

# Sanastotyö tiedonvaihdon tukena

---

Elati-seminaari: Tietojen  
standardointi ja sanastotyö  
24.9.2008

Katri Seppälä  
Sanastokeskus TSK

# Sanastokeskus TSK ry

---

- Ø vuonna 1974 perustettu yhdistys
    - Ø aik. Tekniikan Sanastokeskus ry
  - Ø tarkoitus: saada aikaan ensisijaisesti suomen- ja ruotsinkielisiä sanastoja sekä kehittää sanastotyön metodiikkaa ja pitää yllä siihen perustuvaa osaamista
  - Ø tavoitteita: viestinnän selkeyttäminen, organisaatioiden toiminnan tehostaminen, suomen kielen käyttöalan kaventumisen ehkäiseminen
  - Ø noin 80 jäsentä: yhteisö- ja henkilöjäseniä
  - Ø työntekijät: kielten ja muiden erikoisalojen ammattilaisia
- <http://www.tsk.fi>, [etunimi.sukunimi@tsk.fi](mailto:etunimi.sukunimi@tsk.fi)

# Täyden palvelun sanastokeskus

---

- Ø erikoisalojen sanastoaineistojen tuottaminen
  - Ø oma XML-pohjainen sanastonhallintajärjestelmä
- Ø sanastotyöhön liittyvä koulutus ja konsultointi
- Ø tiedotus- ja julkaisutoiminta
  - Ø Heidi Suonuuti: Sanastotyön opas
- Ø julkisen termipankin ylläpito ([www.tsk.fi/tepa/](http://www.tsk.fi/tepa/))
- Ø erikoisalojen sanastoihin erikoistunut kirjasto
- Ø termineuvonta ja erikoiskielisten tekstien tarkistus
- Ø kansainvälinen yhteistyö
  - Ø Nordterm
    - Ø esim. NTRF-tietuemalli:  
<http://www.nordterm.net/info/Ntrf.pdf>
  - Ø EAFT (European Association for Terminology)

# Täyden palvelun sanastokeskus

---

- ∅ sanastotyön menetelmien tutkimus ja kehitys
  - ∅ ISO/TC 37 Terminology and other language and content resources
    - ∅ ISO 639-1:2002 Codes for the representation of names of languages -- Part 1: Alpha-2 code
    - ∅ ISO 704 Terminology work – Principles and methods
    - ∅ ISO 860:2007 Terminology work -- Harmonization of concepts and terms
    - ∅ ISO 10241:1992 International terminology standards -- Preparation and layout
    - ∅ ISO 12620:1999 Computer applications in terminology -- Data categories
    - ∅ ISO 15188:2001 Project management guidelines for terminology standardization

# Miksi sanastotyötä tehdään?

---

- Ø tavoitteena omakielisten termisuositusten ja vastineiden antaminen sekä käsitteiden määrittely
  - Ø viestinnän helpottamiseksi ja laadun parantamiseksi
    - Ø väärinkäsitysten välttämiseksi
  - Ø tietojärjestelmien yhteiskäytön helpottamiseksi
  - Ø käännöstyön helpottamiseksi
  - Ø toiminnan tehokkuuden parantamiseksi
  - Ø kustannusten säästämiseksi

# Projektimuotoinen sanastotyö

---

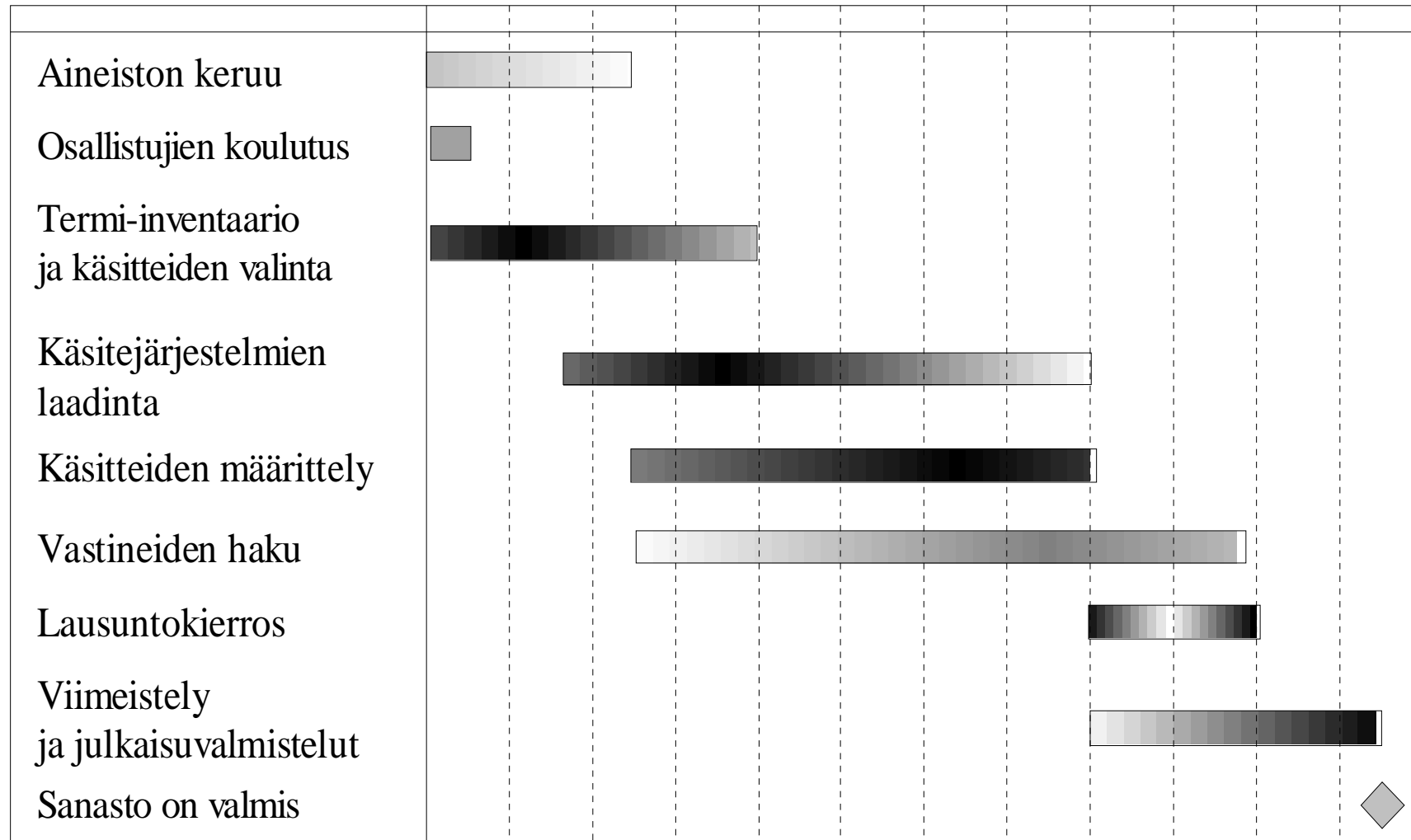
- Ø työryhmässä tehtävät ja vastuu jakautuvat seuraavasti
  - Ø kohdealan asiantuntijat: alakohtainen tietämys – vastuu asiasisällöstä
  - Ø terminologi(t): sanastotyön menetelmien tuntemus – vastuu terminologisesta laadusta
- Ø tyypillisesti sanastotyöryhmään kuuluu 1-2 terminologia ja 6-10 kohdealan asiantuntijaa
  - Ø lisäksi projektilla voi olla koordinaattori, joka huolehtii mm. kokousten järjestämisestä
- Ø kun kohdealan asiantuntijat ja terminologit tekevät yhteistyötä, molemmat voivat keskittyä omaan osaamisalueeseensa ja säästyy aikaa

# Työryhmän toimintatapa sanastoprojektissa

---

- Ø työnaikaisen aineiston hallinta
  - Ø tietojen tallennus- ja ylläpitovastuu sovitulla henkilöllä
  - Ø sovitut merkintätavat ja käytännöt
  - Ø sanastoluonnoksia tarpeen mukaan
- Ø yhteydenpito
  - Ø kokousajankohdat sovitaan ajoissa etukäteen
  - Ø kokousten välillä yhteydenpitoa sähköpostitse
- Ø kokoukset
  - Ø kokouskohtaiset tavoitteet
  - Ø tarpeeksi aikaa, tehokas ajankäyttö

# Sanastoprojektin vaiheet



# Sanastoprojektin suunnittelu

---

- Ø kohderyhmän valinta ja huomioon ottaminen
- Ø alustava rajausta (aiheen jako osa-alueisiin)
- Ø laajuus (käsitteiden määrä)
- Ø käsitteiden kuvaustapa (määritelmät, huomautukset, käsitejärjestelmäkaaviot)
- Ø määritelmä- ja vastinekielet
- Ø organisointitapa
  - Ø työnjako terminologin ja asiantuntijoiden kesken
    - Ø riittävästi aihealueiden ja vastinekielten tuntemusta
    - Ø edustajia tärkeistä intressiryhmistä
  - Ø osallistujien perehdyttäminen työmenetelmään

# Sanastoprojektin suunnittelu

---

- Ø rahoitus
- Ø oikeudet sanastoaineistoon
- Ø sanaston julkaisutavat

Oikein valitut työmenetelmät säästävät jopa 30—50 % sanastotyön kokonaiskustannuksista.

# Inventaariovaihe

---

- ∅ valittua aihealuetta käsittelevän lähdeaineiston kartoitus ja hankinta
- ∅ tallennustavan suunnittelu
- ∅ esitermistön ja käsitetietojen tallennus
- ∅ sanaston alustava jäsentely
- ∅ alustavien käsitejärjestelmien hahmottelu
- ∅ sanaston rajausta eli käsitteiden valinta

# Käsiteanalyysi- ja määrittelyvaihe

---

- Ø käsitteiden sisällön selvittäminen (olennaispiirteiden analysointi)
- Ø käsitejärjestelmien kuvaaminen (esim. kaavioina)
- Ø määritelmien ja niitä täydentävien huomautusten laatiminen
- Ø vastineiden haku (yleensä ainakin osittain samanaikaisesti)
- Ø termien valinta ja käyttösuositusten tai -rajoitusten asettaminen
- Ø Vie jopa 50 % koko projektiin varattavasta työajasta!

# Lausuntokierros

---

## Ø Miksi?

- Ø mahdollisten määrittelyvirheiden, puutteiden yms. havaitsemiseksi
- Ø vastineiden tarkistamiseksi tai keräämiseksi
- Ø sanaston yleisen hyväksyttävyyden varmistamiseksi (erityisesti termivalintojen osalta)
- Ø tiedottamiseksi (ennakkomarkkinointi)

## Ø Keneltä ja miten kysytään?

- Ø jakelu esim. 10—50 asiantuntijataholle ja erityisesti sanaston kohderyhmälle sekä kielenhuoltajille
- Ø täsmällinen vastausaika (esim. 1 kk) ja ohjeet

# Viimeistelyvaihe

---

- Ø terminologinen tarkastus
  - Ø mahdollisten määrittelyvirheiden etsintä
  - Ø merkintätapojen ja kielellisen johdonmukaisuuden tarkastus
- Ø lausuntojen vaatimat tarkistukset
  - Ø sisältömuutokset, täsmennykset, täydennykset ym. korjaukset
  - Ø sanamuotojen viimeistely
  - Ø termivalintojen tarkistus
- Ø esipuheen, käyttöohjeiden jne. laadinta
- Ø taitto ja korjausluku

# Käsitetietojen käsittely ja esittäminen rakenteisessa muodossa

---

- Ø rakenne suunniteltava etukäteen
  - Ø käsitekohtaiset tiedot esitetään termitietueena
  - Ø tietuerakenteeseen sisällytettävät
    - Ø tietotyypit
    - Ø tietokenttien määrä
    - Ø tietokenttien järjestys
  - valitaan loppukäyttäjän tarpeet ja käytettävät ohjelmistot huomioon ottaen
- Ø merkintöjen yksiselitteisyys ja johdonmukaisuus on tärkeää
  - Ø tukena XML-editori sekä DTD- (Document type definition) ja CSS-määrittelykset (Cascading Style Sheet)

# Käsitteiden käsittely ja esittäminen rakenteisessa muodossa

---

- Ø rakennetta suunniteltaessa on tärkeää tuntea hyvin etukäteen
  - Ø tallennettavan tiedon luonne ja looginen rakenne
  - Ø tietojen keskinäiset riippuvuussuhteet

# Esimerkkejä tietotyypeistä

---

- ∅ lähtökielen termit (käsitteiden nimitykset)
  - ∅ päätermi (vrt. esitermi)
  - ∅ synonyymit, lähisynonyymit
  - ∅ hylättävät termit
- ∅ lähtökielen termejä koskeva aputieto
  - ∅ homonymia ja polysemia
  - ∅ kielioppitieto
    - ∅ esim. sanaluokka, suku, luku, taivutus
  - ∅ maantieteelliset rajoitukset
  - ∅ käyttöalaa koskevat tiedot
- ∅ monikielisessä aineistossa vastaavat tiedot vastinekielistä

# Esimerkkejä tietotyypeistä

---

- ∅ käsitteiden sisältöä koskeva tieto
  - ∅ systemaattiset määritelmät ja niitä täydentävät huomautukset
- ∅ termitietueiden luokitus
  - ∅ sisältöä (aihetta) tai laatua kuvaava luokitus
- ∅ viittaukset
  - ∅ toisiin käsitteisiin
  - ∅ lähdeviittaukset
  - ∅ kuviin tai käsitekaavioihin
  - ∅ hakemistoviittaukset
- ∅ hallinnolliset tiedot kuten tietueen luonti- ja päivitysajankohta ja tallentaja

# Elintarvikelain mukaisten rekisterien sanasto – Valvontakohteet (KUTI 1)

- Ø MIKSI? kuntien elintarvikevalvonnan tiedonkeruun tueksi
- Ø KENELLE? erityisesti kuntien valvontaviranomaisille ja muille elintarvikevalvonnassa työskenteleville
- Ø MITÄ? suomenkieliset termisuositukset ja määritelmät sekä ruotsin- ja englanninkieliset vastineet
- Ø MILLOIN? vuoden määräajassa (v. 2007)
- Ø MITEN? järjestelmällisen sanastotyön periaatteiden mukaisesti, sähköiseksi julkaisuksi, ks.  
[http://www.evira.fi/attachments/elintarvikkeet/valvonta\\_ja\\_yrittajat/valvontasuunnitelmat/kuti/sanasto\\_kuti1.pdf](http://www.evira.fi/attachments/elintarvikkeet/valvonta_ja_yrittajat/valvontasuunnitelmat/kuti/sanasto_kuti1.pdf)
- Ø KUKA? TSK yhteistyössä Eviran ja Eviran yhteistyökumppaneiden kanssa
- Ø KENEN RAHOITUKSELLA? Eviran rahoituksella

# Termitietuemalli (Morphon/XML)

Morphon XML-Editor 3.1.4 - Eviratyö.xml

File Edit View Tools Layout Help AddElement TSK

AddElement: RECORD SY SV SOURF NOTE REMK LINK DEF LANG CRDAT

TSK: split up Rename hrefs Import XPATH REF Set Diag Attributes Draw Diagrams Export Translitter

Eviratyö.xml

6 Valvonta ja rekisterit

```
DIAG name='Valvonta ja rekisterit'  
LINK valvonta  
LINK tuotevalvonta  
LINK olosuhdevalvonta  
LINK valvontakohde  
LINK valvontaviranomainen  
LINK valvontakohteen tunnistetieto  
LINK laitoksen hyväksymisnumero  
LINK maastilatunnus  
LINK rekisteri  
LINK tuotantotapa
```

NUMB= ID63, NR0= 59

```
fi TE valvonta [pt] {SY: perusvalvonta [CAP]}  
DEF toiminta, jolla selvitetään ja todetaan, täyttääkö valvontakohde tai tuote sille asetetut vaatimukset [CAP]  
NOTE Valvonta voidaan jakaa valvontaviranomaisen suorittamaan viralliseen valvontaan ja toimijan suorittamaan omavalvontaan. [tr5]  
NOTE Valvonta voidaan jakaa tuotevalvontaan ja olosuhdevalvontaan. []  
STE virallinen valvonta [tr5]  
STE omavalvonta [tr5]  
sv TE kontroll [hw]  
SY övervakning [CAP]  
SY tillsyn [CAP]  
en TE control [CAP]
```

NUMB= ID160, NR0= 60

```
fi TE tuotevalvonta [pt]  
DEF tuotteen vaatimusten mukaisuuden valvonta [tr5]  
sv TE produktkontroll [Eurlex, sä, hw]  
en TE product control [Eurlex, sä]
```

NUMB= ID122, NR0= 61

```
fi TE olosuhdevalvonta [pt]  
DEF toiminnan ja toimintaympäristön vaatimusten mukaisuuden valvonta [tr5, sä]  
sv TE kontroll av förhållanden [hw]  
en TE control of production conditions [sä]
```

NUMB= ID92, NR0= 62, SUBJ= <valvontakohderekisteri>

```
fi TE valvontakohde [pt]  
DEF toiminnan, toimijan tai toimipaikan yhdistelmä, jota valvontaviranomainen tiettyinä ajanjaksona valvoo [CAP, sä]  
NOTE Valvontakohde voi olla esimerkiksi myymälä tai ravintola. [tr5]  
NOTE Yleisemmin valvontakohteella voidaan tarkoittaa myös elintarvikevalvonnan kohteena olevaa tuotetta. [tr5]
```

Set	Name	Value	Type
<input type="checkbox"/>	typr		CDATA
<input checked="" type="checkbox"/>	href	#ID146	CDATA (#RI

Apply Revert

Mode  
 In  Before  After  Replace  Inline

Positions

Namespace

Elements  
HOGR  
[Comment]  
[Entity]  
[Processing Instruction]  
[toggle TEXT/CDATA]

Show valid results only Insert

Loaded

/VOCABULARY/DIAG[8]/LINK[8]/text()

Käynnistä Saapuneet : katri.sep... Reittipias - Mozilla Fir... Microsoft PowerPoint ... Sanasto - Microsoft I... Morphon XML-Editor ... FI 10:55

# Termitietuemalli (OpenOffice > PDF)

---

## **valvonta**

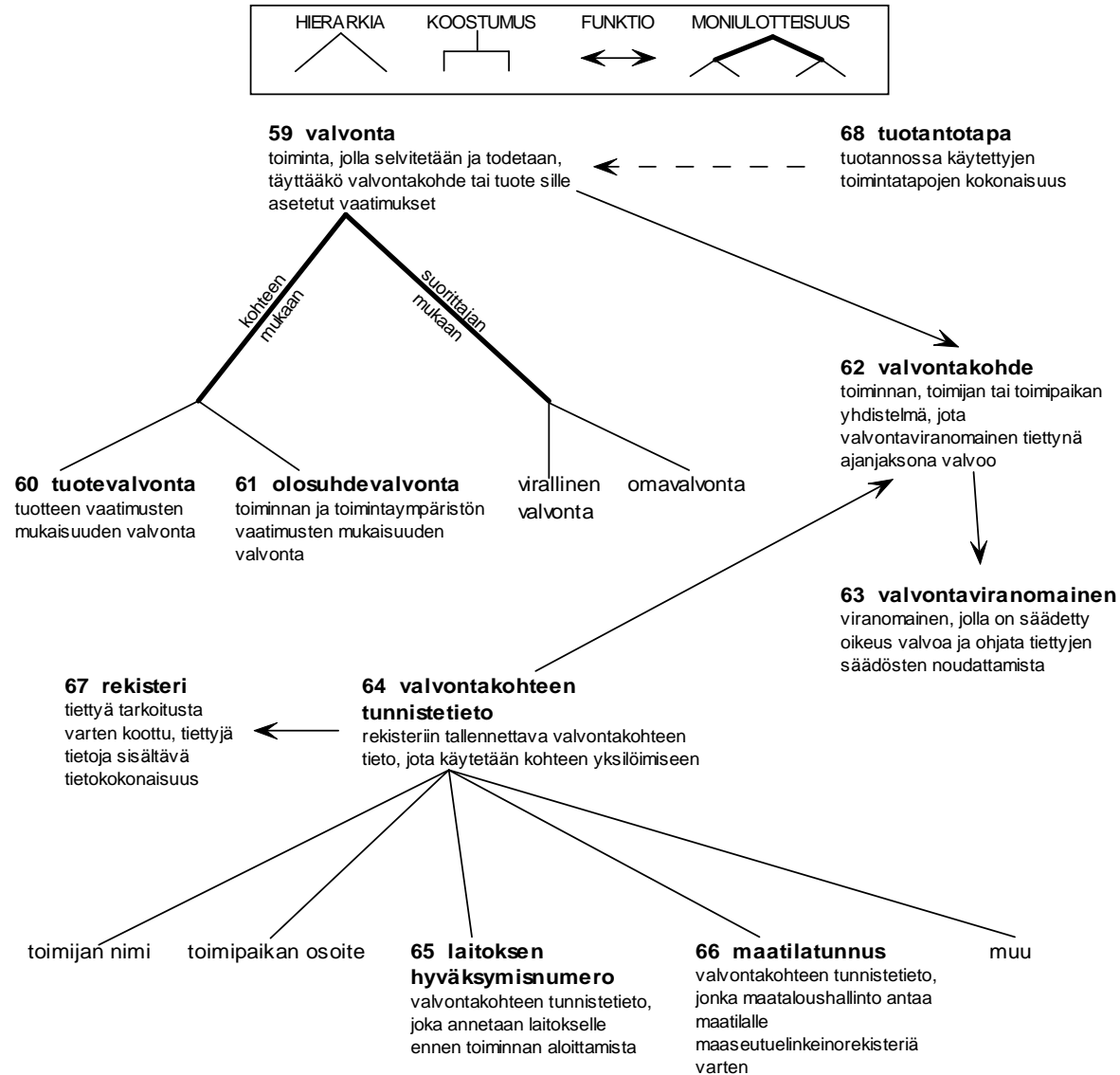
sv kontroll; övervakning; tillsyn

en control

*toiminta*, jolla selvitetään ja todetaan, täyttääkö *valvontakohde* tai tuote sille asetetut vaatimukset

Valvonta voidaan jakaa *valvontaviranomaisen* suorittamaan viralliseen valvontaan ja *toimijan* suorittamaan omavalvontaan.

# Käsitejärjestelmäkaavio



# Julkaisu TEPA-termipankissa

The screenshot shows a Mozilla Firefox browser window displaying the TEPA - Sanastokeskus TSK:n termipankki website. The address bar shows the URL: <http://www.tsk.fi/tepa/netmot.exe?UI=figr&height=159>. The search results for the term 'tuotevalvonta' are shown in a two-column layout.

**tuotevalvonta**

**tuotevalvonta** ⓘ

Elintarvikelain mukaisten rekisterien sanasto - Valvontakoht

fi **tuotevalvonta**

sv produktkontroll

en product control

määntelmä

tuotteen vaatimusten mukaisuuden valvonta

Käsittekaavio: Valvonta ja rekisterit.

Lähde: Elintarvikelain mukaisten rekisterien sanasto - Valvontakohteet (KUTI 1, 2008)

Etusivu Sanastokeskus TSK ry

Valmis

Käynnistä Microsoft PowerPoint ... TEPA - Sanastokesku... FI 11:32

# Yhteenveto

---

- Ø rakenteinen tiedon esitystapa helpottaa tiedonvaihtoa, uudelleenkäyttöä ja ylläpitoa
  - Ø sanastotyössä määritellyt käsitteet ja sovitut termit tukevat esim. tietojärjestelmien rakentamista ja varmistavat näin tiedonsiirtoa
- Ø standardoidut menetelmät ja keskeisten toimijoiden välinen yhteistyö antavat hyvät edellytykset korkealle laadulle
- Ø kun termeistä ja määritelmistä on sovittu yhdessä, niiden vakiintumismahdollisuudet ovat hyvät ja tiedonvälitys tehostuu
  - Ø useampi taho levittää samansisältöistä tietoa
  - Ø tiedonhaku on helpompaa
- Ø kertaalleen systemaattisesti tehty sanastotyö säästää kaikkien kielenkäyttäjien aikaa ja rahaa