
    <title>
```

```
      Titolo della pagina
```

```
    </title>
```

```
  </head>
```

```
  <body>
```

```
    <p>Testo della pagina</p>
```

```
  </body>
```

```
</html>
```

- Il rientro variabile delle righe (*indentazione*) è ignorato dal browser e serve per leggibilità.

Un elemento HTML: tabella

```
<table>
  <tr>
    <td>Titolo colonna</td>
  </tr>
  <tr>
    <td>Contenuto cella</td>
  </tr>
</table>
```

Titolo colonna
Contenuto cella

Le immagini digitali bitmap

- Composte da matrici rettangolari di pixel (picture element)
- Ogni pixel esprime il valore di luminosità o colore nel punto corrispondente dell'immagine
- Risoluzione = numero di pixel: oggi si misura in megapixel
- Profondità = numero di colori o toni di grigio

Tipi di immagini digitali bitmap

- Immagini monocromatiche con 1 bit per pixel (bianco o nero)
- Immagini a toni di grigio (di solito 256)
- Immagini a colori (da 16 a 16,7 milioni di colori)
 - 1 byte (256 colori con colormap)
 - 2 byte (65536 colori con colormap)
 - 3 byte (16,7 milioni di colori)
 - 4 byte (come 3 + le trasparenze)

Esempio di codifiche: immagini 8 bit

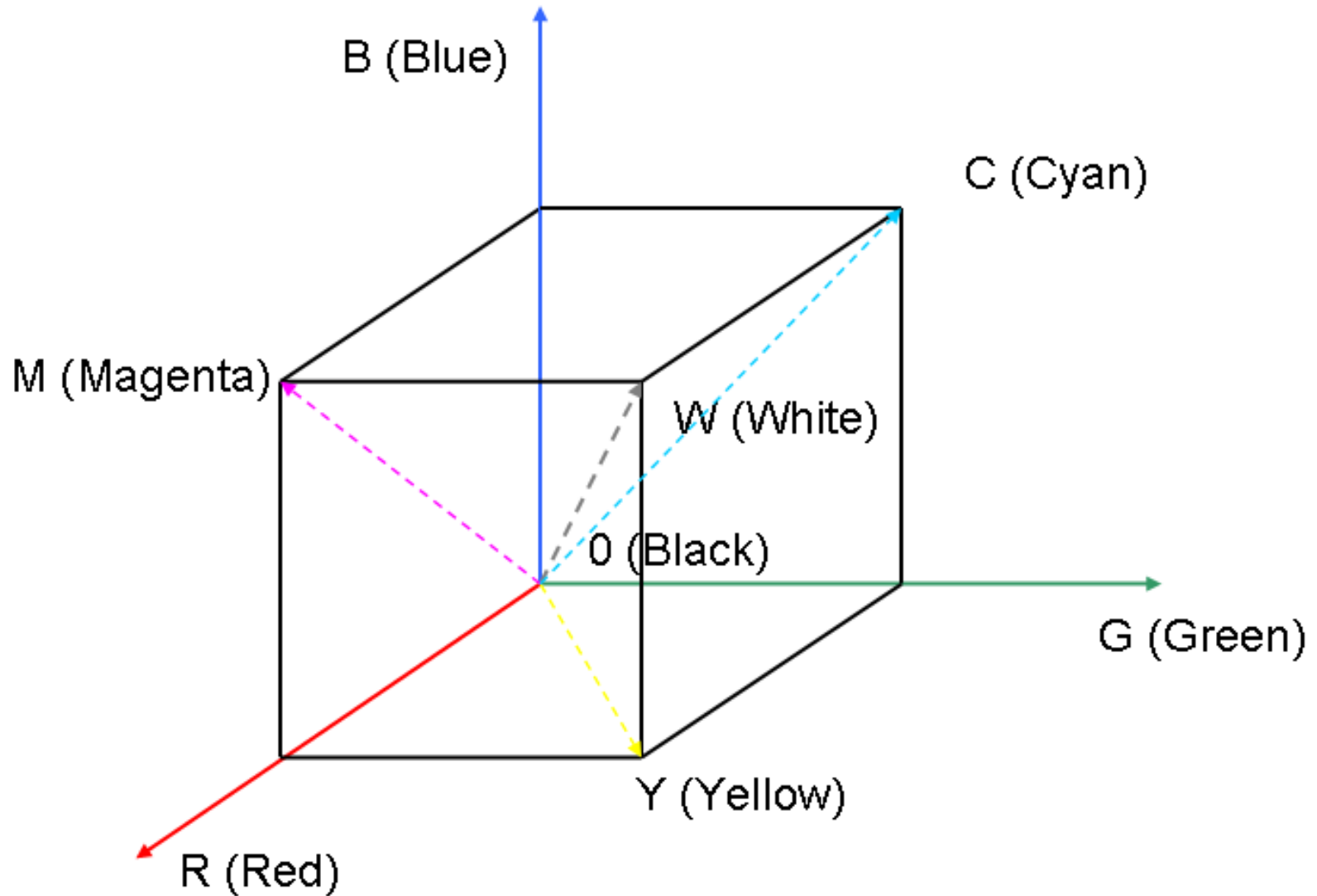


Diretta: 0 = nero, 255 = bianco

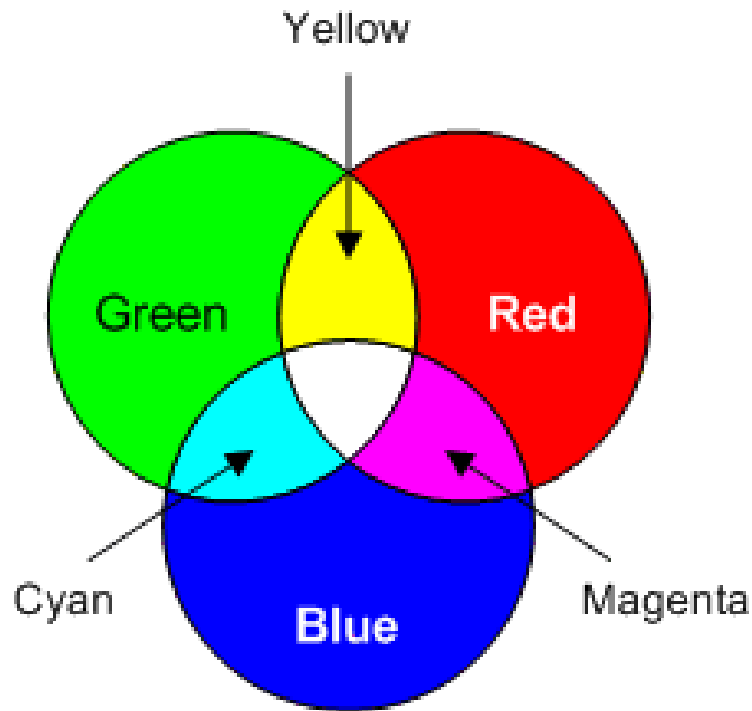


Inversa: 0 = bianco, 255 = nero

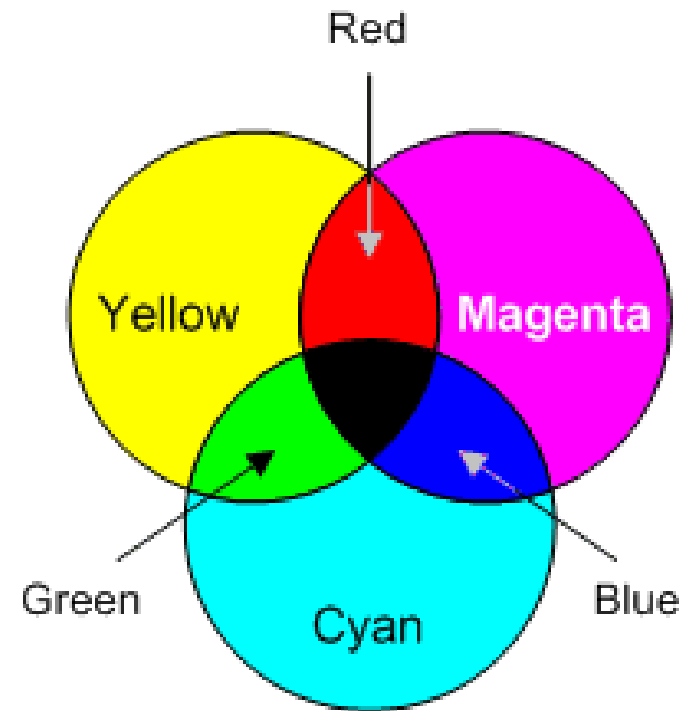
Il cubo RGB



Immagini a colori: sintesi additiva e sottrattiva

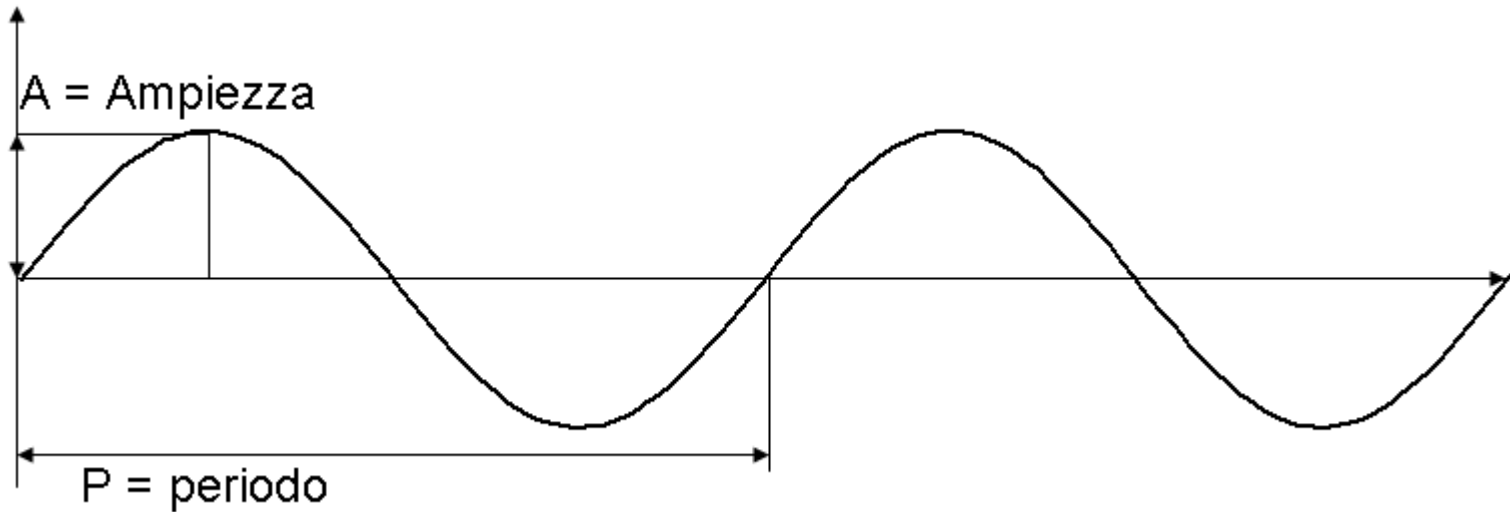


Sintesi additiva (monitor, tv...)



Sintesi sottrattiva (stampa...)

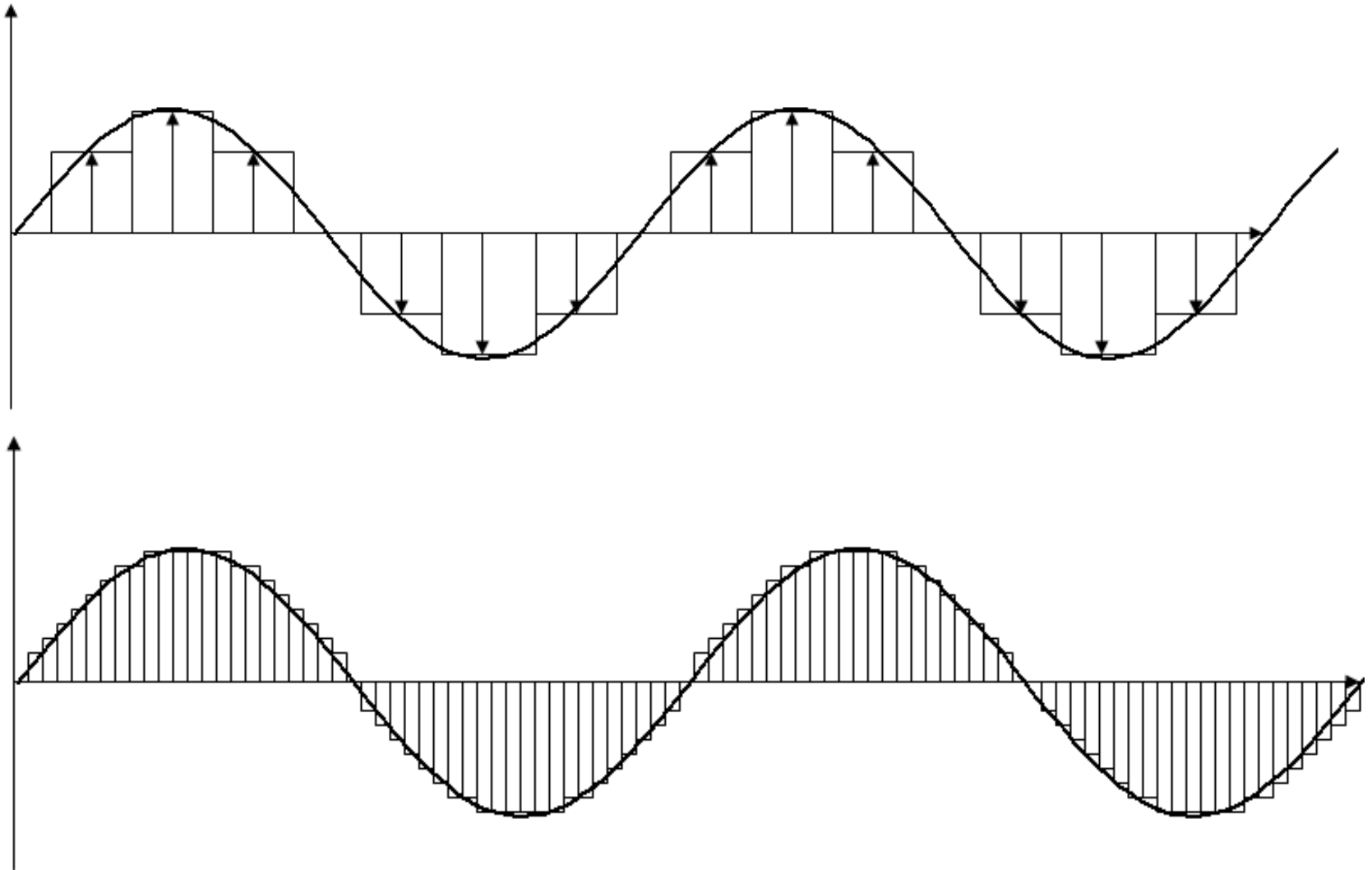
Il suono digitale: originale



$$v(t) = A \sin(2\pi t / P)$$

$$v(t) = A \sin(2\pi f t)$$

Il suono digitale: campionamento



Il suono digitale

- Flusso di campioni
- Teorema di Nyquist $f_c > 2f_M$
- Qualità CD: 44.100 Hz di campionamento
- Qualità Telefono: 8000 Hz di campionamento

Il filmato digitale

- Flusso di immagini statiche (+ audio)
- Per audio valgono le stesse proprietà già viste
- Qualità video lento: 10-15 frame al sec.
- Qualità normale: 24-25 frame al sec.
- Altri formati: 50-100 frame al sec.

I dati: visione IT

- Sono una rappresentazione della realtà di interesse
- Devono essere organizzati, classificati e archiviati in modo da poter essere facilmente reperiti e trasformati da “fatti grezzi” in informazioni utili a significative decisioni
- Rendono persistenti i risultati delle elaborazioni

Dati e Applicazioni

- Le Applicazioni elaborano dati
- Le Applicazioni trasformano le informazioni
- Le informazioni in ingresso sono dette dati di Input
- Il prodotto dell'elaborazione sono i dati di Output
- I dati devono durare oltre l'applicazione che li ha generati/acquisiti

Dati strutturati

- Tracciato a record con campi separati da delimitatori
- Tracciato a record con campi a lunghezza fissa
- Tracciati binari
- Basi di dati non relazionali
- Basi di dati relazionali
- Export di basi di dati

I tipi di dati: file

- Testi ASCII, UniCode, EBCDIC...
- Documenti di Word processor (.doc, .rtf, .sxw...)
- Fogli elettronici (.xls, .sxc, .csv...)
- Archivi di database personale (.mdb...)
- Rubriche (.wab, .nab...)
- Immagini (.gif, .jpg, .png, .tif...)
- Suoni (.mp3, .wav, .au...)
- Filmati (.avi, .mpeg...)
- Altri...



Flussi informativi e flussi informatici

I flussi informativi

- Sono i flussi di informazioni che vengono trasferite tra diverse componenti di un'impresa o tra l'impresa e i propri clienti
- Possono avvenire attraverso diversi mezzi fisici di comunicazione
 - Voce, Telefono
 - Fax
 - Posta cartacea
 - EDI, Internet, reti informatiche (flussi informatici)

I flussi informativi: esempi

- Aggiornamento del contenuto di portali Web
- Agenzie di stampa
- Scambi azionari, transazioni elettroniche bancarie
- Compravendita di prodotti e servizi (richiesta d'ordine, offerta, conferma d'ordine...)
- Trasferimento di progetti alla produzione
- Spedizione di dati sanitari tra ospedali

Flussi informativi e flussi informatici

- Quando un flusso informativo avviene totalmente attraverso strutture ICT prende il nome di **flusso informatico**
- I flussi informatici sono convenienti per l'azienda in quanto più rapidi e a prova di errore
- Evitano attività parassitiche come, ad esempio, ribattitura di testi o dati

Tipi di flussi

- Invio semplice
- Richiesta/risposta sincrona
- Richiesta/risposta asincrona
- Ricezione semplice

Sistema Nervoso Digitale

- Il Digital Nervous System è la visione di un flusso informativo ideale,
- che collega tra loro le organizzazioni
- e che le attraversa singolarmente al loro interno,
- mettendole in condizione di agire, rispondere e adattarsi alle esigenze degli utenti e del mercato più rapidamente e meglio rispetto ai concorrenti.

(B. Gates)

Sistema Nervoso Digitale - 2

- Un efficace Digital Nervous System si basa su una combinazione di elementi.
- In primo luogo, è necessario allevare e promuovere una cultura organizzativa in cui le informazioni circolano e le conoscenze vengono condivise.
- Quindi è necessario che i sottostanti processi aziendali supportino la visione complessiva del flusso informativo ideale.
- Infine, serve l'Information technology - PC, software, comunicazioni, reti etc. - che fornisce l'infrastruttura sulla quale il flusso delle informazioni viene trasmesso.

Sistema Nervoso Digitale - 3

- E' importante comprendere che un Digital Nervous System efficace non può essere creato senza la presenza di tutti questi componenti - ognuno di essi è infatti correlato a tutti gli altri.
- La tecnologia da sola "non può sostituire le qualità di una buona leadership, che resta l'origine prima di ogni buon management".
(B. Gates)



XML: l'esperanto elettronico

XML

Modalità chiaramente definita per

- strutturare
 - descrivere
 - interscambiare
 - i dati
-
- Formato dati, leggibile sia da operatori umani sia da macchine

Vantaggi di XML

- Standard W3C
- Libero da licenze
- Indipendente dalle piattaforme
- Ben supportato dai vendor e dai linguaggi (COBOL, Java, .NET, C/C++, Delphi)
- Definisce solo la semantica e non la presentazione

XML e XSL

- XML (eXtensible Markup Language) usa i tag per il significato semantico del contenuto (es. <AUTORE> ... </AUTORE>)
- XSL (eXstensible Style Language) contiene i comandi che servono ai parser XML per convertire un documento XML in qualsiasi altro formato (es. HTML)

In pratica cos'è l'XML?

Uno standard sviluppato dal W3C
per la gestione di testo
strutturato

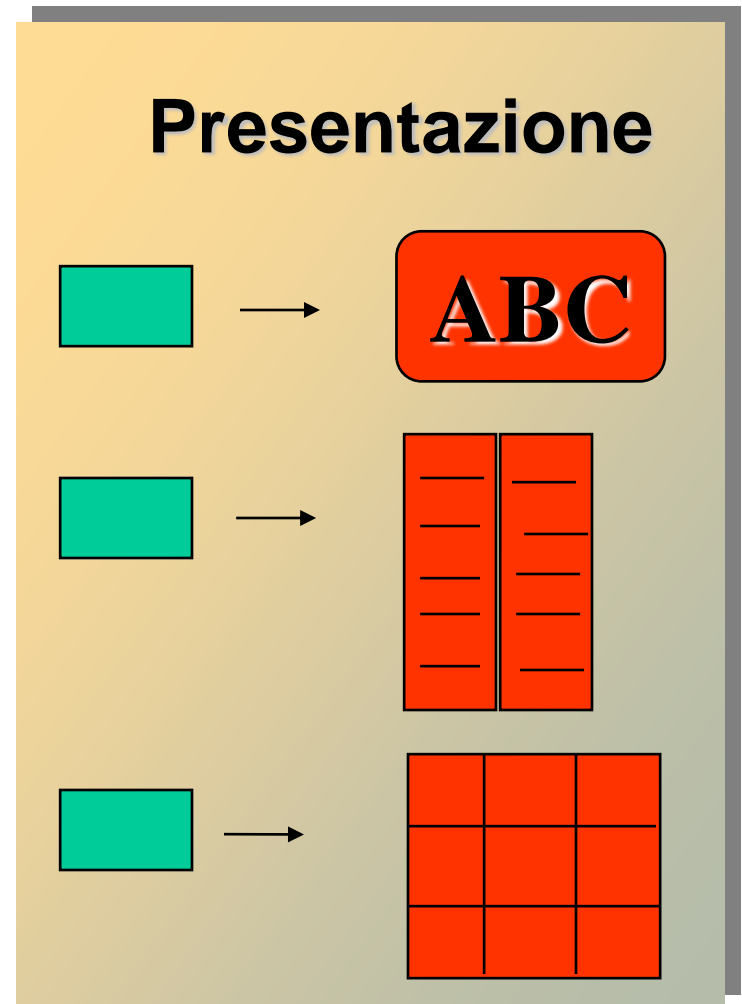
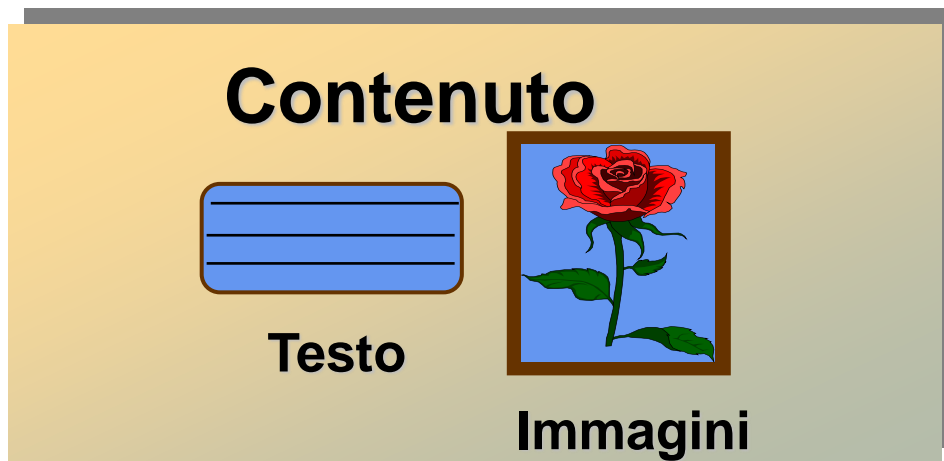
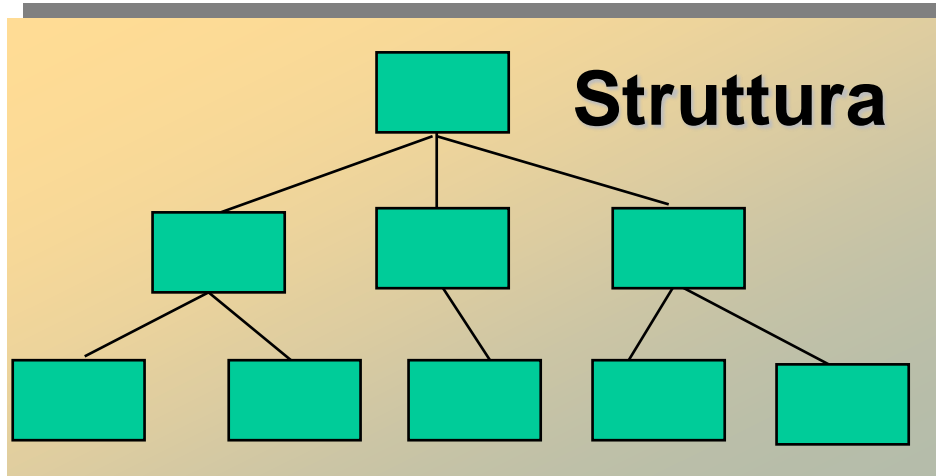
L' "Extensible
Markup Language"

<?XML!>

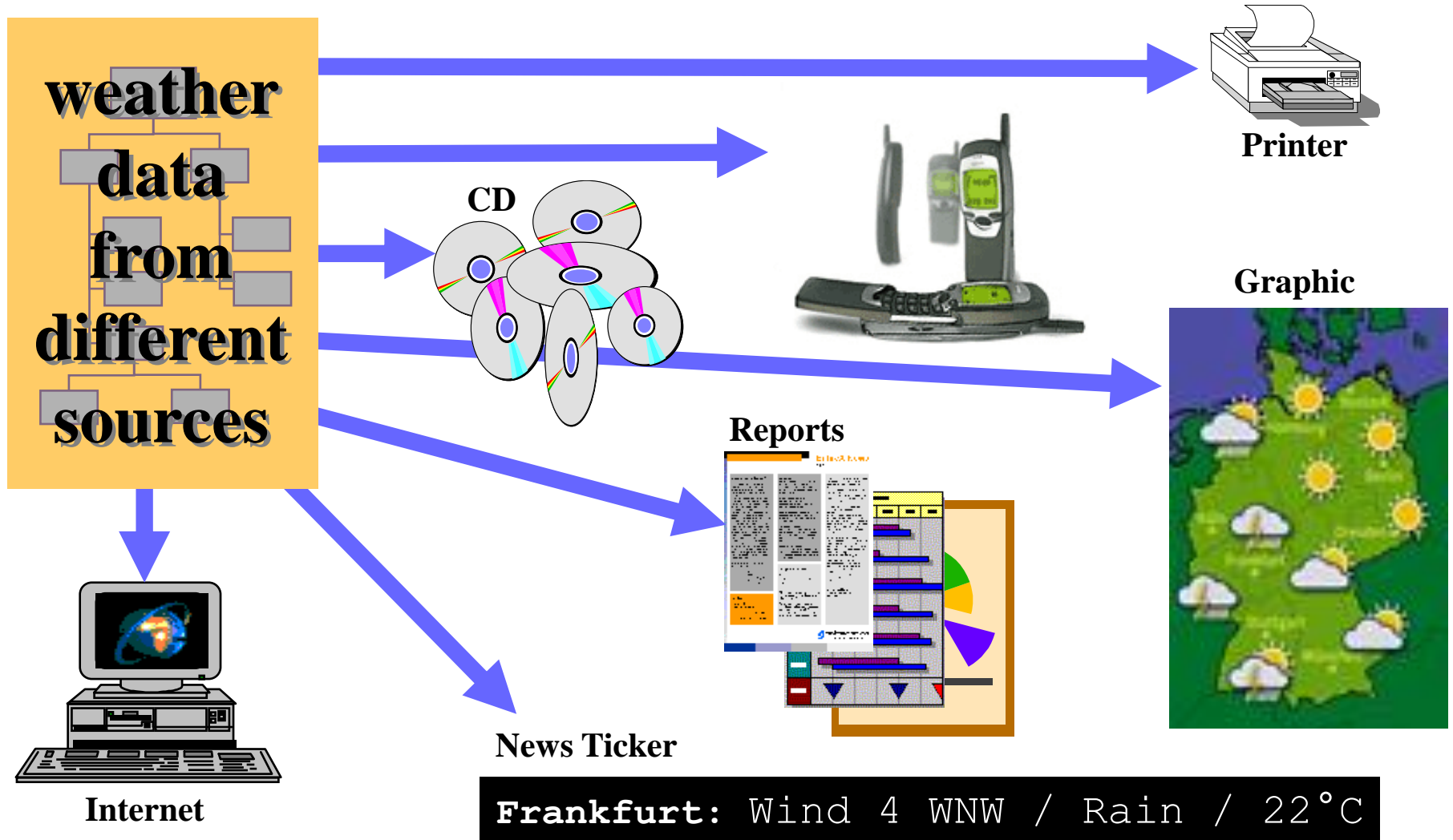
XML separa contenuto,
presentazione e struttura
permettendo che i
documenti siano facilmente
manipolabili

E' un Meta Linguaggio di Markup

Caratteristiche di un documento XML



Vantaggi di XML



Struttura di XML

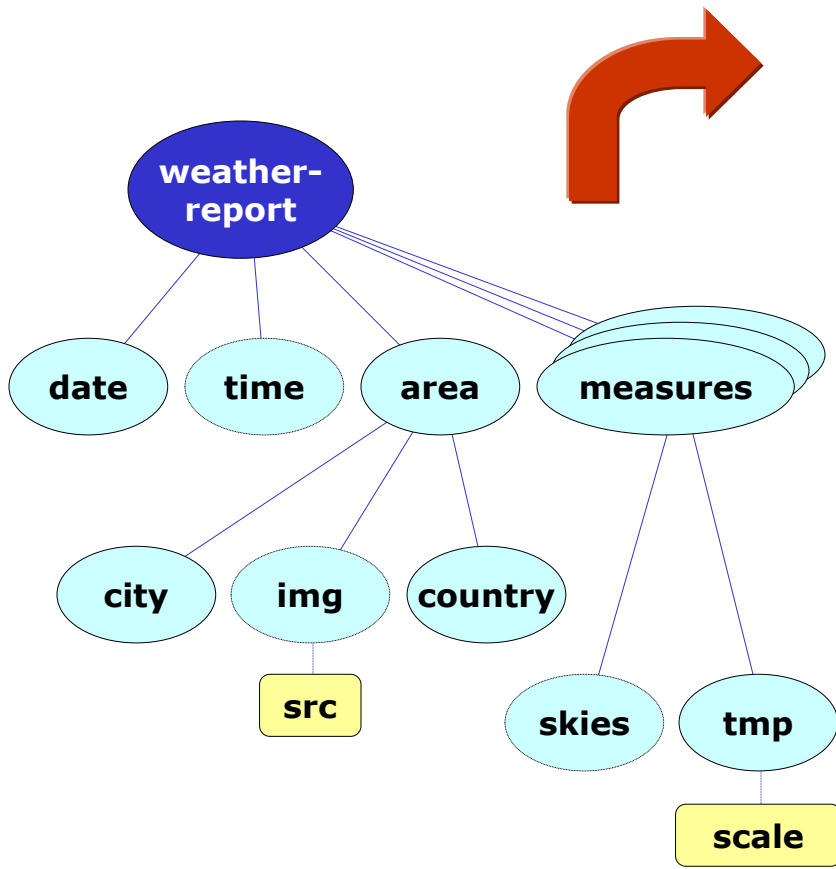
```
<?xml version="1.0"?>
<!DOCTYPE weather-report
  SYSTEM "WeatherBasic.dtd">
<weather-report>
  <date>2, April, 1999</date>
  <time>10am</time>
  <area>
    <city>Madrid</city>
    
    <country>Spain</country>
  </area>
  <measures>
    <skies>cloudy</skies>
    <temp scale="C">18</temp>
  </measures>
</ weather-report >
```

Leggibile & Semplice

***Contenuto e Struttura
Auto descrittivi***

***Gestione
Documenti,
Interscambio
Dati, ...***

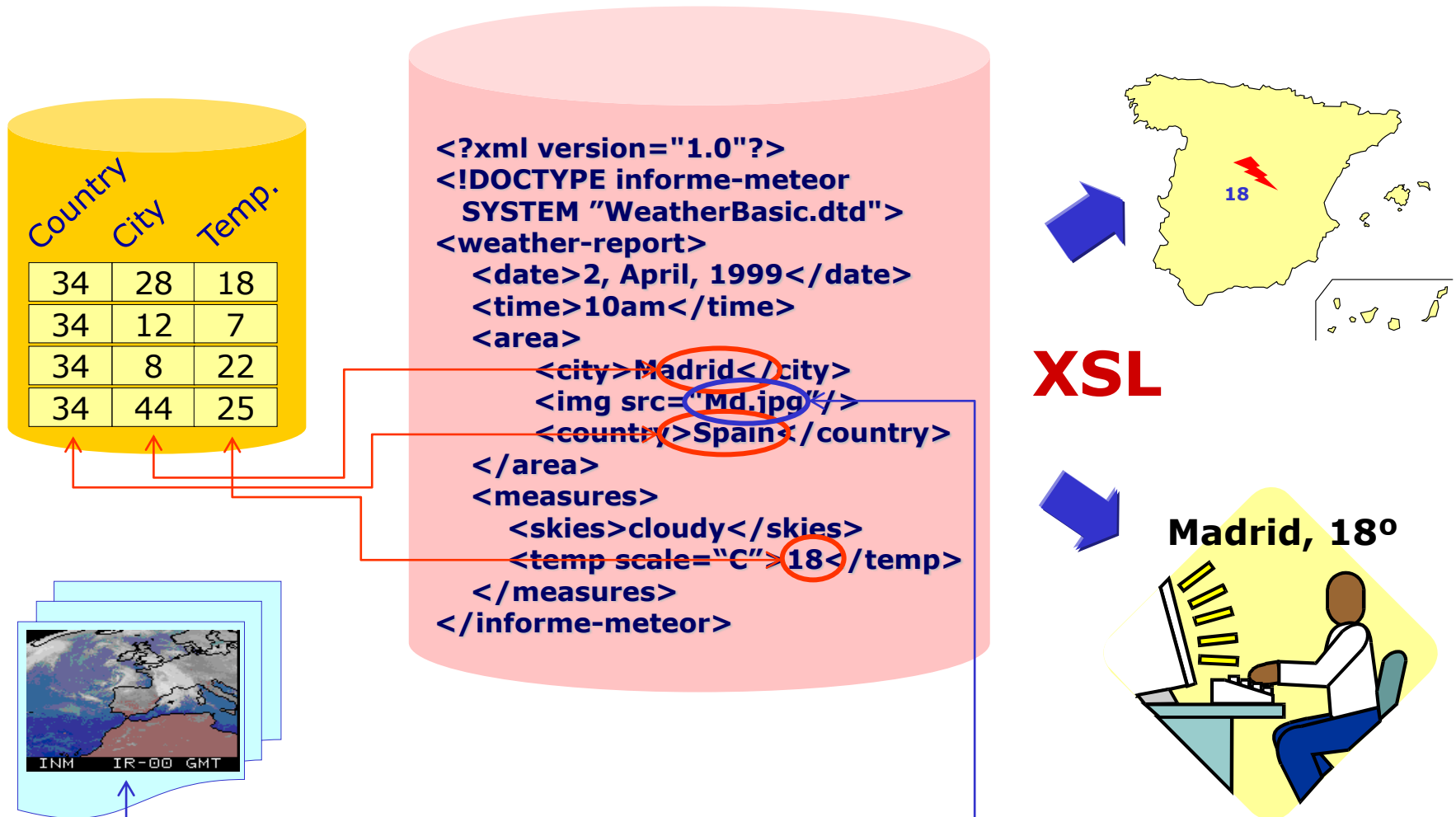
XML: standard per strutturare dati



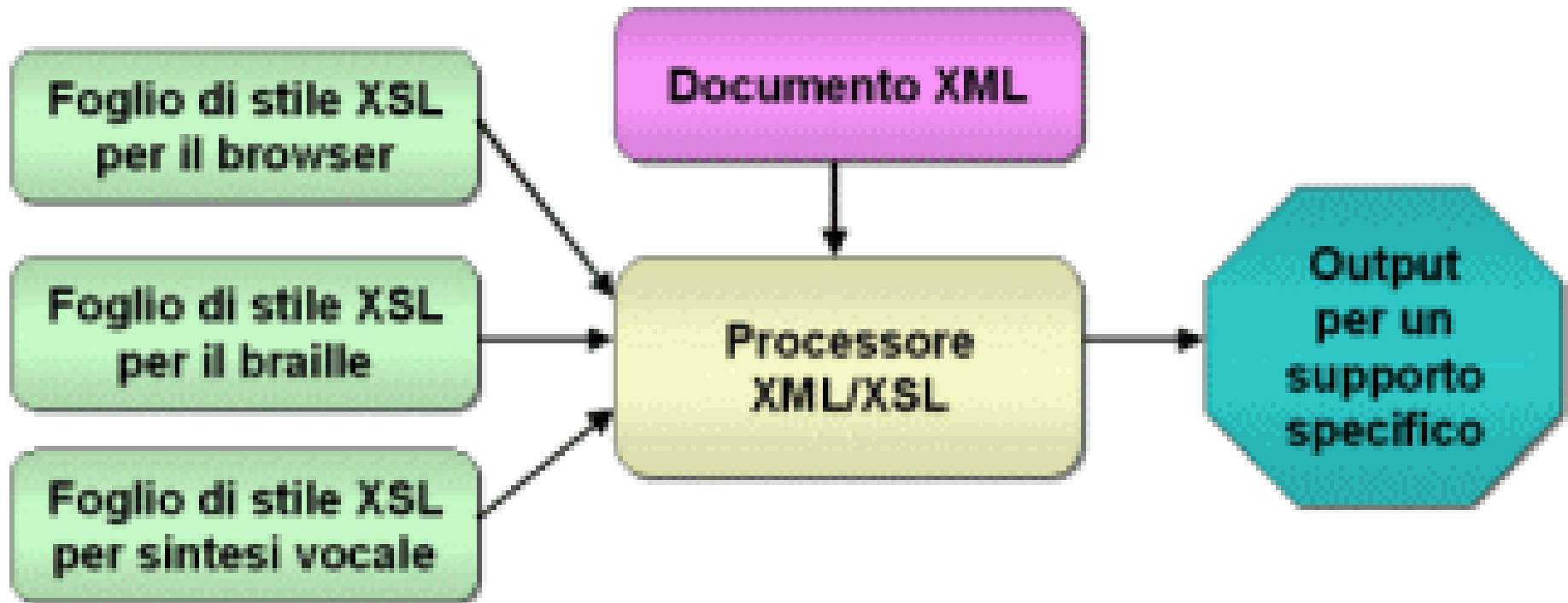
Document Type Definition (DTD)

```
<!ELEMENT weather-report  
    (date, time?, area, measures+)>  
<!ELEMENT area  
    (city, img?, country)>  
<!ELEMENT measures  
    (skies?, temp)>  
  
<!ELEMENT date (#PCDATA)>  
<!ELEMENT time (#PCDATA)>  
<!ELEMENT city (#PCDATA)>  
  
<!ELEMENT img EMPTY>  
<!ATTLIST img  
    src CDATA #REQUIRED>  
  
<!ELEMENT country (#PCDATA)>  
<!ELEMENT skies (#PCDATA)>  
  
<!ELEMENT temp (#PCDATA)>  
<!ATTLIST temp  
    scale (C | F) "C">
```

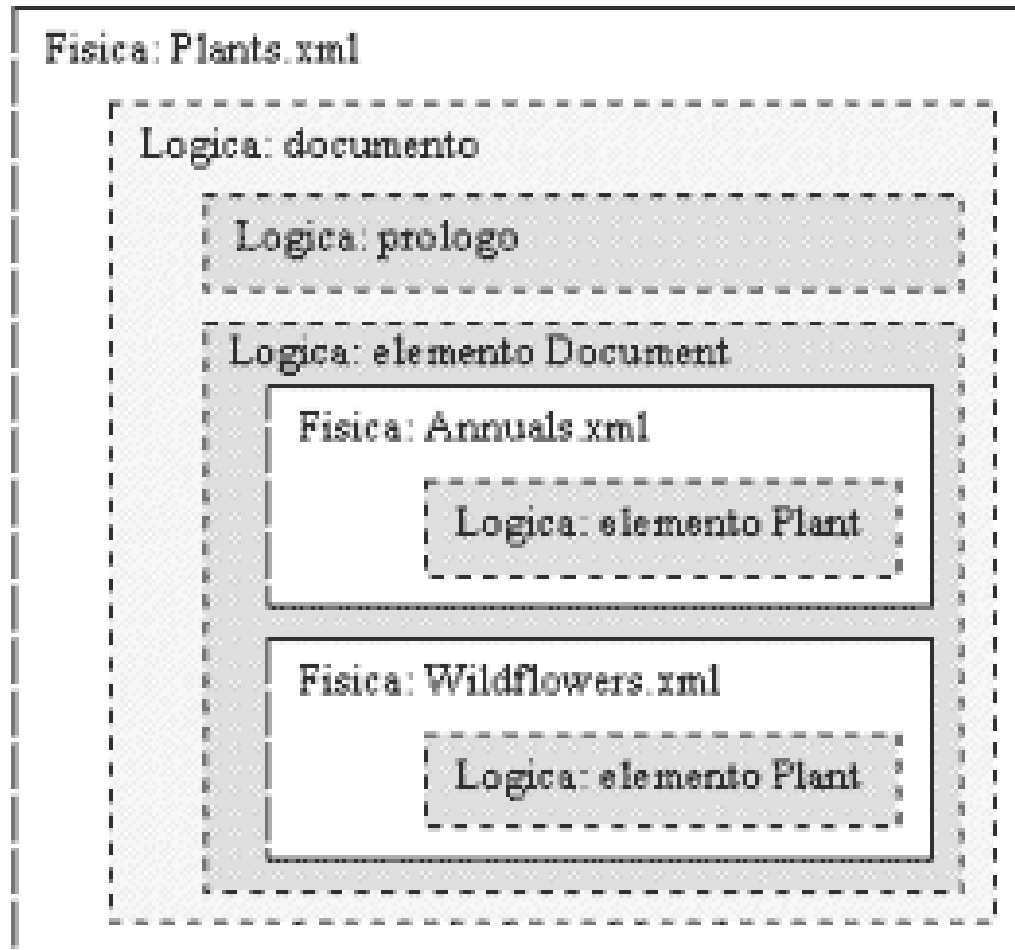
XML: standard per pubblicare



XML/XSL



XML/XSL



Il Prologo

- Dichiarazione XML
 - `<?xml version="1.0"?>`
- Dichiarazione del tipo di documento
 - `<!DOCTYPE Wildflowers SYSTEM "Wldflr.dtd">`

L'elemento Document

<PLANT>

<COMMON>Columbine</COMMON>

<BOTANICAL>Aquilegia

canadensis</BOTANICAL>

</PLANT>

La nidificazione dei dati XML

<DOCUMENT>

<PARENT1>

<CHILD1>...</CHILD1>

<CHILD2>...</CHILD2>

</PARENT1>

</DOCUMENT>

I tag in XML

Sono sempre doppi

`<mioTag> ... </mioTag>`

Oppure vuoti

`<mioTag />`

XML è case-sensitive

Attributi dei tag in XML

Gli ***attributi*** consentono di associare valori a un elemento senza che siano considerati parte del contenuto dell'elemento stesso

```
<PLANT ZONE=3> <COMMON>Columbine</COMMON>  
  <BOTANICAL>Aquilegia  
  canadensis</BOTANICAL>  
</PLANT>
```

Documento XML valido

- La definizione del tipo di documento DTD specificata nel prologo delinea tutte le regole relative a un documento.
- Un documento XML *valido* segue tutte queste regole rigidamente.
- Un documento valido è conforme anche a tutti i limiti di validità identificati dalle specifiche relative all'XML.

Documento XML ben formato

- Tutti i tag di apertura e di chiusura corrispondono.
- I tag vuoti utilizzano una sintassi XML speciale.
- Tutti i valori degli attributi sono racchiusi tra virgolette.
- Tutte le entità sono dichiarate.

DTD

- Dichiarazione del Tipo di Documento o DTD
- Un sottoinsieme DTD esterno e un sottoinsieme DTD interno
- L'interno sovrascrive (e quindi ha priorità) sull'esterno

Un semplice DTD

```
<?xml version="1.0"?>
```

```
<!DOCTYPE EMAIL [ <!ELEMENT EMAIL  
  (TO, FROM, CC, SUBJECT, BODY)>
```

```
  <!ELEMENT TO (#PCDATA)>
```

```
  <!ELEMENT FROM (#PCDATA)>
```

```
  <!ELEMENT CC (#PCDATA)>
```

```
  <!ELEMENT SUBJECT (#PCDATA)>
```

```
  <!ELEMENT BODY (#PCDATA)>
```

...e un XML associato

```
<?xml version="1.0"?>
<!DOCTYPE EMAIL SYSTEM "Email.dtd">
<EMAIL>
  <TO>Jodie@msn.com</TO>
  <FROM>Bill@msn.com</FROM>
  <CC>Philip@msn.com</CC>
  <SUBJECT>My first DTD</SUBJECT>
  <BODY>Hello, World</BODY>
</EMAIL>
```

Dichiarazione di Elemento

Ogni dichiarazione di elemento contiene il nome dell'elemento e il tipo di dati definito *specifiche di contenuto* costituite da uno tra i quattro tipi seguenti:

- Un elenco di altri elementi, denominato modello di contenuto
- La parola chiave EMPTY
- La parola chiave ANY
- Contenuto di vario tipo

Esempi di Dichiarazioni

Elenco

```
<!ELEMENT EMAIL (TO, FROM, CC,  
    SUBJECT, BODY)>
```

Elemento vuoto

```
<!ELEMENT TEST EMPTY>
```

Elemento “universale”

```
<!ELEMENT TEST ANY>
```

Esempi di Dichiarazioni

Contenuto specificato

```
<!ELEMENT EXAMPLE (#PCDATA | x | y | z) *>
```

Contenuto specificato semplice

```
<!ELEMENT EXAMPLE CDATA>
```

Molteplicità

- ‘?’ = un elemento deve essere visualizzato una sola volta o non apparire mai
- ‘*’ = Indica che l’elemento può essere visualizzato ogni volta che l’autore desidera
- ‘+’ = Indica che un elemento deve essere visualizzato una o più volte
- “ (nessun simbolo)= Indica che deve essere visualizzato un solo elemento

Schemi XML e DTD

Gli Schemi XML sono nati con lo stesso scopo dei DTD:

- Specificare la struttura dei documenti.
- Modelli di contenuto, elementi radice, ...
- Specificare il tipo dei dati utilizzabili all'interno di elementi e attributi.
- Nei DTD, il tipo di dato per gli elementi poteva essere solo testo e/o un particolare modello di contenuto.

Perché gli Schemi?

- I DTD usano una sintassi non XML.
 - Perché costringere gli sviluppatori ad imparare le regole di un nuovo linguaggio?
 - Perché scrivere parser XML che debbano leggere anche formati non XML per la validazione?
- I DTD hanno pochi tipi di dato.
 - Il controllo sui domini dei dati è parte integrante del controllo di un documento “ben formato”.
 - Esistono tipi di dato molto noti che si vorrebbero poter sfruttare (interi, reali, data/ora, ecc.).

Novità negli Schemi

- Sono supportati più **tipi di dato** e la possibilità di definirne di nuovi o derivarne altri da tipi già esistenti applicando regole e **restrizioni**.
- Sono presenti nuovi **modelli di contenuto**, tra cui l'insieme ("tutti questi elementi, in qualsiasi ordine").
- Si possono definire più **campi chiave** diversi.
- Si possono dichiarare **classi di equivalenza** tra elementi.

Vantaggi degli Schemi

- Generalmente, le applicazioni che devono basarsi su documenti o dati provenienti dall'esterno "sprecano" una grossa quantità di codice per controllarne la validità.
- Più complessi sono i dati, più il codice sarà laborioso da scrivere.
- Se i dati sono strutturati secondo un preciso Schema XML, l'applicazione potrà avvalersi delle funzionalità di un qualsiasi validatore di schemi in commercio.

Cosa Forniscono gli Schemi

- Un modello per i dati.
 - Descrivono cioè l'organizzazione e i tipi dell'informazione.
- Un contratto.
 - Cioè un protocollo molto specifico per lo scambio di informazioni.
- Un insieme di Metadati.
 - Lo schema contiene molte informazioni valide per l'interpretazione dei dati strutturati sulla sua base.

Cosa Forniscono gli Schemi

- Oltre che per validare i documenti istanza, si possono immaginare molti altri impieghi per gli schemi:
 - Creazione automatica di interfacce per la compilazione dei documenti XML associati.
 - Creazione di interfacce grafiche per la rappresentazione dei dati.
 - Uso degli schemi per definire strutture dati e protocolli per la loro manipolazione e trasmissione.

Un DTD di esempio...

```
<!ELEMENT artist (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT year (#PCDATA)>
<!ELEMENT comment (#PCDATA)>
<!ELEMENT length (#PCDATA)>

<!ELEMENT song (artist?, title, year?, comment?, length)>
<!ELEMENT CD song+>
```

- Gli elementi blu fanno parte della sintassi DTD.
- Gli elementi neri fanno parte della nuova sintassi che stiamo definendo.
- Poiché i DTD sono stati creati prima dei namespaces, non esiste questo concetto nei DTD e gli elementi delle due sintassi sono mescolati.

...e lo Schema Corrispondente...

```
<xs:element name="artist" type="xs:string"/>
<xs:element name="comment" type="xs:string"/>
<xs:element name="length" type="xs:string"/>
<xs:element name="title" type="xs:string"/>
<xs:element name="year" type="xs:string"/>

<xs:element name="song">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="artist" minOccurs="0"/>
      <xs:element ref="title"/>
      <xs:element ref="year" minOccurs="0"/>
      <xs:element ref="comment" minOccurs="0"/>
      <xs:element ref="length"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>

<xs:element name="CD">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="song" maxOccurs="unbounded"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

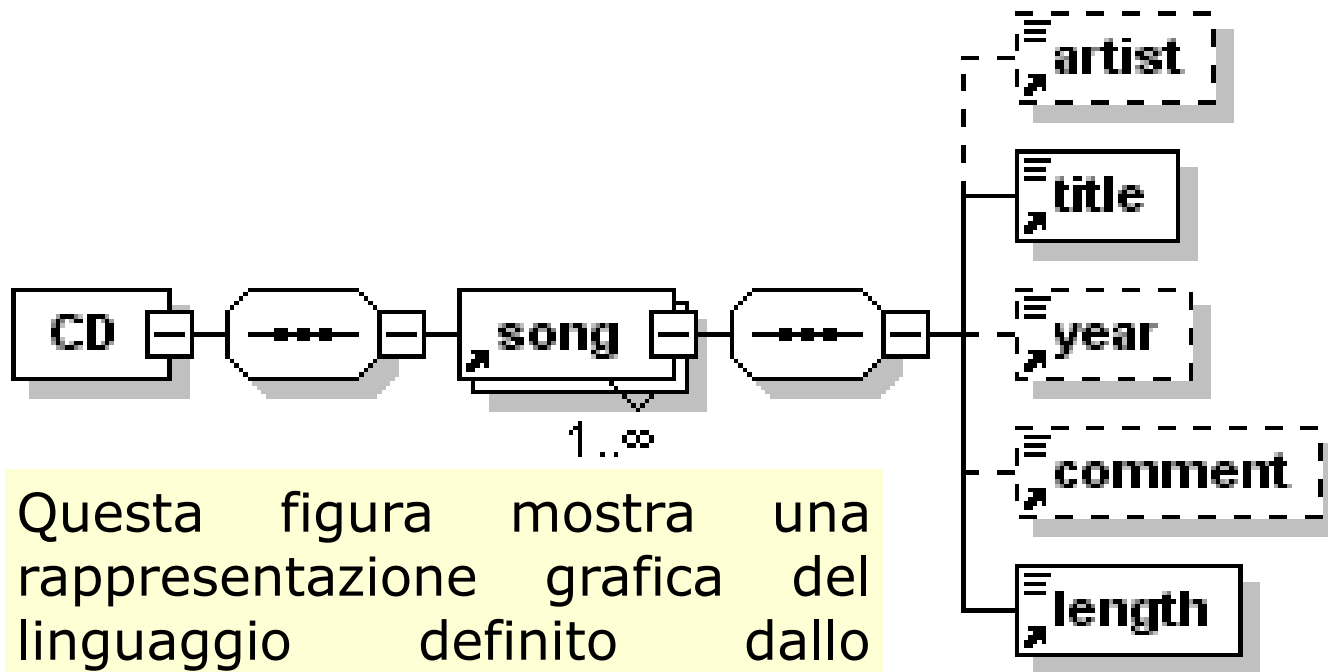
```
<!ELEMENT artist (#PCDATA)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT year (#PCDATA)>
<!ELEMENT comment (#PCDATA)>
<!ELEMENT length (#PCDATA)>
```

```
<!ELEMENT song
(artist?, title, year?, comment?, length)
>
```

```
<!ELEMENT CD song+>
```

Questo Schema definisce lo stesso linguaggio del DTD visto in precedenza. I frammenti del DTD sono affiancati alla parte di schema che li sostituisce.

...Visto Graficamente



Questa figura mostra una rappresentazione grafica del linguaggio definito dallo schema precedente.

La struttura molto raffinata degli Schemi permette di manipolarli con tool grafici avanzati.

I Namespace

- I namespace (spazi dei nomi) XML rappresentano insiemi distinti in cui il nome e la definizione di un tag (e quindi di un elemento XML) sono univoci
- Entro un namespace quindi ogni tag ha una ben specifica ed unica struttura

I Namespace (2)

- Esistono molti elementi XML predefiniti ed altri definiti dall'utente
- I namespace garantiscono una definizione univoca dei nomi, per evitare sovrapposizioni fra gli elementi XML
- Costituiscono infatti una metodologia per
 - la creazione di nomi universalmente univoci in un documento XML
 - identificando i nomi degli elementi con una risorsa esterna univoca.

I Namespace (3)

Nel linguaggio XML uno spazio dei nomi

- è pertanto una raccolta di nomi identificata da un URI
- Può essere
 - qualificato
 - o non qualificato.

I nomi qualificati

Nel linguaggio XML un nome qualificato è composto da:

- il nome dello spazio dei nomi (ovvero un URI) che definisce il namespace
- Un componente locale che identifica l'elemento locale

I nomi qualificati (2)

- E' necessario includere una dichiarazione dello spazio dei nomi nel prologo del documento.
- E' inoltre possibile includere nella dichiarazione un prefisso dello spazio dei nomi.
- Utilizzando i due punti (:), il prefisso può essere aggiunto alla parte locale in modo da associarla al nome dello spazio dei nomi.

I nomi qualificati: esempio

```
<?xml version="1.0"?>
<?xml:namespace
  ns=http://inventory/schema/ns
  prefix="inv"?>
<?xml:namespace
  ns=http://wildflowers/schema/ns
  prefix="wf"?>
<PRODUCT>
  <PNAME>Test1</PNAME>
  <inv:quantity>1</inv:quantity>
  <wf:price>323</wf:price>
  <DATE>6/1</DATE>
</PRODUCT>
```

Nomi non qualificati

- Un nome non qualificato non dispone di nome associato al nome dello spazio dei nomi.
- I nomi di elementi XML tipici non sono qualificati poiché non specificano uno spazio dei nomi.



Dati e informazione entro l'azienda

I dati dentro i sistemi informatici

Cosa si intende per “dati”?

- Contenuto di DB relazionali
- Archivi documentali/multimediali
- Micro applicativi (es. generatori report)
- DB personali (es. elenco indirizzi)
- Archivi di Directory Service
- Configurazioni dei programmi e delle postazioni di lavoro

I dati dentro i sistemi informatici - 2

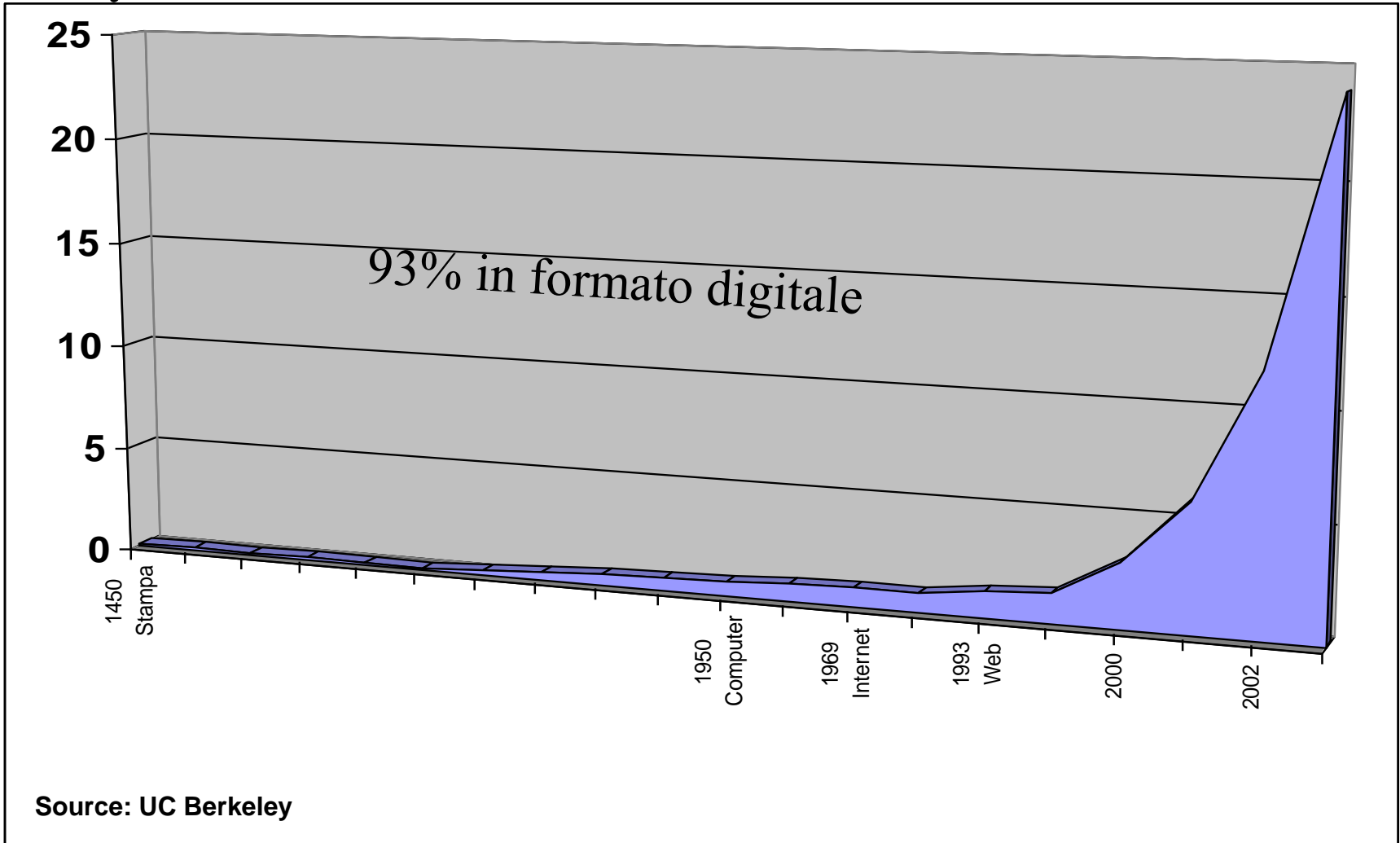
- In un sistema fortemente centralizzato tutti i dati risiedono o nel DB o, comunque, entro file sui dischi del server
- In un sistema distribuito i dati sono ripartiti su più server e hanno una forma molto varia
- Spesso poi ci sono dati importanti “sparsi in giro” per i client

Unità di Misura delle Informazioni

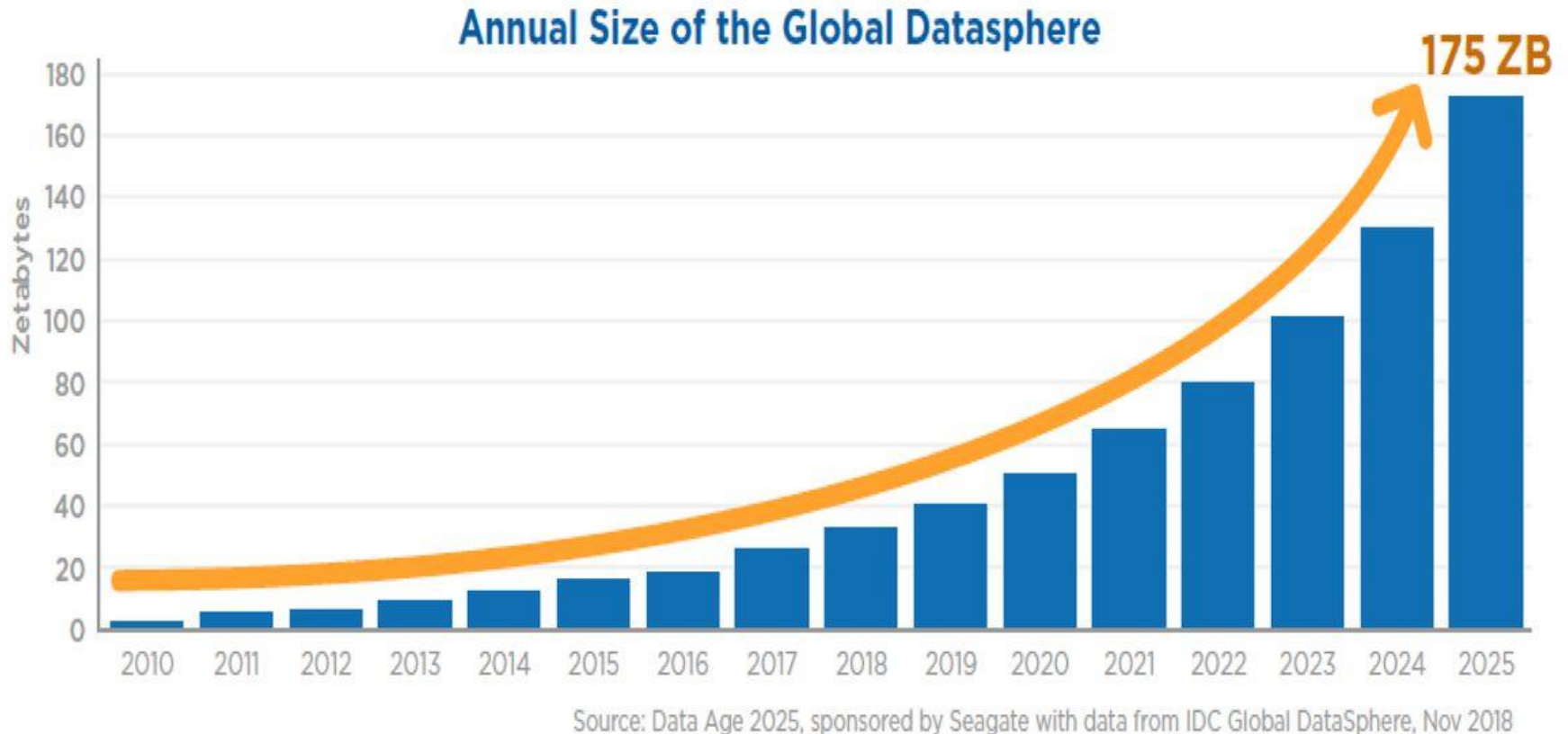
Kilobyte (KB)	1.000 byte 2 Kilobyte: Pagina dattiloscritta 10 Kilobyte: Pagina Web 100 Kilobyte: Fotografia a bassa risoluzione
Megabyte (MB)	1.000 Kilobyte 1 Megabyte: un libro di 250 pagine 5 Megabyte: 30 secondi di video (qualità TV) 500 Megabyte: CD musicale
Gigabyte (GB)	1.000 Megabyte 1 Gigabyte: un intero film (qualità TV) 20 Gigabyte: la registrazione completa di tutte le opere musicali di Beethoven 500 Gigabyte: le dimensioni del più grande sito FTP
Terabyte (TB)	1.000 Gigabyte 10 Terabyte: I libri contenuti nella Biblioteca del Congresso degli Stati Uniti
Petabyte (PB)	1.000 Terabyte 8 Petabyte: l'intero contenuto del web
Exabyte	1.000 Petabyte 1-2 Exabyte: tutta l'informazione generata nel corso dell'anno 1999

La crescita dei dati: nel 2006...

exabyte



...e le previsioni di Forbes per il 2025



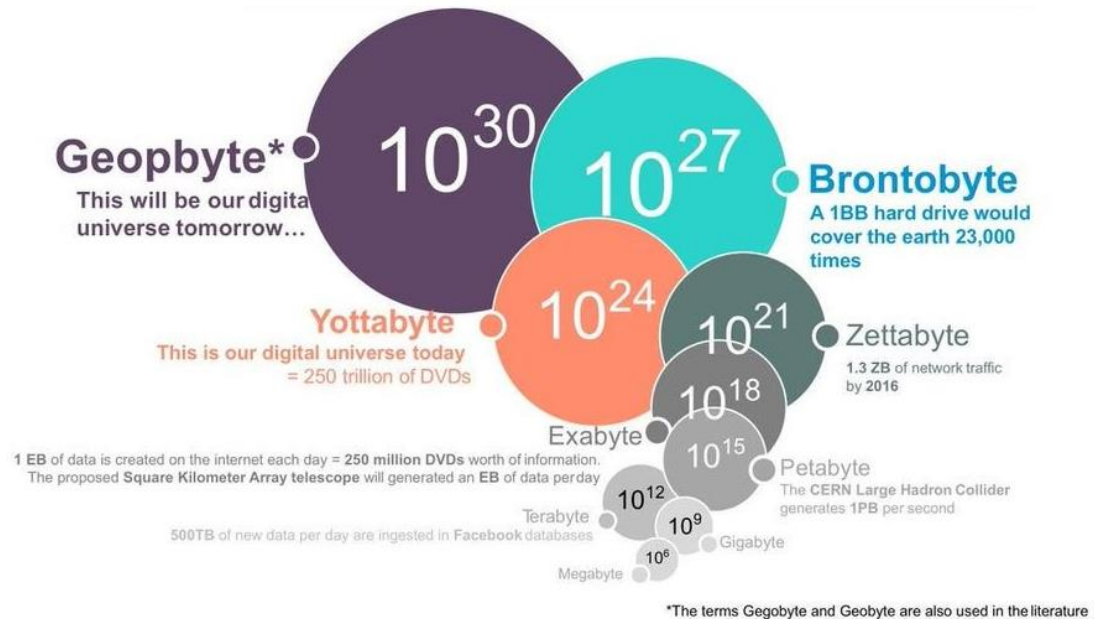
Source: UC Berkeley

L'Internet-minute nel luglio 2020



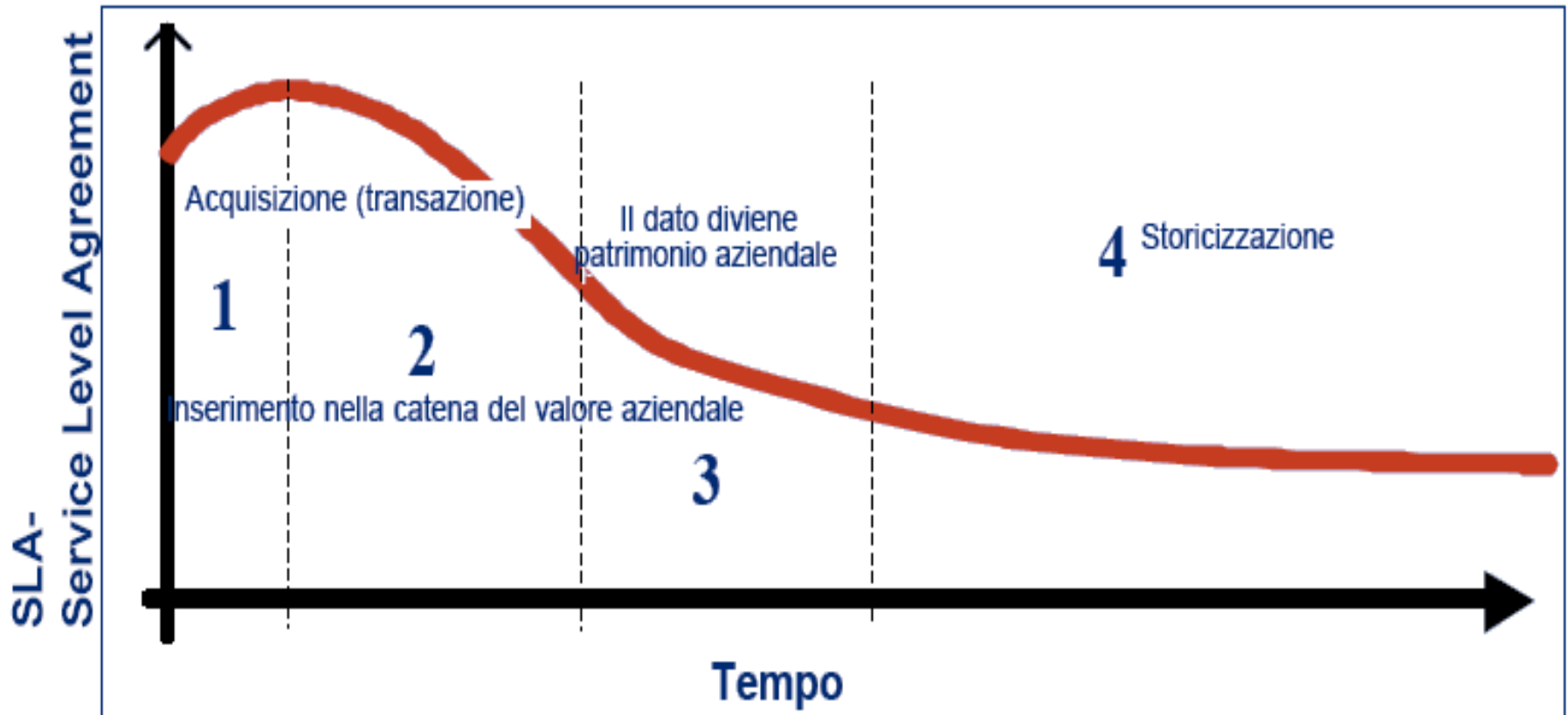
Il futuro della datosfera

- Megabyte (10^6)
- Gigabyte (10^9)
- Terabyte (10^{12})
- Petabyte (10^{15})
- Exabyte (10^{18})
- Zettabyte (10^{21})
- Yottabyte (10^{24})
- Brontobyte (10^{27})
- Geopbyte (10^{30})



Fonte: Simon Kuestenmacher

Ciclo di vita delle Informazioni



Fonte EMC

Overload d'informazione

- Aumento incontrollato di informazione complessivamente disponibile
- Comporta un maggiore sforzo per filtrare l'input estraendo le informazioni utili
- Abuso da parte dei processi organizzativi della possibilità di creare nuova informazione a partire da quella in input



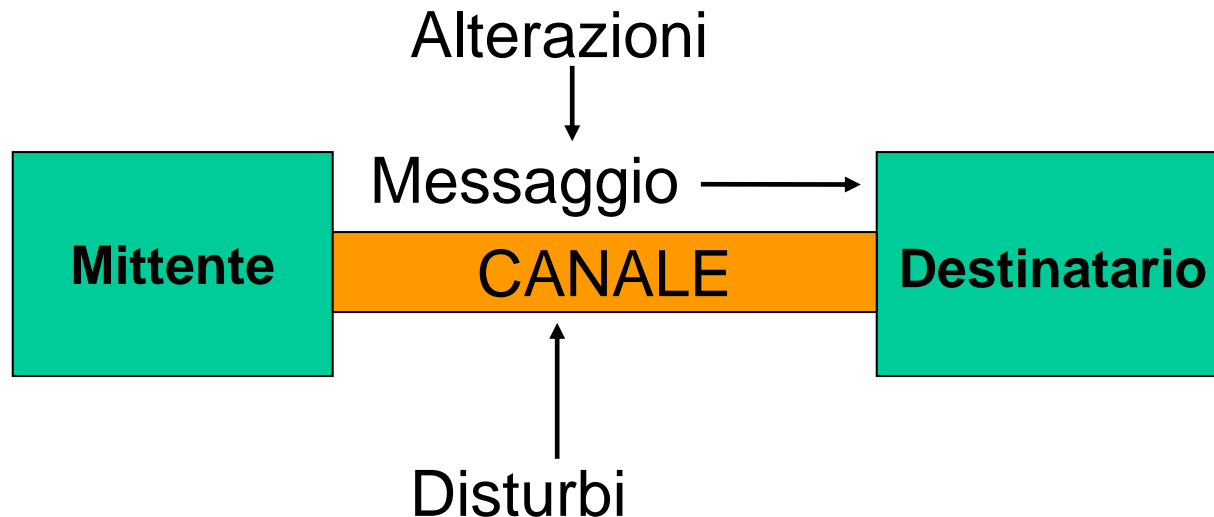
La comunicazione

Il modello teorico della comunicazione



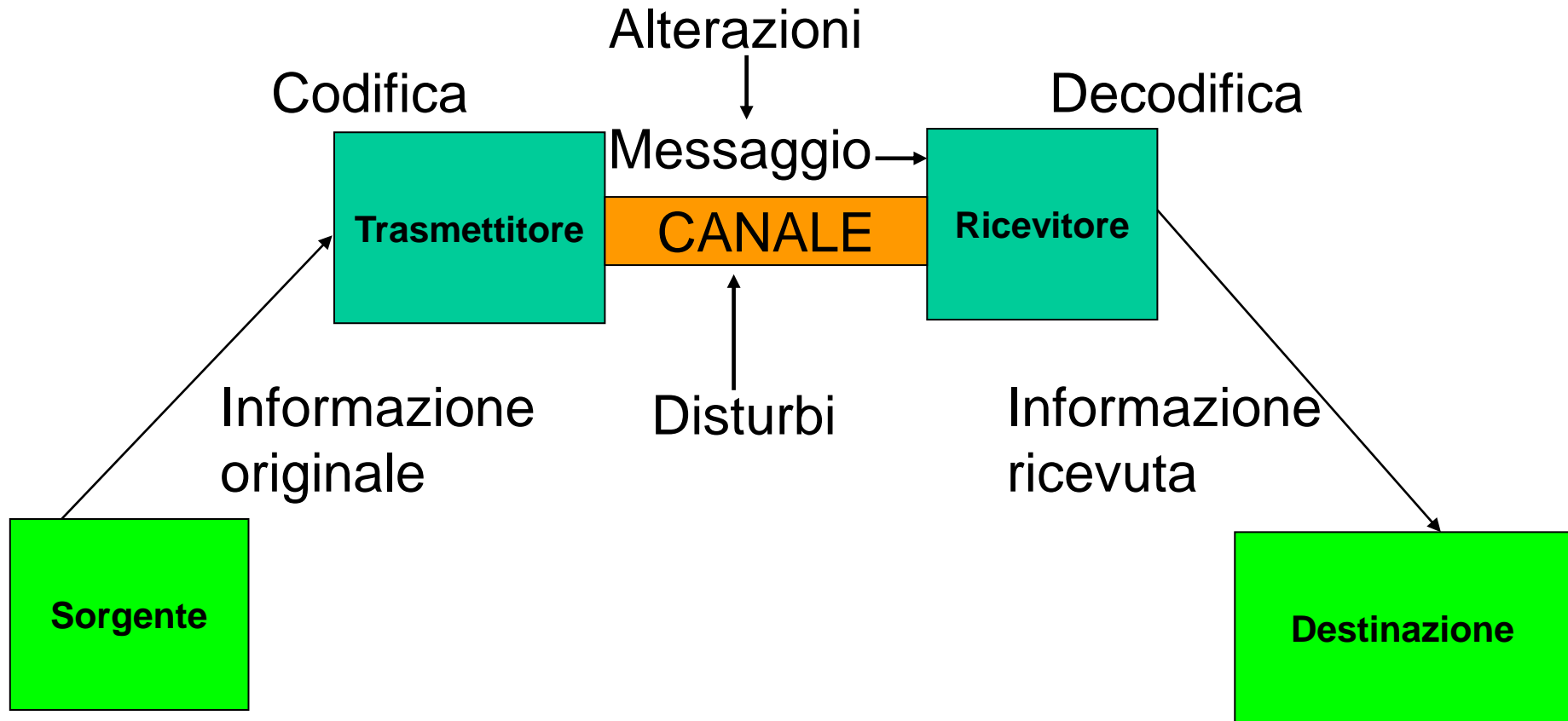
- **un messaggio**, che trasporta una informazione;
- **una sorgente o mittente**, che genera il messaggio e lo colloca sul canale;
- **un canale**, sul quale il messaggio viene trasportato;
- **un destinatario o destinazione**, che deve ricevere il messaggio.

La comunicazione nel mondo reale



- **Disturbi:** segnali “spuri” che si sovrappongono al messaggio
- **Alterazioni:** modifiche che il messaggio subisce per imperfezioni del canale

La comunicazione nel mondo reale - 2



La codifica

- E' l'operazione di trasformazione dell'informazione in un messaggio, trasmissibile sul canale attraverso opportuni segnali
- Prevede diversi passaggi parziali
 - Da informazione a dati elementari
 - Da dati elementari a loro rappresentazioni secondo il codice scelto
 - Da rappresentazioni a segnali fisici

Successo della comunicazione

- La comunicazione ha successo quando l'informazione contenuta nel messaggio arriva a destinazione
- Nella comunicazione analogica può essere un successo parziale
- Nella comunicazione digitale, di solito, o esiste successo o non esiste

La decodifica

- E' l'operazione di trasformazione inversa, che deve estrarre l'informazione da un messaggio, ricevibile dal canale attraverso opportuni segnali che lo rappresentano
- Prevede diversi passaggi parziali
 - Da segnali fisici a rappresentazioni
 - Da rappresentazioni a dati elementari secondo il codice scelto
 - Da dati elementari a informazione

L'Informazione al cervello umano

- La Informazione giunge al cervello umano e viene elaborata
- La percezione è un processo di decodifica
- Che richiede numerosi stadi
- Ciascuno con le sue regole ed i suoi **codici**

Il rapporto segnale/rumore

- Rapporto fra la potenza del segnale e la potenza del rumore
- Solo se il rapporto supera una determinata soglia la comunicazione può avere successo

Sommario

- La risorsa informazione
- La piramide DIKW
- Rappresentazione dell'informazione
- I flussi informativi entro l'azienda
- XML: l'esperanto elettronico
- Dati e informazione entro l'azienda
- La comunicazione