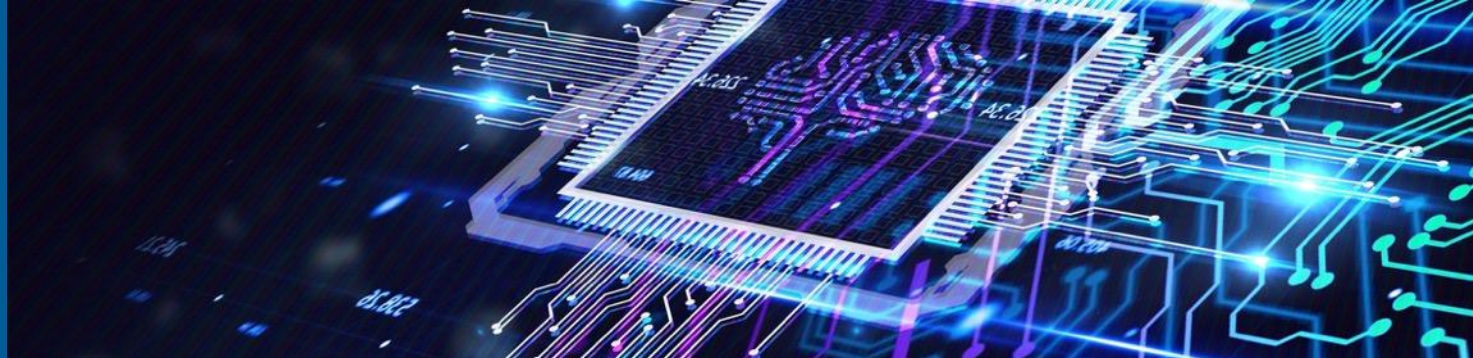




CSC

ICT Solutions for
Brilliant Minds



Hankkeiden aihemallinnus tiedejatutkimus.fi-palvelussa

Lähes 7000 hankekuvausta kertoo, mitä Suomessa tutkitaan

Katja Mankinen / PhD, data scientist at CSC - IT Center for Science
Tutkimustietovarannon ohjausryhmä 3.9.2021



- Mitä tutkimusaiheita tai -teemoja tutkitaan Suomessa tai omassa organisaatiossani?
- Kuinka monta tutkimushanketta rahoitettiin ilmastonmuutoksesta/syövästä/aiheesta X vuosina 2010-2019?
- Mitkä aiheet saavat eniten rahoitusta?
- Rahoittavatko kaikki rahoittajat samankaltaisia aiheita?

→ käytetään ohjaamattomia koneoppimismenetelmiä (aihemallinnusta) ryhmittelemään tekstiaineistoja automaattisesti (usein tieteenalarajat ylittäviin) aiheisiin

Aihemallinnus Tiedejatutkimus.fi -palvelussa

Year of research-based knowledge

2021

Search for information on research in Finland

Search target ▼

For example, publication, field of science, keyword

SEARCH

Search help ▼



Publications
566 686



People
Coming soon



Projects
6 957



Research data
10 362



Infrastructures
133



Other research activities
Coming soon



Organizations
73

- Tällä hetkellä n. 30% hankkeista puuttuu tieteenalatieto
- Aihemallinnus voi parantaa hankkeiden löydettävyyttä ja tarjota uusia näkökulmia suomalaiseen tutkimuskenttään
- Tällä hetkellä testauksessa: tulossa käytettäväksi palveluun pian

Aihemallinnus hankekuvauksilla

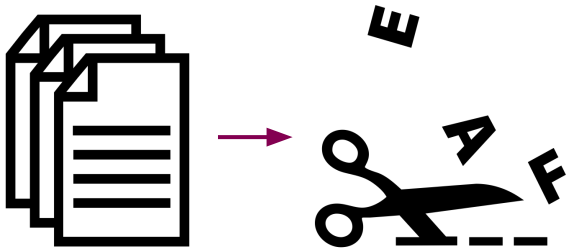


Aineisto

Lähes 7000 hanketta (otsikko, kuvaus, mahdolliset avainsanat ja tieteenala)

Sisältää mm. EU Horizon -hankkeita vuodesta 2014 sekä tietoa muista rahoituspäätöksistä niiltä rahoittajilta, jotka tuottavat tietoa tutkimustietovarantoon

Aihemallinnus hankekuvauksilla



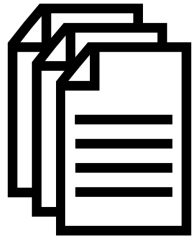
Aineisto

Tekstin käsittely

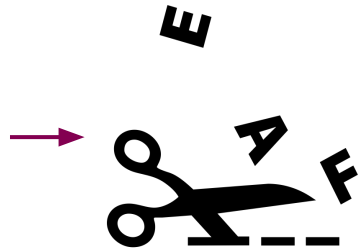
Poistetaan ns. sulkusanat ja liian yleiset termit ("results", "method"...), ja tunnistetaan kuvauksen kieli

Muunnetaan teksti numeeriseen muotoon (sana- ja dokumenttivektoreihin) käsittelyn helpottamiseksi

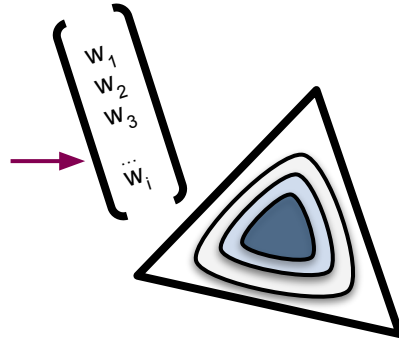
Aihemallinnus hankekuvauksilla



Aineisto



Tekstin käsittely

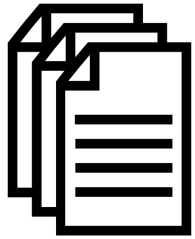


Aihemallinnus

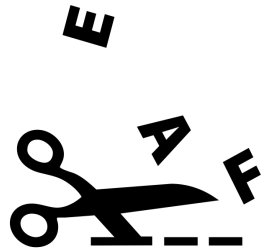
Käytetään ja vertaillaan eri menetelmiä ([contextualized topic models](#), [top2vec](#), [BERTopic](#), [hierarchical stochastic block model](#)) seulomaan mahdollisia aiheita kuvausten sanojen välisiin suhteisiin perustuen

Aihe = joukko sanoja, joilla on jokin semanttinen yhteys

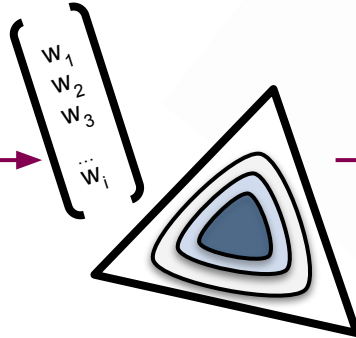
Aihemallinnus hankekuvauksilla



Aineisto



Tekstin käsittely



Aihemallinnus

Valitaan 92 aihetta, ja luokitellaan kukin hanke merkittävimpään aiheeseen



pieniä, suuria, vanhoja, nousevia, monialaisia...
 → aiheet voivat kehittyä, ja uusia aiheita voi ilmetä
 → mallia päivitetään, kun uutta dataa saatavilla

Käyttäjä voi visualisoida, filteröidä ja tutkia hankkeita eri kenttiin perustuen

aineistossa valmiina (ei koneoppimista)

tulossa uutta: aihehallinnuksessa tunnistetut aiheet

Hankkeet - 6 358

Näytetään tulokset 1 - 10 / 6358
10 tulosta / sivu
Mitä hanketietoja palvelu sisältää?

Näytä kuvana

Rajaa hakua

Aloitusvuosi ▼

+ Organisaatio 1

Rahoittaja ▼

+ Rahoitusmuoto 1

+ Tieteenala 1

- Aihe

Avainsanat ▼

Teemat ▼

Tunnistettu aihe 1

Valitse kaikki

business, investments 499

health care 496

sustainable ecosystems, development 475

manufacturing, production 272

HANKKEEN NIMI	RAHOITTAJA	SAAJA	ALOITUSVUOSI
Augmenting Human and Artificial Intelligence to Scale Innovation	Euroopan uniio	Tampereen yliopisto	2022
FAST and energy efficient Learned image and video CompressiON	Euroopan uniio	Tampereen yliopisto	2022
Circulating microRNAs in lynch syndrome	Euroopan uniio	Jyväskylän yliopisto	2022
Mitokondriometabolia hermoston terveyden ja sairauden säätelijänä	Suomen Akatemia	Anu Wartiovaara, Helsingin yliopisto	2022
New, realistic and robust models for cryptocurrency volatility	Euroopan uniio	Tampereen yliopisto	2022
Kestävyystoimijuus terveydenhuollossa	Maj ja Tor Nesslingin Säätiö	Larissa Niemi	2022
Linking ILLegal WILDlife trade and biodiversity loss: new insights from digital media	Euroopan uniio	Helsingin yliopisto	2022
Suomen korkean suojeluarvon metsien (HCV) kartoitus ja tiedon käyttöönotto	Koneen Säätiö	Aalokas Oy	2021
Eläinoikeusraportti 2021: Suomen ensimmäinen vuosiraportti eläinten oikeuksista	Koneen Säätiö	-	2021
AERial RObotic TRAINing for the next generation of European infrastructure and asset maintenance technologies	Euroopan uniio	Tampereen yliopisto	2021

Rajaukset (1):

sustainable ecosystems, development ✕

Tyhjennä rajaukset

	HANKKEEN NIMI ⇅	RAHOITTAJA ⇅	SAAJA ⇅	ALOITUSVUOSI ⇅
	<u>Sopeutuminen sinisen kasvun avaimena</u>	Suomen Akatemia	Mikael Sokero, Demos Research Institute Oy	2021
	<u>Turvemaiden vettäminen typen kuormittamissa ympäristöissä: luonnon monimuotoisuus, ilmasto, veden laatu ja yhteiskunta fokuksessa</u>	Suomen Akatemia	Kristiina Regina, Luonnonvarakeskus	2021
	<u>Metsäluonnon monimuotoisuuden suojele ja hiilen sitominen muuttuvassa ympäristössä (IBC-CARBON)</u>	Suomen Akatemia	Atte Moilanen, Helsingin yliopisto	2021
	<u>Ympäristöliikkeen monimuotoisuus Venäjällä</u>	Koneen Säätiö	Juha Kotilainen, Itä-Suomen yliopisto	2021
	<u>Vihreän kullan kulttuuritien kehittämishanke</u>	Euroopan unioni	Kaakkois-Suomen ammattikorkeakoulu; Lusto - Suomen Metsämuseo	2021
	<u>Sopeutuminen sinisen kasvun avaimena</u>	Suomen Akatemia	Jukka Similä, Lapin yliopisto	2021
	<u>Satakuntalaisen ruokaketjun ilmastokestävyysohjelma</u>	Euroopan unioni	Satafood Kehittämisyhdistys ry	2021

Näytä kuvana

Rajaa hakua

Aloitusvuosi

Organisaatio

Rahoittaja

Rahoitusmuoto

Tieteenala

Aihe

Avainsanat

Teemat

Tunnistettu aihe

Hae

Valitse kaikki

business, investments 499

health care 496

sustainable ecosystems, development 475

manufacturing, production 272

Näytetään tulokset 1 - 10 / 6358

10

tulosta / sivu

Mitä hanketietoja palvelu sisältää?

HANKKEEN NIMI	RAHOITTAJA	SAAJA	ALOITUSVUOSI
 Augmenting Human and Artificial Intelligence to Scale Innovation	Euroopan unioni	Tampereen yliopisto	2022
 FAst and energy efficient Learned image and video CompresION	Euroopan unioni	Tampereen yliopisto	2022
 Circulating microRNAs in lynch syndrome	Euroopan unioni	Jyväskylän yliopisto	2022
 Mitokondriometabolia hermoston terveyden ja sairauden säätelijänä	Suomen Akatemia	Anu Wartiovaara, Helsingin yliopisto	2022
 New realistic and robust models for cryptocurrency volatility	Euroopan unioni	Tampereen yliopisto	2022
 Kestävyystoimijuus terveydenhuollossa	Maj ja Tor Nesslingin Säätiö	Larissa Niemi	2022
 Linking ILLEGAL WILDLife trade and biodiversity loss: new insights from digital media	Euroopan unioni	Helsingin yliopisto	2022
 Suomen korkean suojeluarvon metsien (HCV) kartoitus ja tiedon käyttöönotto	Koneen Säätiö	Aallokas Oy	2021
 Eläinoikeusraportti 2021: Suomen ensimmäinen vuosiraportti eläinten oikeuksista	Koneen Säätiö	-	2021
 AERial RObotic TRAINing for the next generation of European infrastructure and asset maintenance technologies	Euroopan unioni	Tampereen yliopisto	2021



Hankkeiden jakautuminen tunnistetun aiheen mukaan Lisätietoja

Sulje

Valitse teema
Tunnistettu aihe

Pylväskuvaaja Piirakuvaaja

Tunnistettu aihe

Hae

Valitse kaikki

business, investments 499

health care 496

sustainable ecosystems, development 475

manufacturing, production 272

energy, electricity, power 226

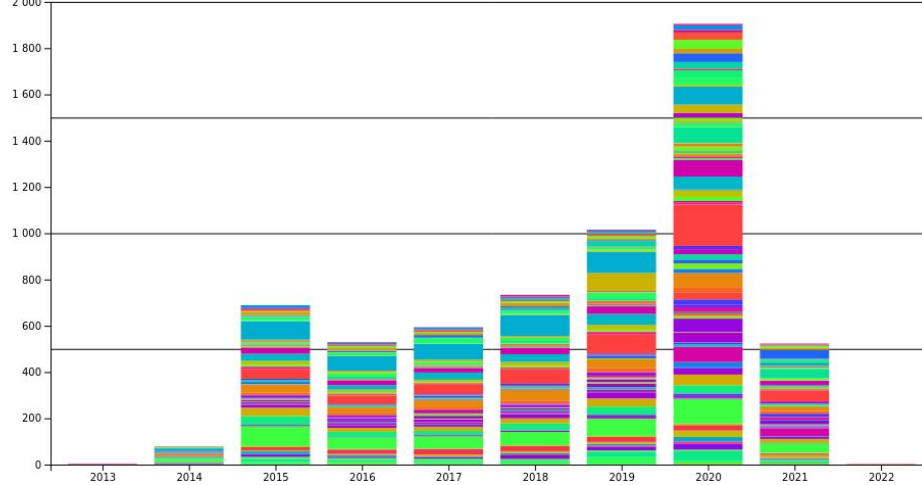
education 224

software development 195

Näytä enemmän Näytä vähemmän

Tyhjennä rajaukset

Ei hakusanaa



- 3d printing, printing
- SMEs, management, growth, entreprene...
- aerosol, atmospheric
- agriculture, farming, dairy
- air quality
- animals
- arctic region
- art, culture
- artificial intelligence, machine learning
- bacteria, antibiotic infections
- batteries, lithium batteries, battery
- biodiversity, species, ecosystems, evolut...
- bioeconomy, biorefinery, biomass, biofue...
- biomass, lignin, hemicellulose, cellulose
- brain diseases, Alzheimer, dementia
- business, investments
- cancer
- carbon, methane, carbon footprint

Hankkeiden määrä

% hankkeiden määrästä

Lataa kuvana (tulossa)

tulossa: kuinka rahoitus jakautuu eri aiheiden välille

Valitse teema
Tunnistettu aihe

Pylväskuvaaja

Piire

Rajaukset (1):

cancer ✕ [Tyhjennä rajaukset](#)

Rahoittaja 1

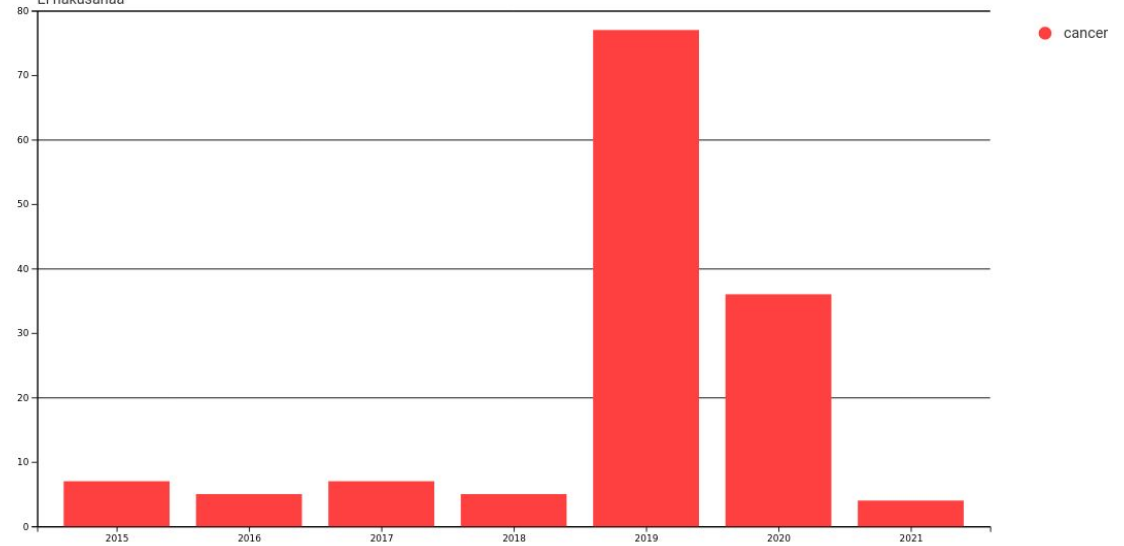
Hae

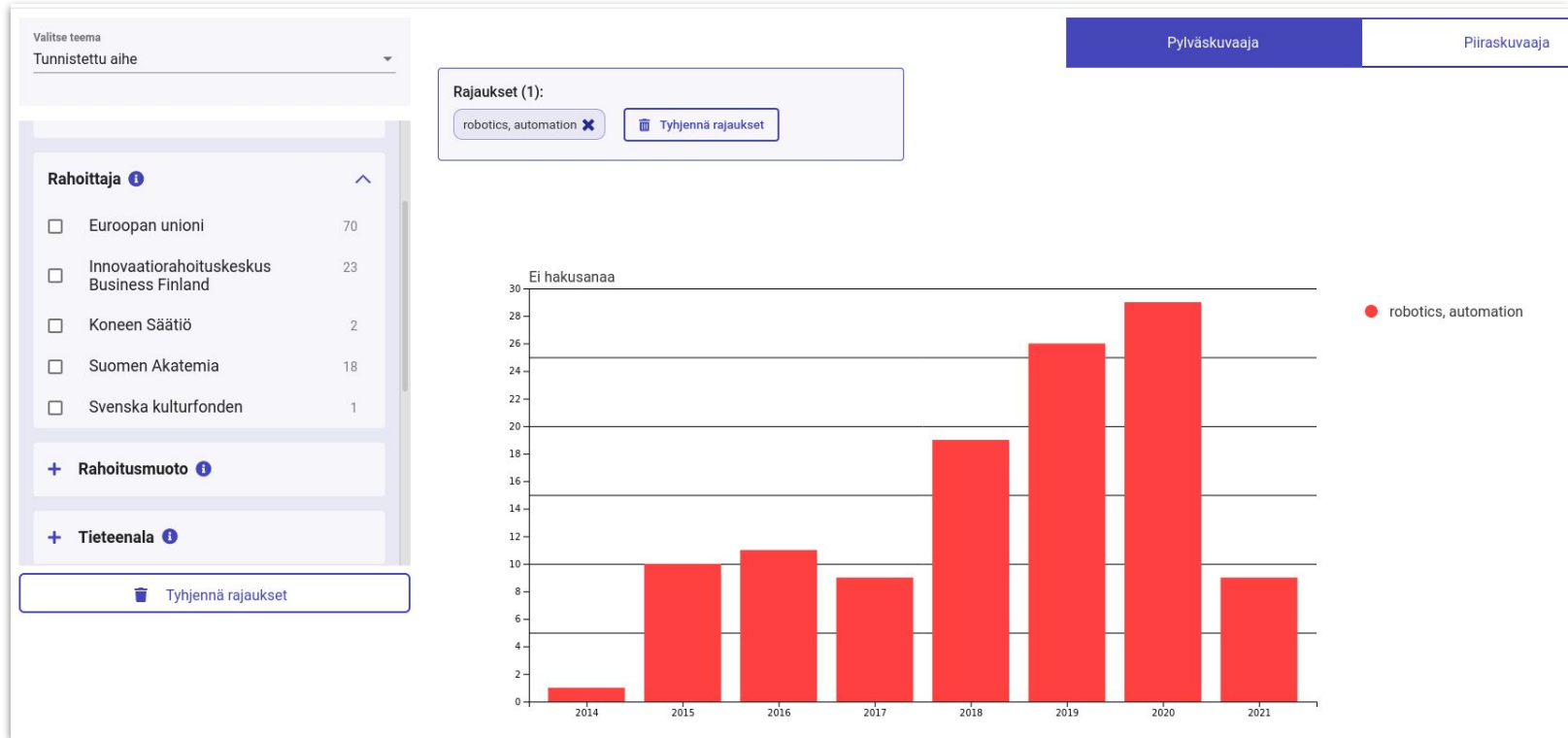
- Euroopan unioni 37
- Innovaatorahoituskeskus Business Finland 8
- KAUTE-säätiö 1
- Suomen Akatemia 26
- Suomen Lääketieteen Säätiö 1
- Suomen Syöpäyhdistys ry 5
- Svenska kulturfonden 6
- Syöpäsäätiö 57

[Näytä enemmän](#) [Näytä vähemmän](#)

[Tyhjennä rajaukset](#)

Ei hakusanaa





Rajoitukset ja huomioitavaa

- Kaikkia aiheeseen liittyviä hankkeita ei ole luokiteltu aiheeseen
- Kaikki aiheeseen luokitellut hankkeet eivät välttämättä tiukasti liity kyseiseen aiheeseen
- Jopa merkittäviä aiheita voi jäädä tunnistamatta!
- Rajoitteita sekä aineistossa että menetelmissä
 - pieni mutta monikielinen aineisto
 - hyvin spesifit tutkimusaiheet
 - osassa kuvauksista pelkkä otsikko ja/tai hyvin lyhyt kuvaus
 - aihemallinnus perustuu ohjaamattomaan koneoppimiseen, ja pienet parametrimuutokset voivat tuottaa erilaisia tuloksia
 - tekstin esikäsittely



Aihemallinnus sopii tutkimusteemojen kartoittamiseen suuressa mittakaavassa, mutta EI yksittäisten aihealueiden tarkkaan kokojen määrittämiseen

Aihemallinnus tarjoaa uusia näkökulmia tutkimukseen

- Koneoppimista ja aihemallinnusta voidaan käyttää datalähtöiseen aiheiden tunnistamiseen
- Löydetyt aiheet voivat ylittää tieteenalarajoja, ja tarjota uusia näkökulmia niiden lisäksi
- Tulossa tiedejatutkimus.fi-palveluun:
 - Aihemallinnustulokset ja koneoppimisen avulla generoidut avainsanat parantavat hakuominaisuuksia
 - Lisää visualisointeja (mm. rahoituksen määrä hankkeittain)