



LÄPINÄKYVYYDESTÄ

Anna-Mari Rusanen

Kognitiotiede

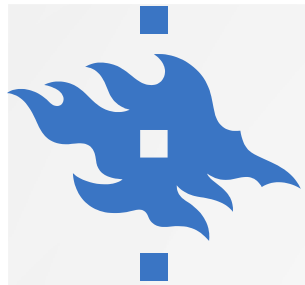
Digitaalisten ihmistieteiden osasto

Helsingin Yliopisto

JulkiCT

Valtionvarainministeriö





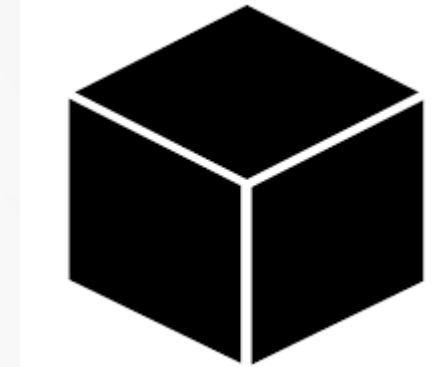
MIKSI LÄPINÄKYVYYS - KESTOTEEMA

- Ns. eettiset ohjeistukset: 75/84 mainitsee (Jobin & al, 2019)
- Useita eri muotoiluja
- Tausta: ns. musta laatikko- keskustelu
- mm. teknologian ja tieteen tutkimus 1990- luvun lopulla
- ”Tekoälybuumi”: ML ja DNN-sovellukset 2012->, laaja-alainen soveltaminen 2015 (?)>
- Olennaista: Järjestelmien monimutkaistuminen

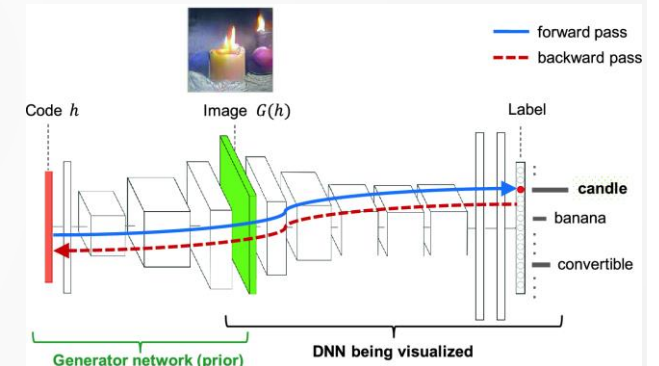




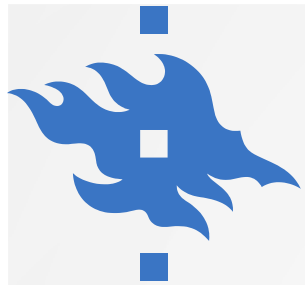
LÄPINÄKYVYYS – MITÄ/KENELLE



- Humphreys (2009):
- **Olemuksellinen ("periaatteellinen")**
- Ei voida "nähdä" sisälle (edes periaatteessa)
- Esim. tietyt DNN- sovellukset, ehkä myös tietty "laskenta"
- **Tiedollinen ("hahmotettavuus")**
- Kuinka läpinäkyvä X on jollekulle toimijalle?
- Mitä edellytyksiä toimijalla on hahmottaa X *läpinäkyvänä*?
- Miten tehdä järjestelmä läpinäkyväksi toimijalle?

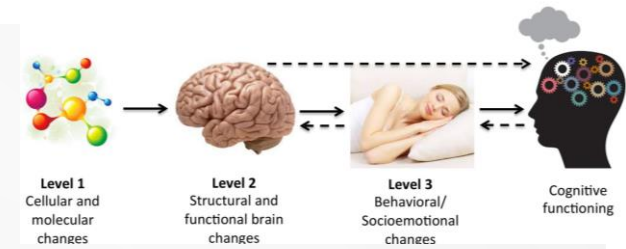


Nguyen & al, 2019



MIKÄ VAIKUTTAÄ TIEDOLLISEEN LÄPINÄKYVYYTEEN?

- KOGNITIONTUTKIJAN NÄKÖKULMA

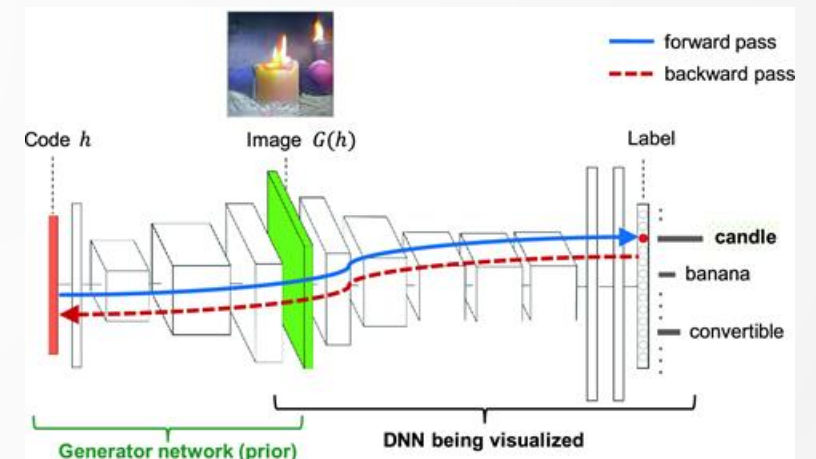


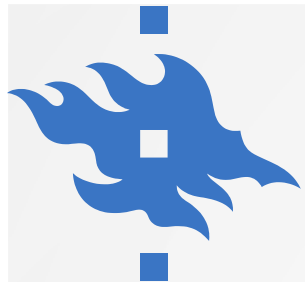
- Ihmisen lajityypillinen tiedonkäsittely (toimintatapa, ”muoto”, rajoitukset)
 - esim. muisti-, havainto-, tarkkaavaisuus-, päättely-, kieli- yms.
 - Pääsääntö: Samankaltaisuus/käännettävyys ihmiselle ”tutulle” tiedonkäsittelylle
- Yksilöllinen vaihtelu
- Kulttuuri (kieli, käsitteistö, myös kulttuuriset koodistot tms.)
- Oppiminen, asiantuntijuus
 - Algoritminen lukutaito
 - Asiantuntijuus
 - Olennaista: luonne ja aste erilainen eri toimijoilla/eri tehtävissä



LÄPINÄKYVYYDEN HAASTEITA

- Eri toimijoilla eri vaatimukset läpinäkyvyydestä
 - Juridiset/eettiset normistot, toisaalta asiantuntijuuskysymykset
- Tietojenkäsittelytieteilijä:
 - Mitä läpinäkyvyys tarkoittaa?
 - Miten lisätä läpinäkyvyyttä tietoteknisesti?
 - Esim. dissektiomenetelmät
- Käyttäjä:
 - Ei ymmärrä sovelluksia algoritmisesta näkökulmasta
 - Mikä oikea läpinäkyvyyden muoto/aste? Ymmärrettävyys?





LÄPINÄKYVYYDESTÄ "TIEDÄT, MITÄ TEET" - PERIAATTEESEEN

- Miten lisätä "hahmotettavuutta"?
- Usein: tehdään järjestelmät kognitiivisesti "tutummiksi"
- Esim. mielen teoria
- Visualisointimenetelmät tms.
- Algoritminen lukutaito, koulutus

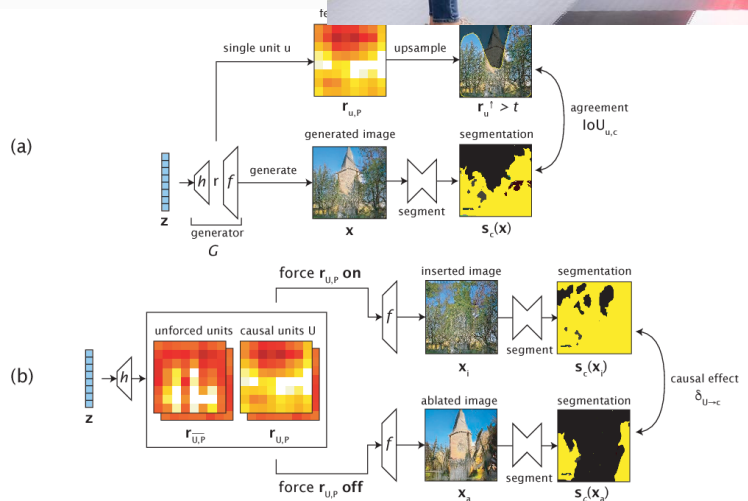
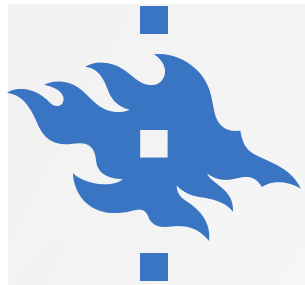


Figure 2: Measuring the relationship between representation units and trees in the output using (



MIKSI LÄPINÄKYVYYDESTÄ PUHUTAAN?

- Aiempaa monimutkaisten järjestelmien soveltaminen
 - Esim. automaattinen päätöksenteko viranomaistoiminnassa
 - Erityisvaatimukset (normit, lainsäädäntö): esim. päättely tukeutumisesta
 - ”Jos tukeudumme X:n -> toiminnan oltava läpinäkyvää -> jäljitettävää...”
- Nostaa kysymyksen läpinäkyvyydestä usealla tasolla:
 - Eri toimijoilla erilaiset vaatimukset läpinäkyvyydestä
 - Juridinen, eettinen, käyttäjän ominaisuudet (ml. Inklusiokysymykset), toiminto...
 - Myös ”algoritmisen lukutaidon vaatimus” vaihtelee
 - Koodarin läpinäkyvyys vs. käyttäjän/viranomaisen/lainsäätäjän läpinäkyvyys



60° 10 1.2 N, 24° 57 18 E

anna-mari.rusanen@helsinki.fi
anna-mari.rusanen@vm.fi