

Verkostoanalyysiä

Case: TTY:n kurssien esitiedot 2010-2011

Verkostoanalyysi 2011
TTY

Jarno Marttila
Tampereen teknillinen yliopisto
Hypermedialaboratorio

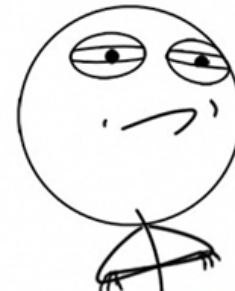
Sisälllys

- Taustaa
- Informaation visualisointiprosessi Casen tapauksessa
- Gephi
- PageRank
- Esimerkki

Mistä liikkeelle ?

- Data-edellä lähestymistapa
 - Mielenkiintoinen datasetti
- Tavoite/idea
 - Opintojaksojen mallintaminen verkostoksi
 - Tunnuslukujen laskenta
 - Laskennallisesti merkittävimpien kurssien löytäminen

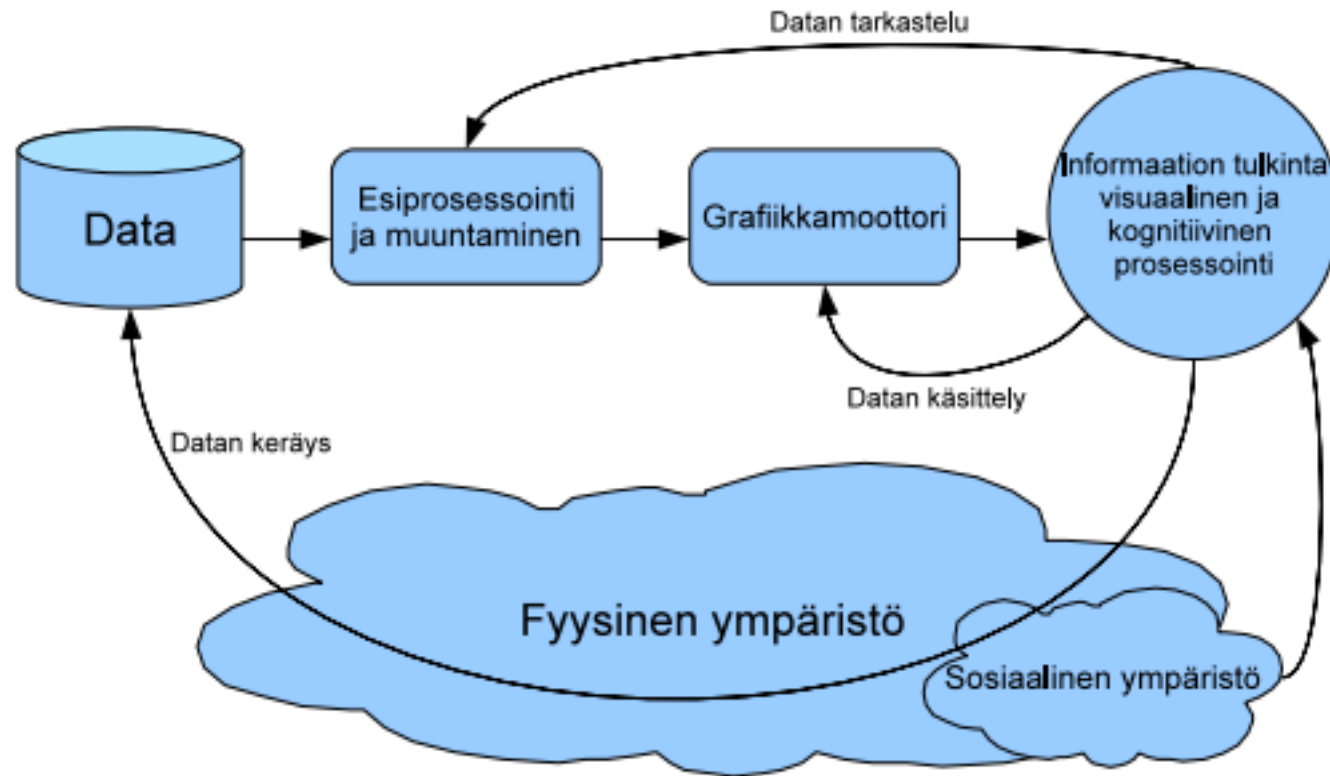
NETWORK ANALYSIS



CHALLENGE ACCEPTED

meme-generator.net

Informaation visualisointiprosessi



Ware 2004

Data

Opinto-opas 2010-2011

MATHM-37200 Hypermediajärjestelmät, 3 op Hypermedia Systems

Vastuuhenkilö

Jaakko Salonen

Opetus

Opetusmuoto	P1	P2	P3	P4	Kesä	Toteutuskerrat	Luentoajat ja -paikat
Luennot		3 h/vko				Toteutus 1	Keskiviikko 12 - 15, S1

Suoritusvaatimukset

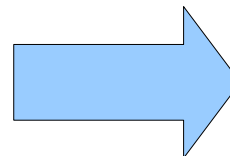
Vuosittain määritettävää kirjallisia tehtäviä sekä käytännöllisiä ja soveltavia laboratorio- ja harjoituksia. Osasuoritusten pitää liittyä samaan toteutuskertaan.

Osaamistavoitteet

Opintojakson suoritettuaan opiskelija ymmärtää hypermedia-, sisällönhallinta-, ja -julkaisujärjestelmien peruskäsitteistä sekä niiden toiminnan perusteet. Hän oppii lisäksi analysimaan järjestelmien piirteitä ja arvioimaan niiden soveltuvuutta erilaisiin käytötapauksiin. Lisäksi opiskelija oppii rajatusti järjestelmien käyttöönottoa ja konfigurointia.

Sisältö

Sisältö	Ydinaines	Täydentävä tietämys	Erityistietämys
1.	Hypermediajärjestelmien historia, mallit ja visioita		
2.	Sisällönhallintajärjestelmien (CMS) ja julkaisujärjestelmien keskeiset periaatteet, erityispiirteet WWW:ssä.	Sisällönhallintajärjestelmien ominaisuudet	Portal-järjestelmät ja Enterprise Content Management
3.	Web-arkkitehtuuri ja -integraatio	Web-pohjainen autentikointi ja auktorisointi, Web-syötteet ja takaisinvirtaustekniikat.	Hajautetut ja Peer-to-peer -pohjaiset tieto- ja palvelusovellukset
4.	Uudet vuorovaikutusmallit ja sosiaalinen media	Wikit, blogit ja verkkoyhteisöt	
5.	Web-analytiikka.	Käytön seuranta, liiketietoanalyysi, tiedon louhinta ja visuaalisointi osana käytön seurantaa	Tietoteolliset tutkimusmenetelmät käytön seurannassa



Opintojakson arviointi

Opintojakson tarkat arviointiperusteet määrittyvät toteutuskertakohtaisesti asetettujen osasuoritteiden perusteella.

Arviointiluokitus:

Opintojaksoilla käytetään numeerista arviointiluokituskaikkoa (1-5)

Osasuoritukset:

Osasuoritusten pitää liittyä samaan toteutuskertaan

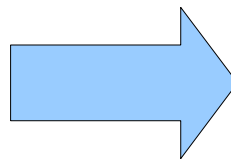
Esitietovaatimukset

Opintojakso	P/S	Selite
MATHM-37100 Johdatus hypermediaan	Pakollinen	

[Esitietovelje](#) (Vaatii kirjautumisen POPiin)

Esiprosessointi

	A	B	C	D	E	F	G	H
1	KOODI	NIMI	LAAJUUS	LAAJUUS_MAX	LAITOS	ESITETOKOODI	ESITETO	PAKOLLINEN/SUOSITELTAVA
2	ACI-11010	Automaatiotekniikan työkurssi	3	3	ASE	ACI-20030	Säädön suunnittelu	pakollinen
3					ASE	ACI-20020	Digitaalinen säätö	suositeltava
4					ASE	ACI-20010	Automaatiotekniikan	pakollinen
5	ACI-20030	Säädön suunnittelu	7	7	ASE	ACI-20010	Automaatiotekniikan	suositeltava
6	ACI-21110	Panosprosessien automaatio	4	4	ASE	ACI-31040	Automaatiolaitteet ja	suositeltava
7					ASE	ACI-10020	Automaation dokumentointi	suositeltava
8	ACI-31040	Automaatiolaitteet ja -verkot	7	7	ASE	ACI-20010	Automaatiotekniikan	suositeltava
9					ASE	ACI-20010	Automaatiotekniikan	suositeltava
10	ACI-32020	Automaation reaaliaikajärjestelmät	7	7	ASE	OHJ-1150	Ohjelmointi II	suositeltava
11					ASE	ACI-31040	Automaatiolaitteet ja	suositeltava
12	ACI-32040	Automaation ohjelmistokomponentit ja sovelluspalvelut	7	7	ASE	ACI-32020	Automaation reaaliaikajärjestelmät	pakollinen
13	ACI-42086	Optimal and Robust Control System Design with Matlab	7	7	ASE	ASE-1250	Järjestelmien ohjaus	suositeltava
14					ASE	ACI-20010	Automaatiotekniikan	suositeltava
15					ASE	ASE-1256	Introduction to Control Systems	suositeltava
16					ASE	ASE-1257	Introduction to Control Systems	suositeltava
17	ACI-51416	Assignment in Microfluidics	1	5	ASE	ACI-51406	Microfluidics	pakollinen
18	AHT-1200	Taide 2	8	8	ARK	AHT-1120	Taide 1	pakollinen
19	AHT-2200	Arkkitehtuurin ammattikurssi	14	14	ARK	AHT-2100	Arkkitehtuurin perusteet	pakollinen



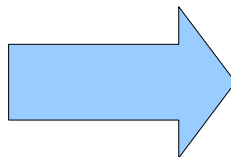
Esiprosessointi ja muuntaminen

```
for i in range(len(dict_list)):
    #Iterate through dictionary
    #print i
    #print dict_list

    if dict_list[i]["KOODI"] != "":

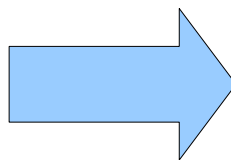
        #solmun yhteyden kohde on KOODI-niminen kurssi
        target = dict_list[i]["KOODI"]
        #lisataan kohdesolmu solmulistaan
        #node_list.append(add_node(dict_list[i]["KOODI"]))
        node_list.append(add_node_with_att(dict_list[i]["KOODI"], dict_list[i]["LAAJUUS"], dict_list[i]["LAAJUUS_MAX"], dict_list[i]["LAITOS"]))
        #add_att_value()
        #lisataan kohdesolmu tarkistuslistaan
        added_node_list.append(add_node(dict_list[i]["KOODI"]))
        if dict_list[i]["ESITIETOKOODI"] != "":
            edge_list.append(add_edge(dict_list[i]["ESITIETOKOODI"], dict_list[i]["KOODI"], dict_list[i]["PAKOLLINEN/SUOSITELTAVA"]))
            if added_node_list.count(dict_list[i]["ESITIETOKOODI"]):
                node_list.append(add_node(dict_list[i]["ESITIETOKOODI"]))
        elif dict_list[i]["KOODI"] == "":

            if dict_list[i]["ESITIETOKOODI"] != "":
                edge_list.append(add_edge(dict_list[i]["ESITIETOKOODI"], target, dict_list[i]["PAKOLLINEN/SUOSITELTAVA"]))
                if added_node_list.count(dict_list[i]["ESITIETOKOODI"]):
                    node_list.append(add_node(dict_list[i]["ESITIETOKOODI"]))
            elif dict_list[i]["ESITIETOKOODI"] == "":
                edge_list.append(add_edge(dict_list[i]["ESITIETO"], target, dict_list[i]["PAKOLLINEN/SUOSITELTAVA"]))
                if added_node_list.count(dict_list[i]["ESITIETO"]):
                    node_list.append(add_node(dict_list[i]["ESITIETO"]))
```

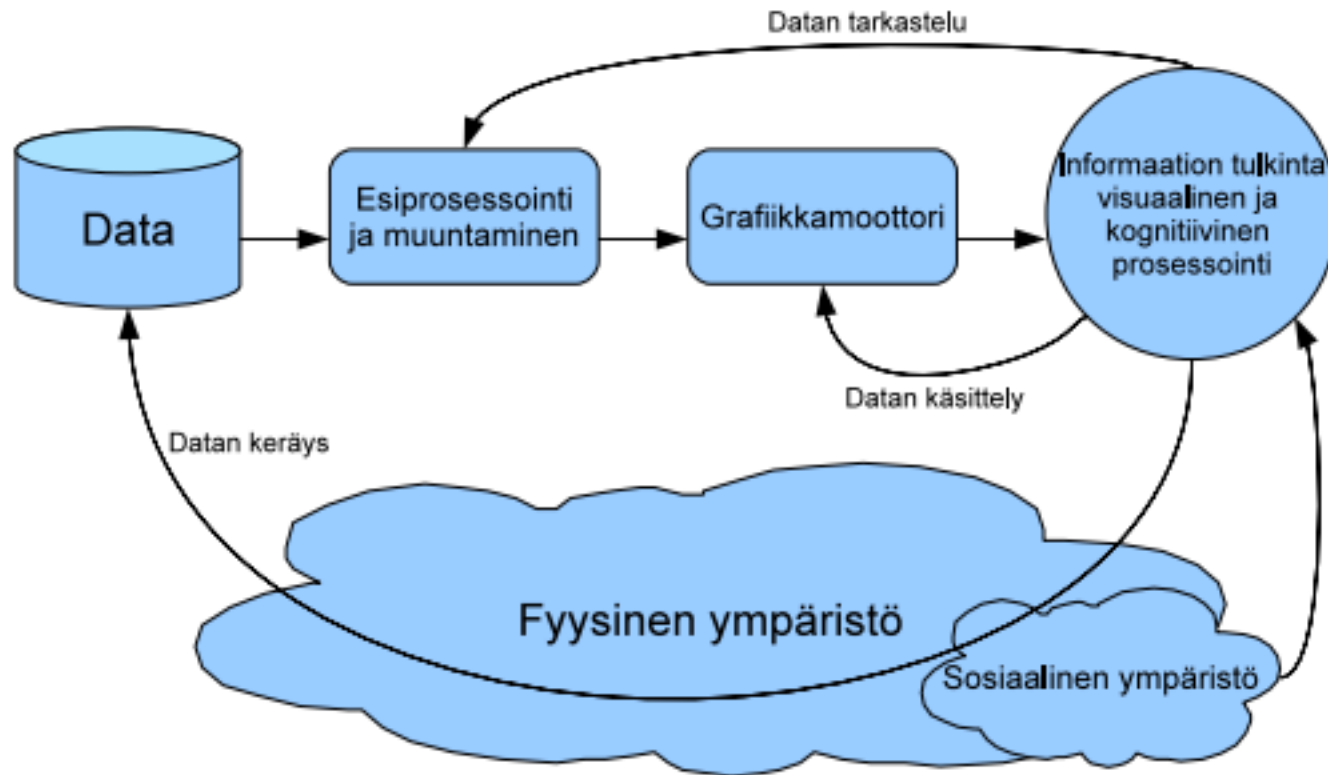


Verkostomalli

```
1 <gexf xmlns="http://www.gexf.net/1.1draft"
2     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3     xsi:schemaLocation="http://www.gexf.net/1.1draft
4     http://www.gexf.net/1.1draft/gexf.xsd"
5     version="1.1">
6 <graph mode="static" defaultedgetype="directed">
7 <nodes>
8 <node id="ACI-11010" label="ACI-11010"/>
9 <node id="ACI-20030" label="ACI-20030"/>
10 <node id="ACI-21110" label="ACI-21110"/>
11 <node id="ACI-31040" label="ACI-31040"/>
12 <node id="ACI-32020" label="ACI-32020"/>
13 <node id="ACI-32040" label="ACI-32040"/>
14 <node id="ACI-42086" label="ACI-42086"/>
15 <node id="ACI-51416" label="ACI-51416"/>
16 <node id="AHT-1200" label="AHT-1200"/>
17 <node id="AHT-2200" label="AHT-2200"/>
```

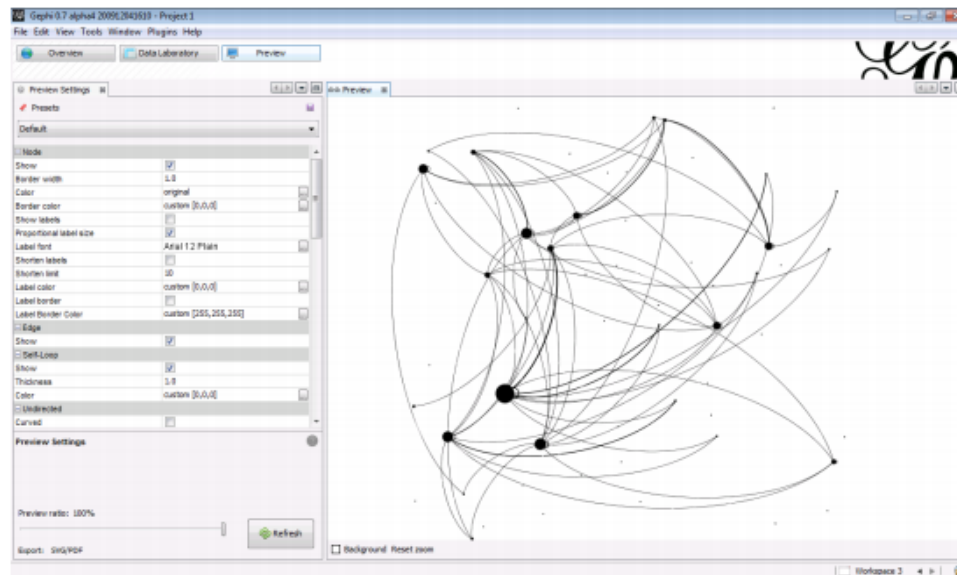


Informaation visualisointiprosessi



Visualisointi ja tunnuslukujen laskenta

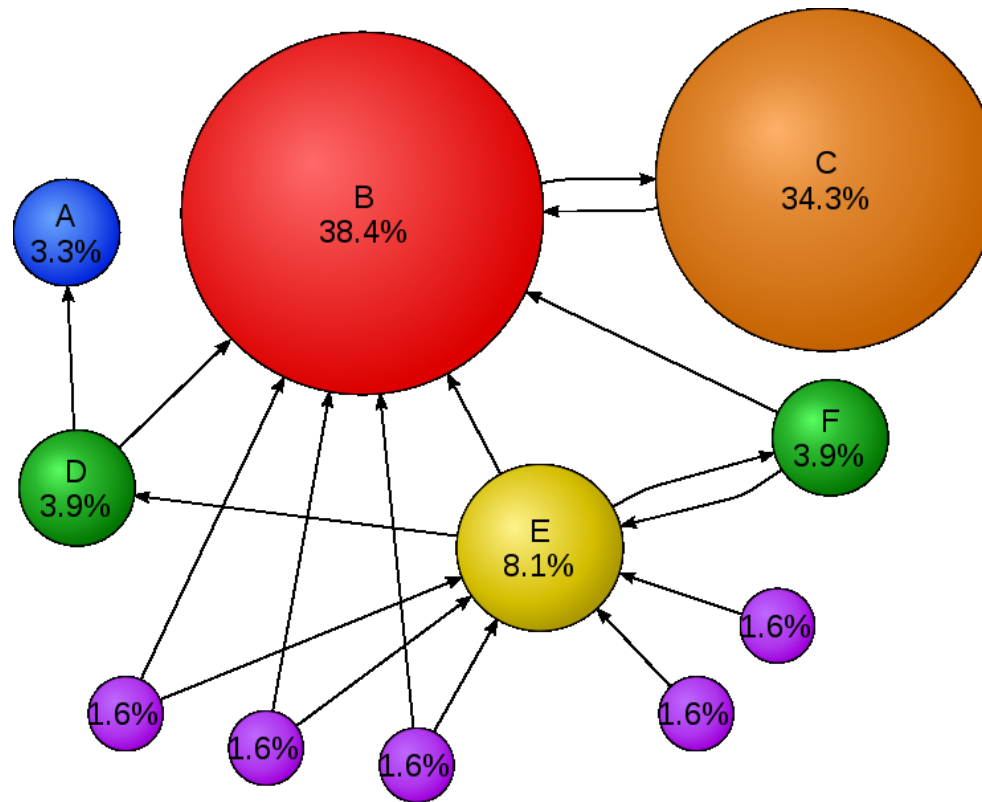
- Visualisointi ja laskenta toteutettu Gephillä
- Gephi
 - Avointa lähdekoodia
 - Graafien analysointi- ja verkostojen visualisointiohjelma



PageRank

- Larry Pagen ja Sergey Brinin kehittämä algoritmi
- Web-sivustojen arvottamiseen niiden muodostamien linkkiverkostojen perusteella
- Numeerinen arvo, esittää kuinka tärkeä toimija on

PageRank

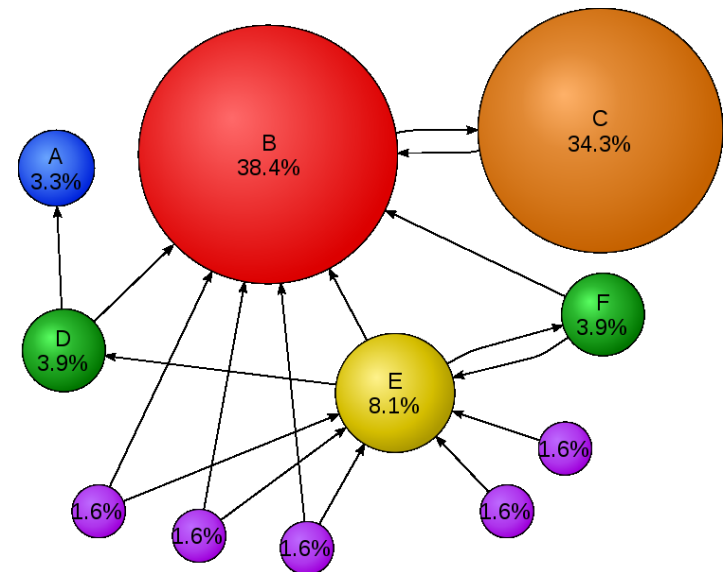


PageRank

- Toimijan A PageRank määritellään yhtälöllä

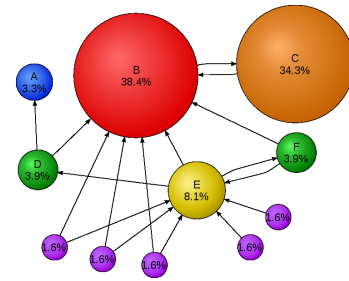
$$PR(A) = (1 - d) + d * \left(\frac{PR(t_1)}{C(t_1)} + \dots + \frac{PR(t_n)}{C(t_n)} \right).$$

- Yhtälö on alkuperäinen Googlen julkaisema PageRank algoritmi



PageRank

- Toimijoiden äänillä tietty painoarvo
 - Riippuvainen PageRankista
- Äänestäessä välitetään eteenpäin osa PageRankista
 - Jaetaan kaikkien toimijoiden kesken joita äänestetään
- Toimijan välittämä PageRank on suhteellinen lähtevien yhteyksien määrään
 - Äänestettävälle toimijalle äänestävä toimija jonka PR = 4 ja 5 äänestyskohdetta tärkeämpi kuin, toimija jonka PR = 10 ja 100 äänestyskohdetta
- Toimijoiden PageRank ei vähene vaikka ne välittävät äänestämällä omaa PageRankiaan eteenpäin



$$PR(A) = (1 - d) + d * \left(\frac{PR(t_1)}{C(t_1)} + \dots + \frac{PR(t_n)}{C(t_n)} \right).$$

Esimerkki



Kiitos!
Kysymyksiä?

- <http://www.amazon.com/Information-Visualization-Second-IntWare> 2004
- <http://ilpubs.stanford.edu:8090/422/>
- <http://dspace.cc.tut.fi/dpub/handle/123456789/6835>

ation-Visualization-Second-Interactive-Technologies/dp/1558608192

422/