



Vihreä data: Tallennus ja käsittely

Toni Taipalus
Tampereen yliopisto
toni.taipalus@tuni.fi



Database Systems -tutkimusryhmä
<https://dbs-research.github.io>



Hiilijalanjälki: prosessointi

“The cloud now has a greater carbon footprint than the airline industry. A single data center can consume the equivalent energy of 50,000 homes.”^[1]

40 kg CO₂ per year per terabyte (TB) of data stored.^[2]

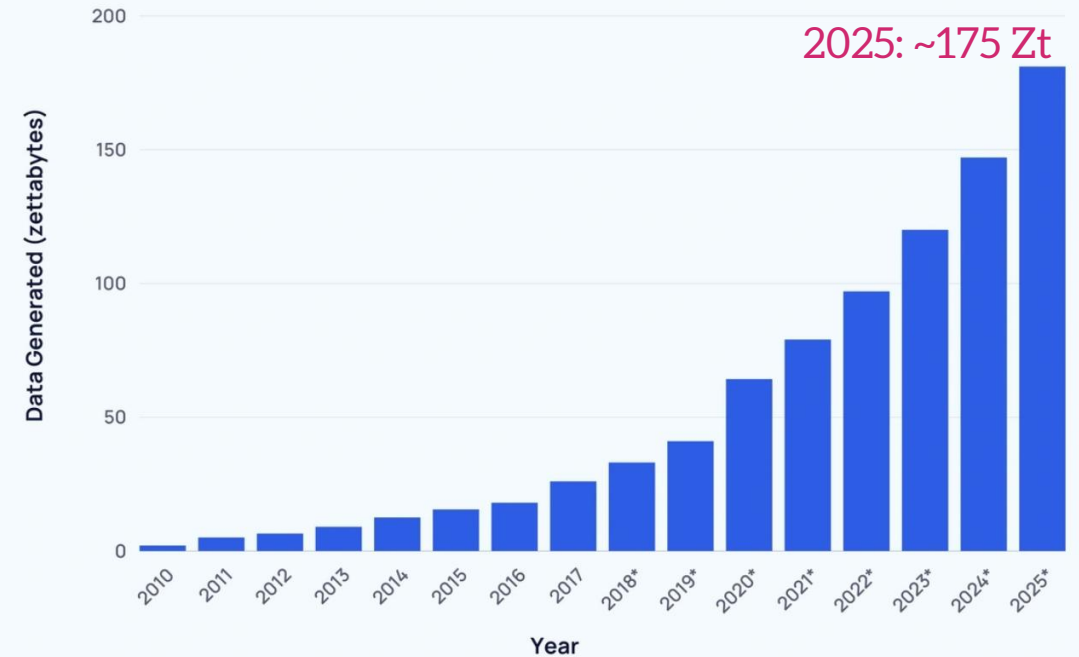
[1] <https://thereader.mitpress.mit.edu/the-staggering-ecological-impacts-of-computation-and-the-cloud/>

[2] <https://greenly.earth/en-gb/blog/industries/what-is-the-carbon-footprint-of-data-storage>

Hiilijalanjälki: data

- 2025: ~175 zettatavua luotua dataa (1 Zt = 1,000,000,000 Tt).
- BluRay-levyille tallennettuna pino yltää Kuuhun...
- ...ja takaisin...
- 10 kertaa.
- 1 Tt -kiintolevyille tallennettuna, jotka ovat vierekkäin:
- 35,000 jalkapallokenttää, tai
- ~2,600 km² (n. yksi Luxemburg)

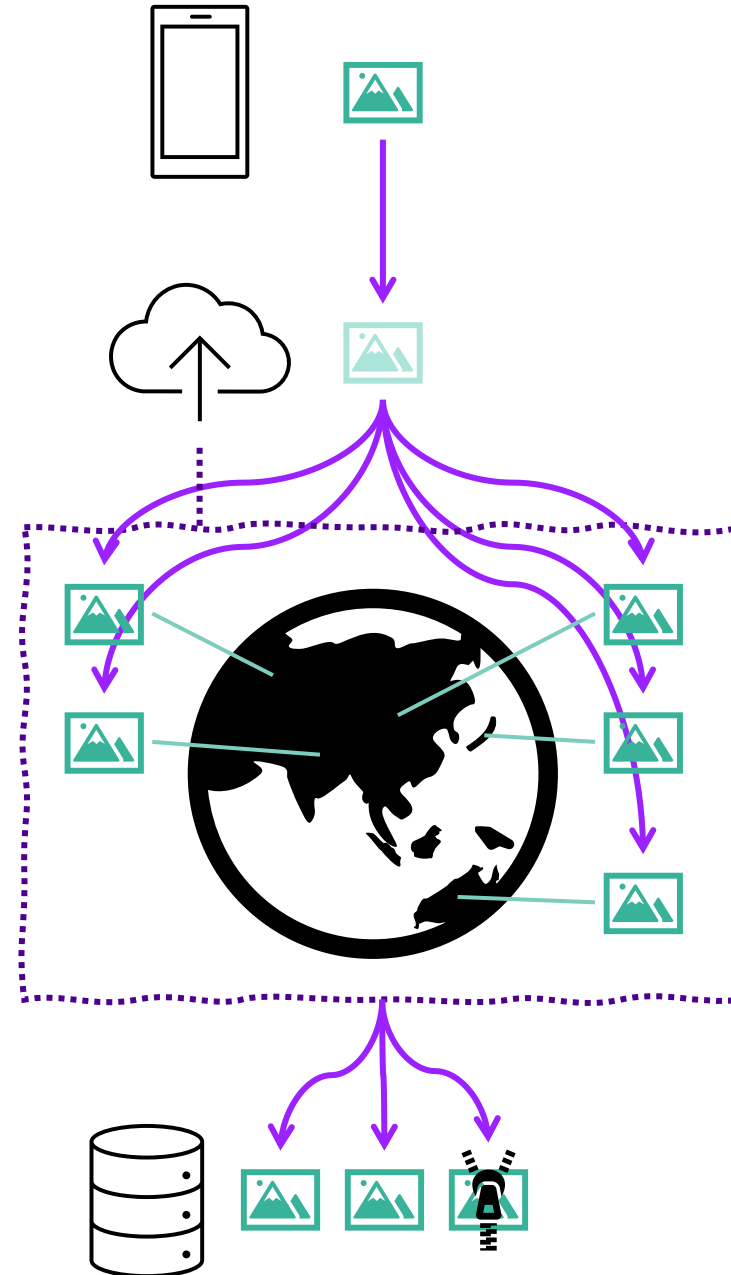
Global Data Generated Annually



Kategoria	% energian käytöstä
Laskenta	50-60%
Infrastrukturi ja jäähdytys	20-30%
Datan tallennus	15-20%
Verkko	5-10%

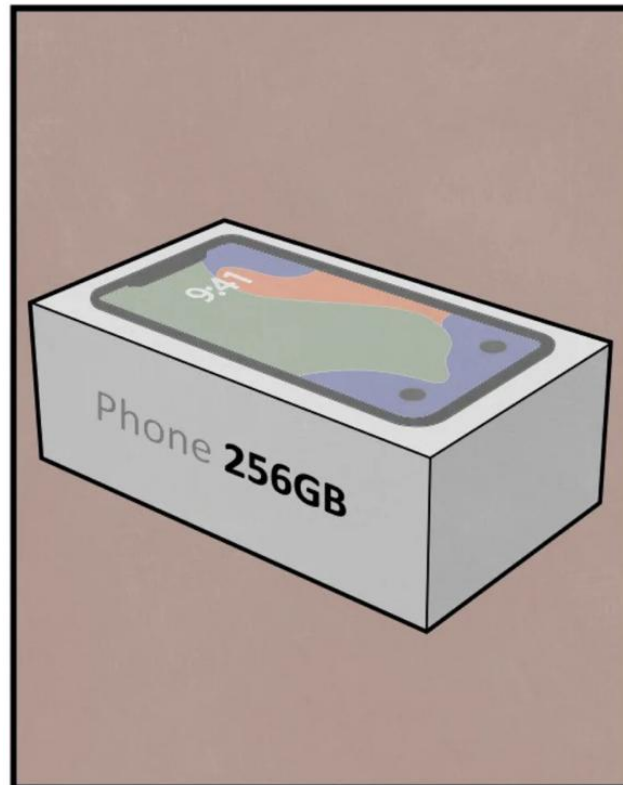
Yksi valokuva älypuhelimella

- Yksi valokuva ~4 Mt
- Kopioituna pilveen, yhteensä 8 Mt
- Hajautettuna viiteen eri datakeskukseen, yhteensä 24 Mt
- Varmuuskopiot, yhteensä 32 Mt
- 1,000 valokuvaa = 32 Gt
- 32 Gt = 1,2 kg CO₂ / vuosi, ainoastaan datan tallentamisesta.

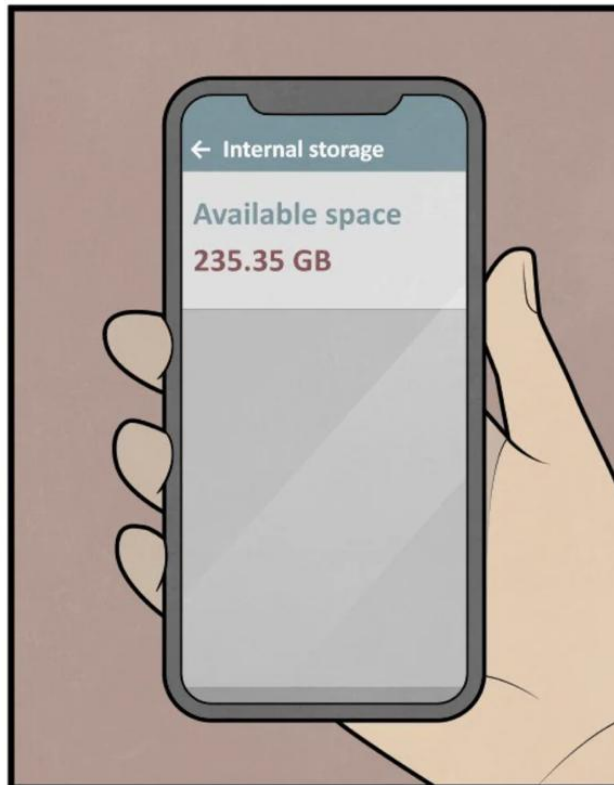


Käyttämätönkin tallennustila kuormittaa

YES,

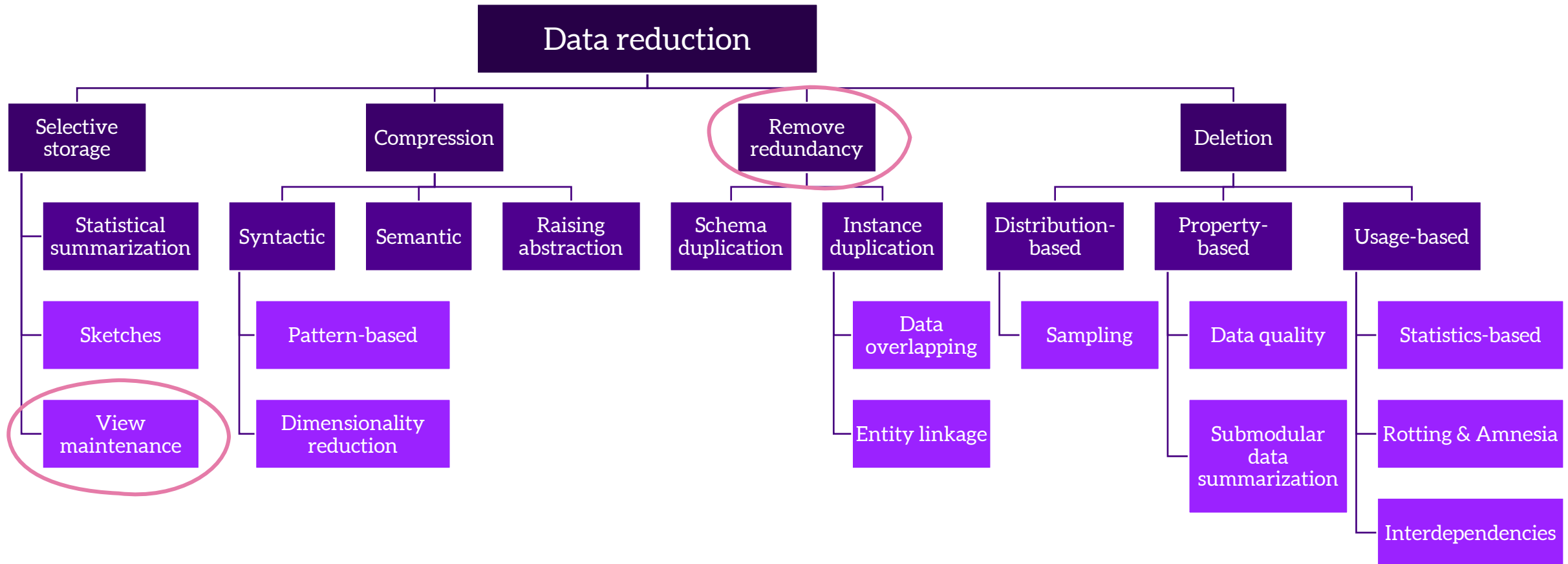


BUT

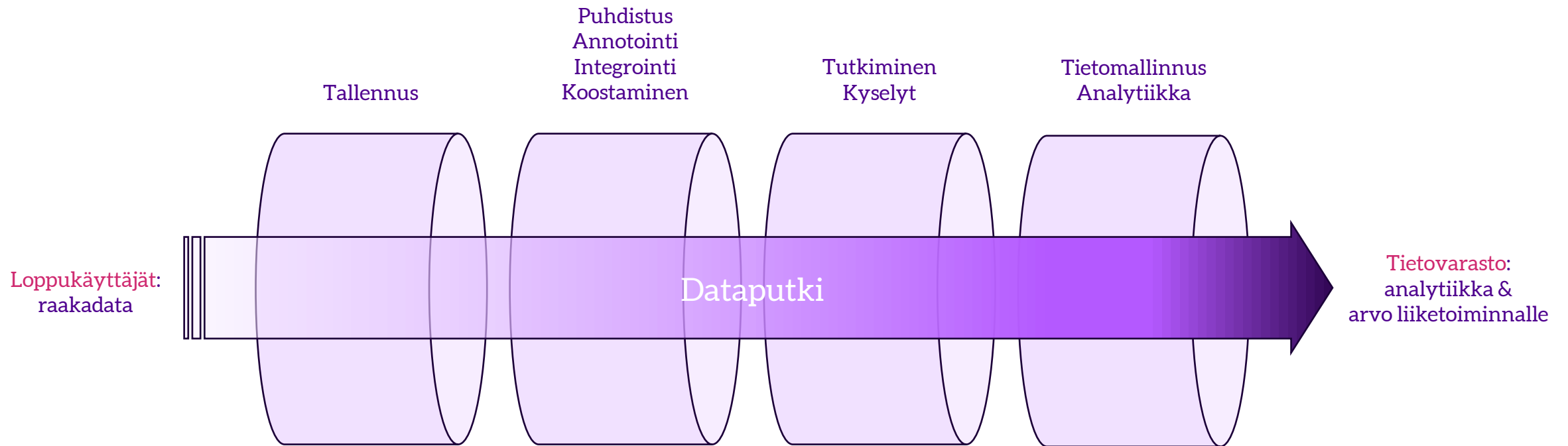


 _yes_but

Datan vaatiman tallennustilan vähentäminen



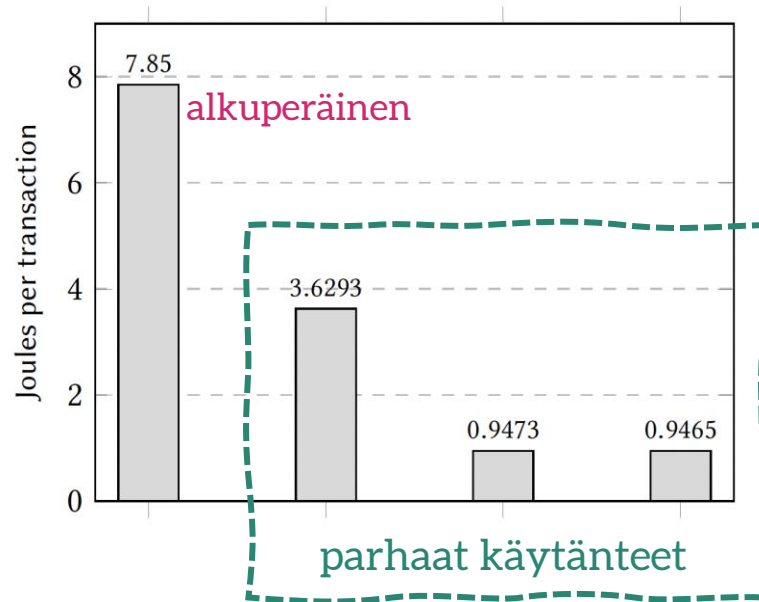
Tallennustilan ja laskennan vähentäminen #1



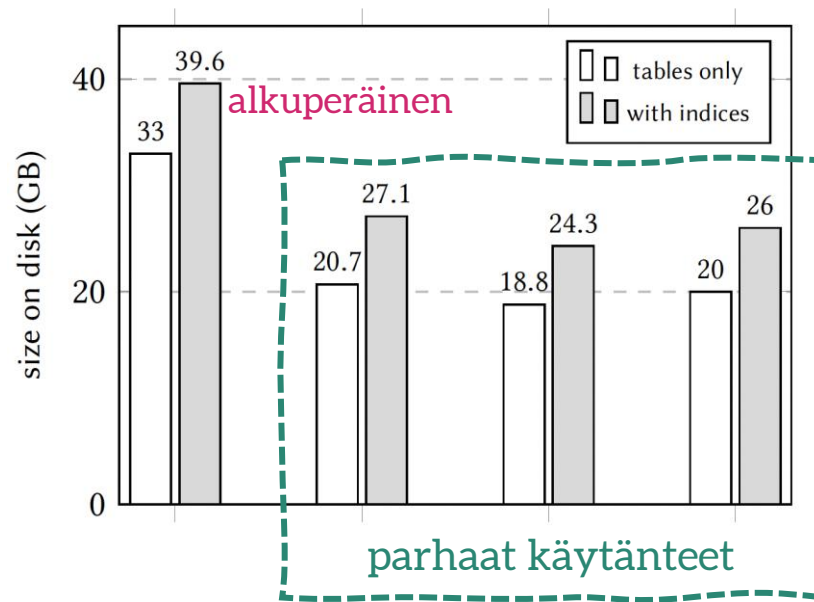
- Datan siirtäminen ja laskenta tarvittaessa, ei “varmuuden vuoksi”.
- Teoreettiset perusteet olemassa, mutta harvoin käytössä.^[1]

Tallennustilan ja laskennan vähentäminen #2

energiankulutus (matalampi = parempi)



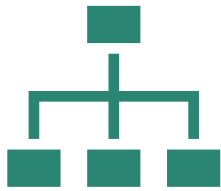
levytila (matalampi = parempi)



- Soveltamalla ~50 vuotta vanhoja parhaita käytänteitä:
 - Vähennettiin energiankulutusta lähes 90 %.
 - Vähennettiin levytilavaatimuksia lähes 40 %.

Kuinka kestävyystavoitteet voidaan saavuttaa

- What would increase your likelihood of adopting *green data engineering* practices? (N = 283) ^[1]



1. Selkeä tuki organisaatiolta
(56 % vastaajista)



4. Ulkoiset säädökset / standardit
(38 %)



2. Paremmat työkalut ja
viitekehykset
(55 %)

DBS-tutkimus



3. Määrälliset metriikat
(43 %)



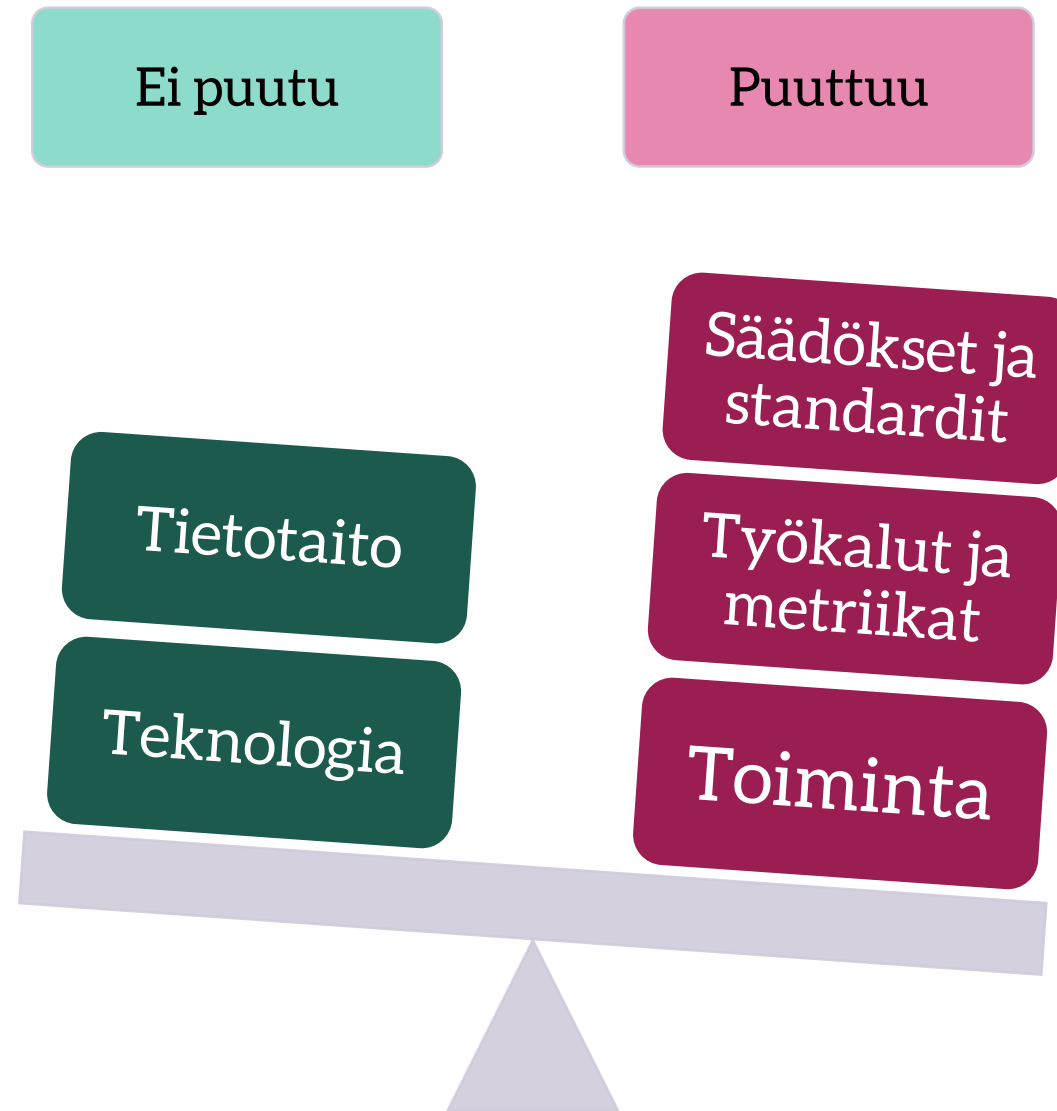
5. Vertaistuki
(30 %)

Yhteenveto

“I know what to do, how to do it,
and why it is important.

There is just no
professional incentive
to do it.”

-Anonymous data engineer ^[1]





Vihreä data: Tallennus ja käsittely

Toni Taipalus
Tampereen yliopisto
toni.taipalus@tuni.fi



Database Systems -tutkimusryhmä
<https://dbs-research.github.io>

