

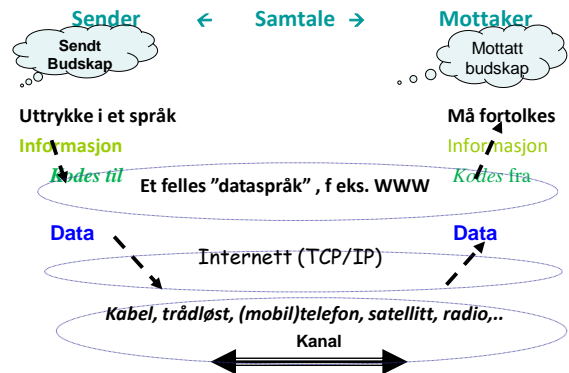
DRI 2020

Internet – fra TCP/IP til semantisk web

- Arkitektur –prinsipper
 - Litt mer om TCP/IP
 - Utviklingsfilosofien
- WWW – Hva er hemmeligheten
- Fra databaser til HTML og semantisk web

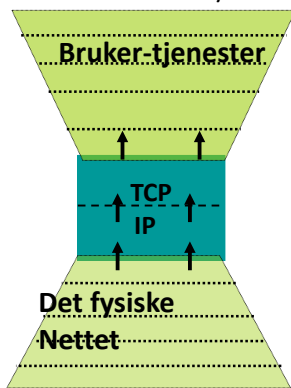
1

Hva er datakommunikasjon – en enkel modell



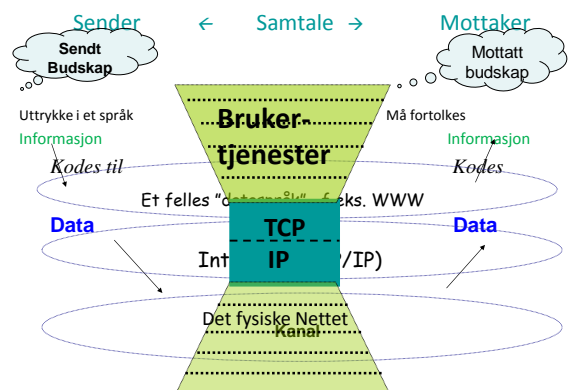
2

TCP/IP-modellen



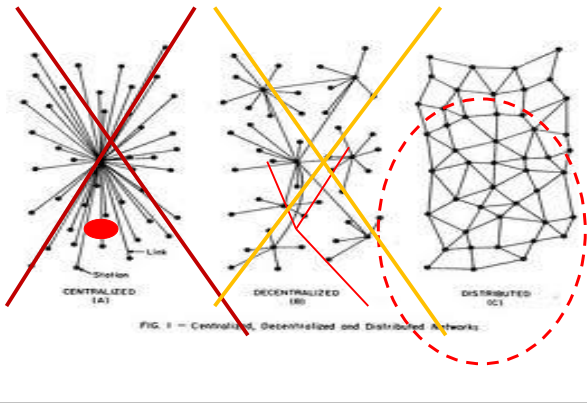
- Modularisert og lagdelt
- Vekt på enkelhet, minimal funksjonalitet og standarder
- Fleksibelt med veldefinerte grensnett, tjenestekvalitet og funksjonalitet
- Skille mellom protokoll og tjeneste.

Hva er datakommunikasjon – en enkel modell

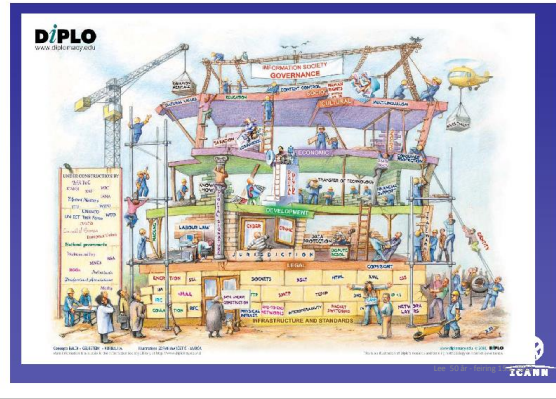


4

Hva slags nett ??



Bygging av Internett

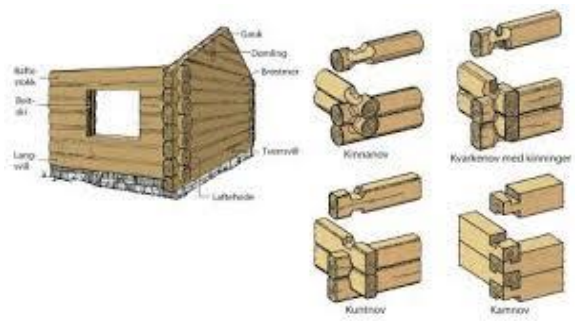


Bygge modulært – etter enkle grunnprinsipper



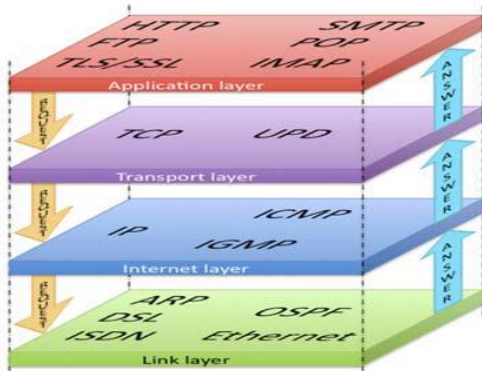
«Uendelige» muligheter og variasjoner, forutsatt at «grensesnitts»-regler overholdes

Lagdeling

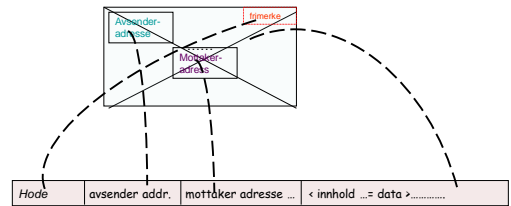


«Enkelt å bygge nye etasjer – eller tilleggsrom

Arkitekturen - skjematisk



Hva er en IP-pakke



10

IPv4-protokollen : Struktur og innhold

- Faste og valgfrie felter : Veldefinert betydning av hvert informasjonselement (Utenfor datadelen)
- Min. lengde 20 tegn (hodet) maksimum er 65535 tegn
- Struktur:

Id.<..>	Avsenderadresse	Mottakeradresse	<annet.>	...data	Kontrollsiffer
---------	-----------------	-----------------	----------	---------------	----------------

- Er syntaks viktig?
- Er formatet viktig?
- Er innholdet viktig

Se f eks. http://no.wikipedia.org/wiki/IPv4#IPv4_hode-format

11

Koding, syntaks og innhold

- Vi må bli enige om
 - Hvor mange bits i en byte (lengde : 6, 8, 16, ...)
 - Hvilke byte som representerer hvilke tegn (ASCII, 'UTF-8',...)
 - Språket jeg vil bruke (norsk, engelsk,
 - Struktur - organisering
 - Syntaks -grammatikken
 - Datamodellen (skjemaet)
 - "Navn" eller "Fornavn" + "etternavn"
 - Semantikken (hva opplysningene betyr)

Hvordan blir vi enige om alt dette ?

- Avtaler (mellom hvem, hvor mange,...)
- Standarder (hvilke, hvor universelle)

12

Ulike tilnæringer til utvikling av nettverk

Topp-styrt utvikling

- Målet er den rasjonelle og fullstendige, perfekte løsningen
 - Søker å inkludere alle gode ønsker og behov
- Spesifikasjonsdrevet: fossefallsmetoden
 - 'Lukket' verden og sentralisert kontroll
 - Monolittisk organisasjon
- Som regel implementert revolusjonært (alt på en gang)

Nedenfra-opp utvikling

- eksperimentell og iterativ
- Involverer flere brukere og miljøer
- Starter med minimumsløsninger
 - Akseptere kompromisser og søker pragmatiske løsninger

13

Semantisk web – Hva er problemet ?

- Semantic web is ...the Web of data (Tim Bernes Li, W3C)
- Tilbyr et felles rammeverk for å utveksle data på tvers av applikasjoner
- Hjelpemidler for å forstå informasjonsinnholdet i data, og derved muliggjøre bruk av felles data
- En ny "AI-bølge" : SW gjør datamaskinene smartere (gjennom å kunne fortolke og "forstå" hva slag informasjon (kunnskap) data representerer ?

DRI2010 H 08 24092008 Arild Jansen , AFIN

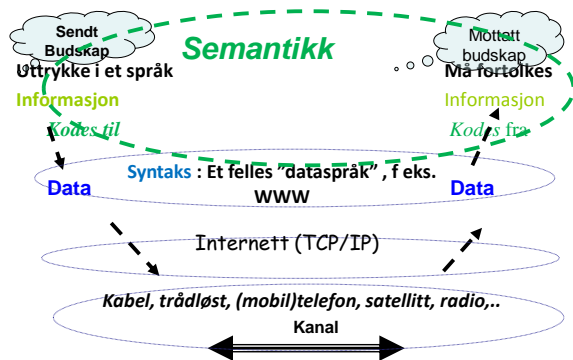
14

Syntaks, semantikk og grammatikk og ontologi

- **Syntaks** er læren om hvordan ord settes sammen til større enheter, dvs. setninger og fraser. Ulike språk har *ulike regler* for hvordan dette skjer på både med hensyn til ordstilling og syntaktiske avhengigheter
- **Semantikk** er den grenen av språkvitenskapen som studerer *ordenes betydning*, både enkeltvis og i sammenstilte strukturer (setninger).
- **Grammatikk** er studiet av reglene som gjelder et språk. Reglene som styrer et spesielt språk kalles språkets grammatikk

15

Hva er datakommunikasjon – en enkel modell



16

Hva forstår dere med dette dette?

- OLA HANSEN
- HANSEN, KARI
- OLA BJØRN
- 11032003
- 031103
- 11/03/03
- 13/03/45
- 0110111010111011 (= 67273)

Hva kreves for å forstå (få mening ut av dette)?

17

IP-Adressering og Url're

- Internett har et felles, globalt adresse-rom, såkalt IP-adresse: 32bits adresse (i dag)
 - Siffer binær form: x.y.z.u, hver et tall 0-255, tilsammen $256*256*256*256 - 1 = 2^{32} - 1$: (hvorfor?)
 - Eks. 129.240.178.69
 - Symbolsk (tekstlig) form) : f eks. f eks, nrk.no, afin.uio.no, google.com, wikipedia.org,
 - Kobling mellom IP-adresse og symbolsk adresse skjer ved DNS (Domain Name Server)
- URL: Uniform Resource Locator : nøyaktig adresse til et dokumentnavn, f eks.

<http://www.uio.no/studier/emner/jus/afin/DRI1002/v04/>

DRI 2020 Arild Jansen , AFIN

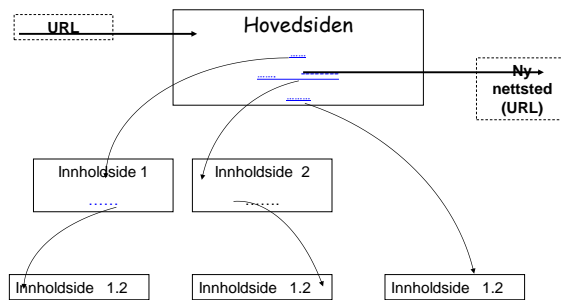
Organisering av informasjon (data) på WWW

Noen hovedbegreper:

- Hjemmeside (home page): Hovedsiden (ofte kalt startside for et nettsted, (web-site)
 - Eks: <http://www.uio.no/>, <http://www.afin.uio.no/>
 - En hjemmeside identifiseres ved en URL
- Lenke : peker til et annet dokument
- Hypertekst : tekst som inneholder lenker til andre dokumenter (*URL'er)

DRI 2020 Høst 13 Arild Jansen , AFIN

Strukturen på ett nettsted



Se på <http://www.jus.uio.no/ifp/om/organisasjon/afin/>

DRI 2020 Høst 13 Arild Jansen , AFIN

Hvordan organiseres informasjon ?

- Informasjon er ofte organisert i form av **'dokumenter'** (et kapittel, en artikkel, en rapport, brev, avisinnlegg, dvs. et avgrenset stykke tekst....).
- Dokumenter lagres digitalt i form av **filer**

Et digitalt dokument omfatter tekst og eventuelt bilder, men kan også ha tilknyttet video eller musikk/tale-innslag ved hyperlenker.

Noen ulike binære representasjonsformer

- Tekstdokumenter
 - Ren tekst , (.txt)
 - Word-dokument (.doc)
 - RTF :rich text format (.rtf)
 - PDF-format (Portable Document Format) : .pdf
 - HTML-format (.html)
- Bilder :
 - GIF (Graphics Interchange Format)
 - JPEG (Joint Photographic Experts Group)

Eksempel på enkel (tabellbasert) database

Arild Johan Jansen, Hofstadgate, 1384 Asker
Dag Wiese Scha **Feltes** d Lovenskiolds v, 0760 Oslo

Pnr	Efternavn	Fornavn	Gate/veinavn	Postnr	Poststed
002	Jansen	Arild Johan	Hofstadgata	1384	Asker
01	Schartum	Dag Wiese	H. Lovenskiold vei	0760	Oslo

Poster
Poster (record) En 'linje' i tabellen som inneholder verdier i de enkelte feltene

Primær-nøkkel
entydig identifikasjon for alle poster

De norske navn på Hall of Fame

