

*Suomen korkeakoulutus- ja tutkimusjärjestelmän  
tulevaisuus – kansainvälisyyttä ja avointa tiedettä  
2.12.2013 – Helsinki*

# **Tutkimuksen tietoaineistojen infrastruktuuri – avoimen tiedon näkökulma**

Esa-Pekka Keskitalo, Kansalliskirjasto

[esa-pekka.keskitalo@helsinki.fi](mailto:esa-pekka.keskitalo@helsinki.fi)



www.oldbaileyonline.org  
197 000 oikeudenkäyntiä  
2 500 kuolemaantuomittua  
Koneluvun kannalta sekamelska  
Osin OCR sekä naputeltu käsin  
170 000 000 sanaa  
100 000 lekseemiä jaettu 1 000 luokkaan  
1000-ulotteinen avaruus –  
onko kyse ystävyydestä, luottamuksesta, vaatetuksesta...  
Tuottanut kymmeniä ellei satoja tutkimuksia.  
Data on avoimesti saatavissa!

The Proceedings of the OLD BAILEY London's Central Criminal Court, 1674 to 1913

Home Search About The Proceedings Historical Background API The Project Contact

Celebrating a Decade of Making History Accessible: The Old Bailey Online, 2003-2013

Home Page  
Search  
About the Proceedings  
Historical Background  
The Project  
Copyright & Citation Guide  
Contact  
Research and Study Guides  
Login / Register

**ON THIS DAY IN... 1795**  
A box containing 39 glass bottles of Spilsbury's antiscorbutic drops was stolen from the 'rumble tumble', the basket at the back of a coach, as it rode out of London. [read more](#)

**The Proceedings of the Old Bailey, 1674-1913**  
A fully searchable edition of the largest body of texts detailing the lives of non-elite people ever published, containing 197,745 criminal trials held at London's central criminal court. If you are new to this site, you may find the [Getting Started](#) and [Guide to Searching](#) videos and tutorials helpful.

To search the Proceedings use the boxes on the right or go to the

**SEARCH**  
the Proceedings  
Keyword(s)  
Reference No.  
Search In  
ALL TEXTS

**ON THIS DAY IN... 1795**  
A box containing 39 glass bottles of Spilsbury's antiscorbutic drops was stolen from the 'rumble tumble', the basket at the back of a coach, as it rode out of London. [read more](#)

# Tutkimusdata on tapetilla

1. Datan määrä
2. Datan tallennusmahdollisuudet
3. Datan analysointimahdollisuudet
4. **Datan uudelleenkäyttö**

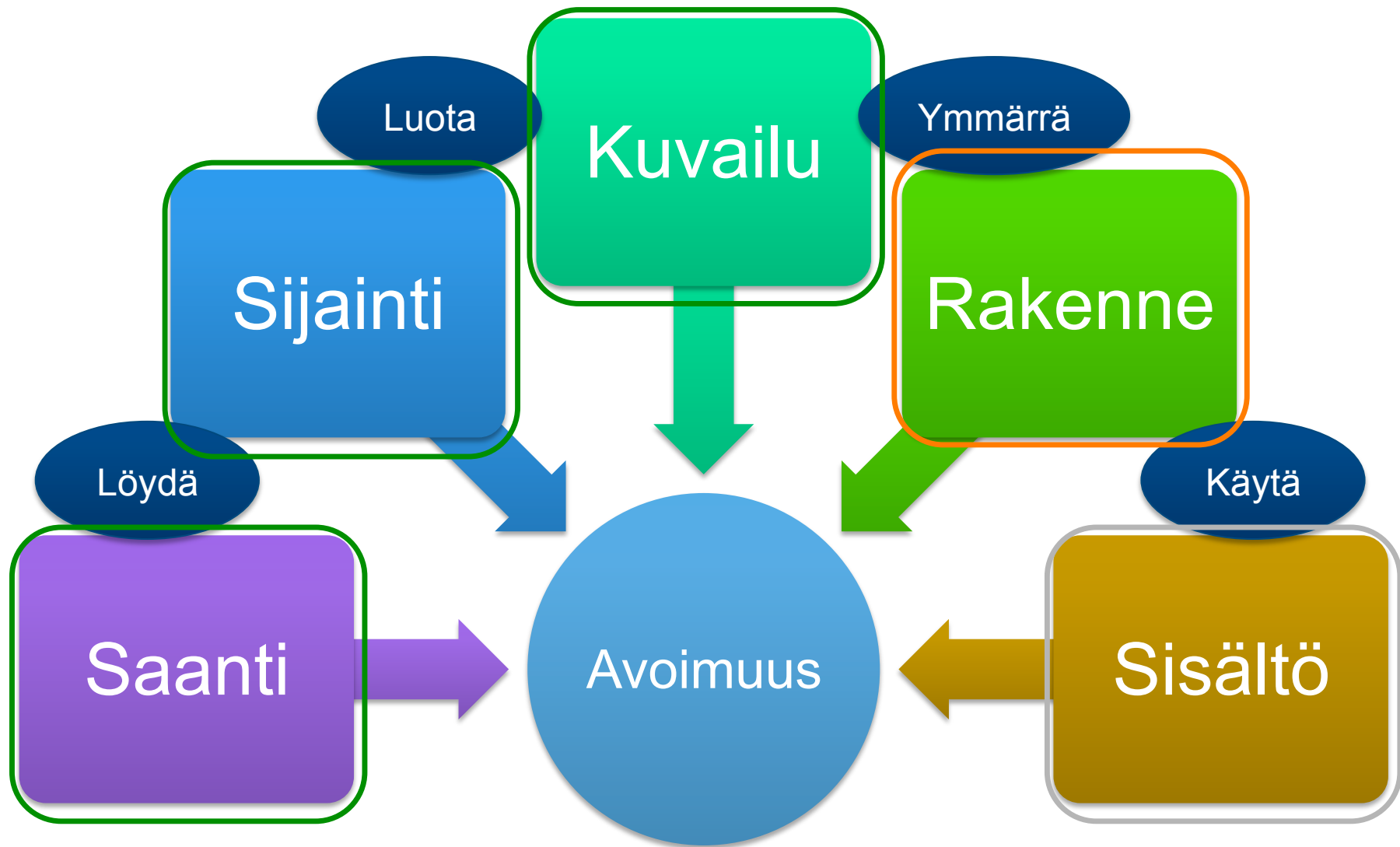
**”(T)here are predictions that by 2020, 80 percent of research projects are expected to begin by mining existing data as opposed to starting afresh.” \***

\* Nimittäin markkinatutkimuksessa: <http://www.wns.com/Resources/Articles/Article-Details/106/The-New-age-Survival-Secret-for-Marketers.aspx>

# Tutkijan työpöydältä julkaisuluetteloon – tutkimusdata?



- Millainen kokonaisuus?
- Millaisessa muodossa?
- Miten selitän tämän datan ulkopuolisille? Miten tämän voi tulkita väärin?
- Onko tässä luottamuksellista tietoa?
- Kuka päättää tämän datan käytöstä?
- Missä se säilytetään?
- Miten kukaan löytää sitä?
- Mitä tämä data on – ja mitä se ei ole?
- Miksi luotan siihen?
- Millä ehdoilla saan käyttää tätä dataa?
- Miten haen lupaa, jos tarvis?
- Miten saan tiedostot auki?
- Miten viitataan tähän dataan?



**“Tiedon jakamisen tehtävän täytämme vain puoliksi,  
jos tieto ei ole laajasti ja helposti  
koko yhteiskunnan käytössä.” - Berliinin julistus**