

Seminar privind diseminarea datelor statistice și lansarea unei noi soluții web pentru BNS Moldova
Chișinău, 27 mai 2010
Markku Huttunen (markku.huttunen@stat.fi)

Publicarea datelor statistice pe mai multe canale (publicații electronice și baze de date) – Experiența Finlandei

1. Conceptul publicării pe mai multe canale

1.1. Trecerea de la sistemul statistic pe hârtie la sistemul statistic electronic

Statistica oficială reprezintă un sistem de date statistice sociale importante, de înaltă calitate, produse în mod regulat și suficient de frecvent și care sunt reprezentative pentru întreaga țară.

Statistica socială generală reprezintă produsul principal al Biroului de Statistică al Finlandei (Statistica Finlanda). Actualmente Biroul de Statistică al Finlandei produce în jur de 200 de seturi de date statistice pe 26 de subiecte. Unele date sunt colectate lunar sau trimestrial, altele – anual sau mai rar. Noile date statistice sunt publicate pe larg pe Internet. Numărul publicărilor anuale de noi date este de aproximativ 650.

Timp de foarte mulți ani, informațiile statistice erau raportate, publicate și diseminate sub formă de publicații tipărite. Cele mai vechi serii de publicații statistice din Finlanda există de peste 250 de ani. Evoluția dezvoltării Internetului a pus în încurcatură modul tradițional de raportare, publicare și diseminare a informațiilor statistice. Forma tradițională de publicare în care tabelele și informația necesară pentru interpretarea acestora sunt compilate într-un volum, o serie de publicații statistice oficiale a început să dispară. Instituțiile naționale au început să adopte diseminarea electronică a datelor, atât a tabelor (bazelor de date), cât și a analizelor textuale (a publicațiilor) necesare pentru interpretarea datelor.

Ce caracteristici trebuie să posede diseminarea electronică a informației statistice în sistemul statistic în întregime, atunci când are loc trecerea de la diseminarea în scris la diseminarea electronică? Informația statistică trebuie să fie disponibilă pe larg și ușor de găsit. Datele publicate trebuie să fie însoțite de metadatele necesare pentru interpretarea lor, la care trebuie să existe acces direct.

1.2. Structura Comună a Informațiilor Statistice (CoSSI)¹

Pentru a implementa un serviciu Internet eficient, datele statistice propriu-zise (tabele, publicații) și metadatele necesare pentru interpretarea lor, trebuie plasate în format electronic. Însă doar acest lucru nu este suficient pentru a avea un serviciu eficient. Conversiunea automată a informațiilor pentru diferite canale de diseminare, arhivarea acestora, mai multe versiuni de limbă și anexarea unui volum suficient de metadate la acestea sunt probleme ce trebuie soluționate pentru a crea un serviciu adecvat.

Pentru a soluționa aceste probleme a fost elaborată o Structură comună a informației statistice (CoSSI). Punctul de plecare la elaborarea CoSSI a fost analiza

¹ Heikki Rouhuvirta, Harri Lehtinen. Structura Comună a Informațiilor Statistice (CoSSI); <http://www.stat.fi/rossi>

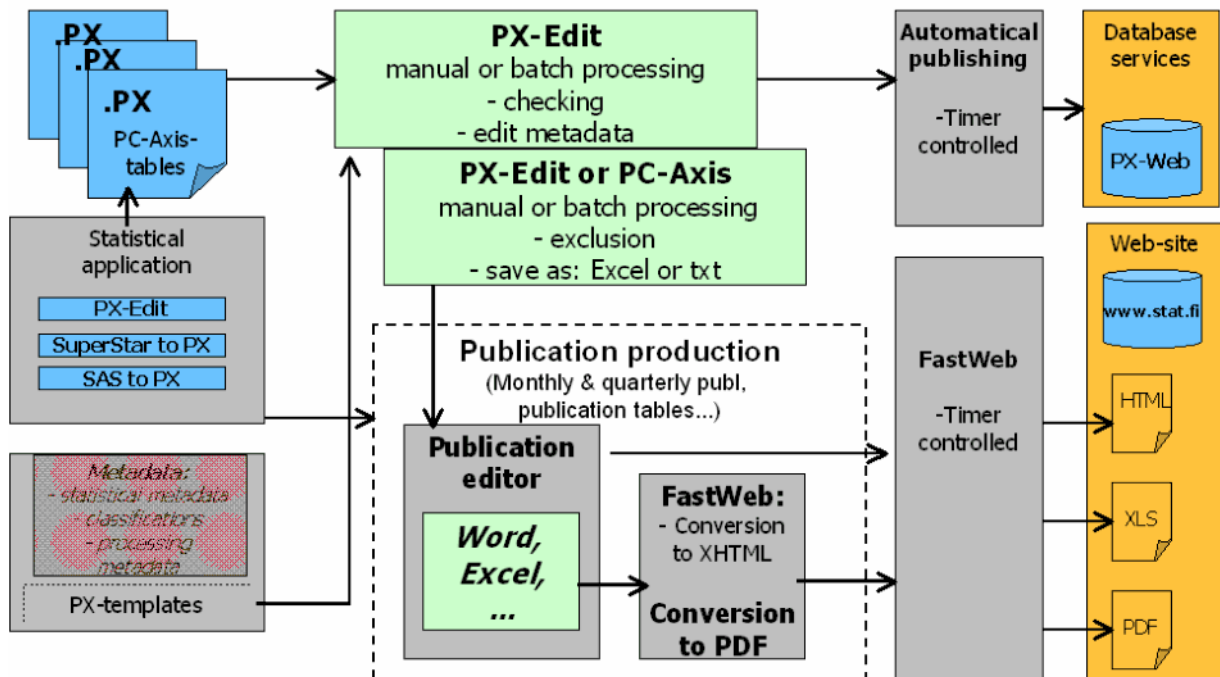
infologic a informaiei statistice. Concluzia analizei a fost că, deși în practică definiția informaiei statistice variază în funcție de situație și de domeniul de aplicare, în realitate informația statistică are o structură universală acceptabilă și simplificată. CoSSI descrie structura comună, care nu depinde de situația informaiei statistice prezentate în diferite formate. CoSSI definește structura datelor statistice, a metadatelor și a publicațiilor.

Modelul CoSSI reprezintă un sistem modular DTD. Acesta constă din definiții-tip ale documentului (DTD) și este bazat pe standardele CALS, XDF, Dublin Core, are un DTD pentru matrici statistice și un DTD pentru tabele statistice, de asemenea, are DTD-uri pentru tabele statistice și pentru publicații și documente. Un fișier XML conține date, metadate și toate versiunile de limbă.

1.3. Procesul de publicare în XML este compatibil cu modelul CoSSI

Anterior, în cadrul Biroului de Statistică al Finlandei, publicarea din bazele de date, actualizarea paginilor web și producerea publicațiilor (tip rite, PDF și HTML) erau fiecare procese de producere individuale. Pentru fiecare proces erau utilizate diferite instrumente, iar munca efectuată în cadrul fiecăruia se dubla (Figura 1).

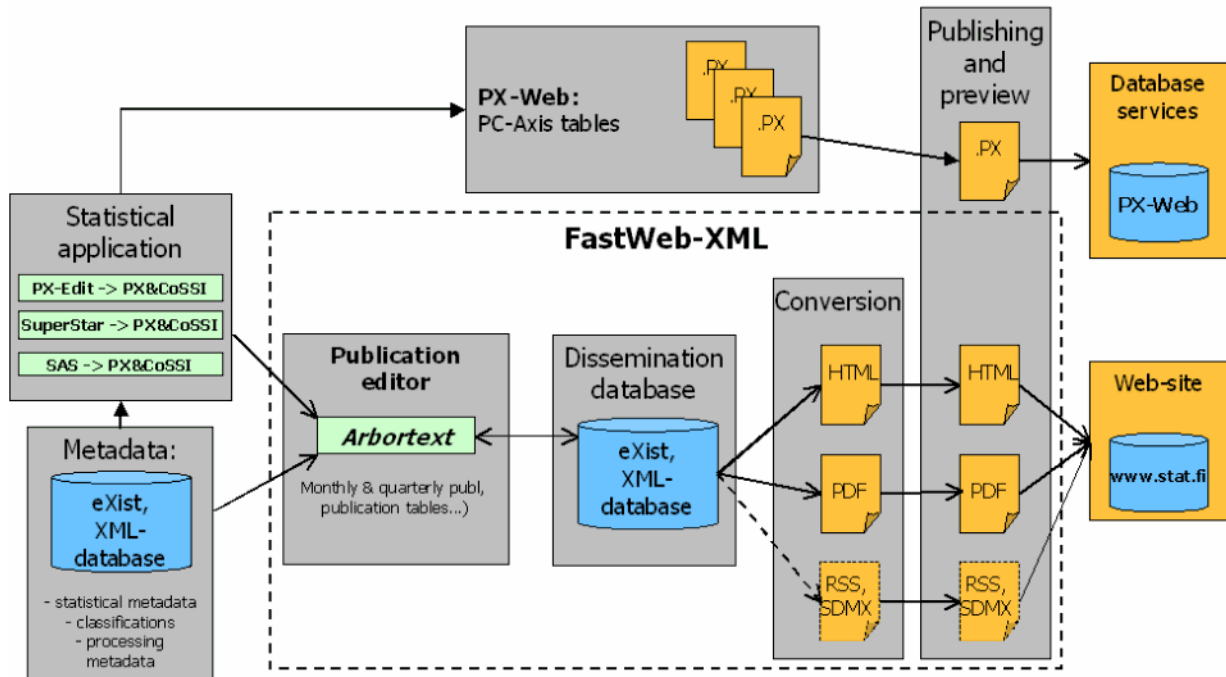
Figura 1. Procesul de diseminare - Office97²



Procesul de publicare în XML care este compatibil cu modelul CoSSI va fi implementat integral în diseminarea datelor de către Statistica Finlanda. În cadrul noului proces, diferite forme de publicații (imprimare, PDF, HTML) sunt generate automat dintr-un fișier original XML. Din același fișier original pot fi realizate și alte forme de diseminare (e-mail, RSS) (Figura 2.).

² În mai 2010 se utiliza în 40 de statistici, de exemplu pentru "Indicele Prețului de Consum"; http://www.stat.fi/til/khi/index_en.html

Figura 2. Procesul de diseminare bazat pe XML – XML i PC-Axis³



În prima fază, tablele bazei de date (.PX) sunt produse cu tablele pentru procesul de publicare în XML. SAS, SuperStar și PX-Edit produc atât tablele în format XML pentru procesul de publicare în XML, cât și tablele în format .PX pentru scopuri de diseminare în baza de date (PX-Web). La această fază, funcțiile de pre-visualizare și publicarea de diferite tipuri (tip rit, PDF, HTML), precum și tablele bazei de date (.PX) sunt combinate în cadrul aceleiași aplicații de transmitere, fișierele fiind salvate într-un singur directoriu, având o structură uniformă. Această structură tematicilor și datelor statistice devine disponibilă atât pe serviciul web, cât și în baza de date PX-Web. Denumirea fișierelor XML, HTML, PDF și .PX a fost standardizată la diferite faze a procesului de publicare.

1.4. Arhivarea unei publicații statistice electronice

În primii ani de existență a internetului, când erau create primele servicii inovative, nu se acorda o atenție foarte mare arhivării informațiilor publicate pe web. Situația nou-creată a fost examinată suficient nici de birourile naționale de statistică, din perspectiva unui sistem statistic integrat, pe termen lung. Cum va putea fi disponibil și accesibil, peste 10 ani, de exemplu, informația statistică publicată astăzi doar în format electronic? Dar peste 30 de ani? Dacă acum avem nevoie de Indicele prețului de consum de acum 30 de ani, am putea găsi datele pentru anul 1980 în seriile de publicații statistice oficiale.

În procesul de producție care este compatibil cu modelul CoSSI, originalul unei publicații este salvat în format XML într-o bază de date eXist-XML. Această publicație originală în XML conține într-un singur fișier informația propriu-zisă, metadatele statistice și metadatele în limbile dorite. Fișierul XML este arhivat în baza de date XML.

La publicare, din acest fișier original XML sunt produse formatele dorite pentru diseminare (tip rit, PDF, HTML, etc.). Volumul metadatelor variază în funcție de formatul publicației. De exemplu, în cazul publicațiilor tip rit numărul de pagini

³ În mai 2010 se utiliza în 150 de statistici, de exemplu "Prețul locuințelor";
http://www.stat.fi/til/ashi/index_en.html

devine în scurt timp un factor limitator. Conținuturile cele mai voluminoase de metadate pot fi atașate unei publicații în format HTML, prin care metadatele dorite sunt tipărite automat ca fișiere HTML, pe lângă fișierele HTML care conțin datele statistice propriu-zise.

Diferitele formate de publicații (PDF, HTML) difuzate pe serviciul web sunt publicate astfel încât adresele URL generate pentru diferite fișiere nu se modifică mai târziu. Acest fapt face posibilă utilizarea adreselor URL și în altă parte, cum ar fi în diverse studii, articole sau alte servicii web, iar utilizatorii de date vor putea verifica în orice moment documentul original indicat în referințe. Asigurarea accesibilității publicațiilor pentru utilizatori reprezintă un element esențial al unui serviciu de informare eficient și al unui sistem statistic oficial, mai ales dacă datele nu sunt publicate și pe suport de hârtie.⁴

1.5. Rolul publicațiilor statistice electronice și a bazelor de date tabelare

Datorită bazelor de date statistice, volumul informației statistice publicate în format tabelar este în continuă creștere semnificativă. Totuși, tabelele publicate în bazele de date au roluri și proprietăți diferite decât cele difuzate în publicațiile electronice sau tipărite. Aceste diferențe trebuie examinate mai îndeaproape.

Tabelele incluse în bazele de date sunt, de regulă, ca tabele pe serii de timp, conținutul cărora este completat regulat. Tabelele din bazele de date nu sunt arhivate la fel ca tabelele publicate în versiune tipărită, HTML sau PDF (sau ca tabelele publicate în alte formate, cum ar fi Excel). De exemplu, revizuirea datelor din seriile de timp, corectarea erorilor sau modificarea clasificărilor modifică conținutul tabelului din baza de date în mod retrospectiv, chiar și în cazul datelor mai vechi, uneori și după mai mulți ani de la prima publicare.

Astfel, în tabelele bazelor de date, chiar și datele observațiilor mai vechi sunt publicate întotdeauna conform celei mai recente situații. Din această cauză, seriile de timp publicate în bazele de date de asemenea pot devia de la datele ce se referă la perioade mai vechi și care au fost publicate ca date preliminare în versiuni electronice sau pe hârtie, sau care au fost revizuite în alt mod sau corectate ulterior primei lor publicări. Din acest motiv, stabilirea legăturii între publicația electronică și tabelele corespunzătoare din baza de date nu este pe deplin justificabilă. Datele corectate ulterior, după care datele din ele și cele din publicațiile electronice nu mai corespund.

La Biroul de Statistică al Finlandei, tabelele create din baza de date PX-Web nu sunt conexe la alte baze de date de referință, în timp ce baza de date PX se creează pe server din fișierele PC-Axis. În legătură cu introducerea noului sistem de publicare XML, fiecare fișier PC-Axis publicat în serverul PX-Web este salvat automat și într-un directoriu de date statistice. Astfel, publicațiile difuzate pe pagina de start a unui set de date statistice și fișierele PC-Axis salvate în același directoriu, pot fi ușor conexe. Ulterior publicațiile de format PDF sau HTML difuzate pe paginile de start ale statisticilor și tabelele corespunzătoare din bazele de date (în format PC-Axis) sunt arhivate nemodificate. Acest fapt are două avantaje semnificative: 1) este rezonabil ca și tabelele publicate în baza de date să fie conexe la publicațiile, deoarece acestea rămân neschimbate, 2) în acest fel se creează și o arhivă pentru tabelele publicate în baza de date PX-Web.

⁴ Arhiva publicațiilor statistice: http://www.stat.fi/til/arkisto/index_en.html

2. Difuzarea prin mai multe canale în practica Biroului de Statistic al Finlandei

2.1. Noi practici de difuzare a statisticilor

Biroul de Statistic al Finlandei și-a înnoit practicile de difuzare a statisticilor în limba finlandeză în 2004, iar în Suedeză și Engleză – în 2005⁵. Potrivit noilor practici de publicare, o publicație de format standard este produsă pentru toate statisticile de fiecare dată când sunt compilate noi statistici și datele sunt puse la dispoziția utilizatorilor. Înainte de 2004, noile statistici erau publicate în comunicate de presă, care erau difuzate în număr de 300 anual. Noile practici de publicare au făcut publicarea datelor mai comprehensivă. În jur de 650-680 de publicații statistice au fost produse anual, fapt ce indică o creștere considerabilă a numărului de publicații.

Pentru publicațiile suedeză și engleză, noile proceduri au fost adoptate în iunie 2005. Odată cu această schimbare, numărul de publicații statistice în suedeză a crescut de la aproximativ 120 de comunicate de presă, până la același număr ca și în cazul publicațiilor în finlandeză – adică în jur de 650-680 de la începutul anului 2006. Numărul publicațiilor statistice în limba engleză a crescut de la 120 la 350 anual în 2010⁶. Obiectivul este de a mări și mult numărul publicațiilor în engleză.

La modernizarea practicilor de publicare s-a ținut cont de obiectivele și cerințele procesului de publicare în XML, care încă este în dezvoltare, din punct de vedere a arhivării și distribuției prin mai multe canale. Toate publicațiile statistice sunt arhivate în serviciul online și adresele URL ale acestora vor rămâne neschimbate.⁷ Structura publicațiilor statistice a fost definită cu exactitate, pentru a satisface cerințele distribuirii prin mai multe canale. De exemplu, titlurile publicațiilor statistice și primul alineat sunt transmise automat prin distribuția prin e-mail și RSS.⁸

Sistemul de publicare al Biroului de Statistic al Finlandei, bazat pe XML a devenit operațional în primăvara 2007. Publicația lunară Sondajului Forței de Muncă, care a servit drept proiect pilot realizat în XML, a fost creată cu editorul Arbortext XML la 22 mai 2007 și publicată pe serviciul online al Biroului de Statistic al Finlandei în formate HTML și PDF, în finlandeză și în engleză (doar publicațiile statistice fiind difuzate în suedeză), precum și în format tipărit în Finlandeză.⁹ Un singur fișier original XML în trei limbi poate fi utilizat pentru a publica automat pagini HTML în finlandeză, engleză și suedeză (aproximativ 90 de fișiere HTML), a compila o publicație PDF în finlandeză și engleză, un fișier PDF în finlandeză pentru tipărire și distribuție prin e-mail și RSS în finlandeză, engleză și suedeză.¹⁰

Până în mai 2010, publicațiile a 150 din cele 200 de statistici ale Biroului de Statistic al Finlandei au fost create prin intermediul noului proces XML, atât în format HTML, cât și PDF. Până acum, noul proces de publicare a fost utilizat pentru a produce peste 1000 de titluri de diferite publicații în finlandeză, engleză și suedeză (în formate HTML și PDF).

⁵ La 26 iulie 2005 au fost lansate paginile web noi în engleză ale Biroului de Statistic al Finlandei; http://tilastokeskus.fi/ajk/poimintoja/2005-07-26_webpages_en.html

⁶ Calendarul publicațiilor 2010: http://tilastokeskus.fi/ajk/julkistamiskalenteri/julkistamiskalenteri_aika2010_en.html

⁷ Arhiva publicațiilor statistice : http://tilastokeskus.fi/til/arkisto/index_en.html

⁸ Ultimele publicații statistice de la Biroul de Statistic al Finlandei : http://tilastokeskus.fi/media/rss/2.0/tk_en.rss

⁹ Conținutul paginii de start a Sondajului Forței de Muncă se extinde: http://tilastokeskus.fi/til/tyti/tyti_2007-05-22_uut_001_en.html

¹⁰ Sondajul Forței de Muncă > 2010 > martie; http://tilastokeskus.fi/til/tyti/2010/03/tyti_2010_03_2010-04-27_tie_001_en.html

2.2. Modificarea designului publica iilor

În vechiul proces de publicare (Figura 1), datele con inute într-un set erau publicate 1) în mod electronic, online (minimum obligatoriu, în caz de necesitate erau publ icate i tabele anexate, figuri i un articol mai mare), 2) pe suport de hârtie i 3) ca tabele în baza de date. Toate cele trei procese de publicare erau separate i utilizau diferite instrumente în producerea tabelor i difuzarea datelor. Datele din diferite canale de distribu ie nu erau coordonate i nu formau un tot integru consecvent.

Sec iunea de text a vechiului Sondaj al For ei de Munc era o publica ie din dou coloane în limba finlandez . Tabelele erau în dou limbi – finlandez i englez , fi- ind optimizate pentru paginile de format A4. Publica ia era disponibil contra plat în format PDF i tip rit . Împreun cu versiunea de tipar, pe pagina de start erau d i- fuzate online, gratuit o publica ie statistic i câteva tabele i figuri anexate, produ- se într-un alt proces de publicare i care con ineau date diferite.¹¹ Sondajul revizuit al For ei de Munc este o publica ie comprehensiv compus din sec iunea de text i o anex cu un tabel mare i o figur , precum i o descriere a calit ii.¹² Publica ia este disponibil gratuit atât în format HTML, cât i în format PDF, iar versiunea pe suport de hârtie poate fi comandat contra plat .

Noul sistem de publicare (Figura 2) va combina con inutul publica iilor statistice difuzate online (i anexele cu tabele i figuri care le completeaz), precum i con- inutul publica iilor tip rite anterior. Aceast revizuire nu este doar o chestiune de r u- tin ce const în copierea datelor anterioare în noi instrumente, ci înseamn modif i- carea designului con inutului în întregime. Obiectivul este de a p stra i în viitor forma publica iei în pachetul de date publicate .

2.3. Publica ii statistice de baz

Transferul publica iilor într-un nou sistem de difuzare este realizat cu scopul de a difuza datele de baz privind statisticile oficiale. O publica ie de baz este o entitate ce descrie un singur set de date statistice într-o anumit perioad de referin i in- clude o sec iune de text comprehensiv , anexe cu tabele i figuri i o descriere a cal- it ii standardizat . Dac este necesar, într-o anumit perioad de referin pot fi difuzate câteva publica ii de baz dintr-un set de statistici necesitate, inclusiv o pu- blica ie preliminar i alte publica ii con inând date finale¹³ sau, de exemplu, câteva publica ii axate pe diferite tematici.

Obiectivul este de a transfera pe larg publica iile de baz din statistic în noul pro- ces de publicare în XML. Toate cele 200 de statistici produse vor fi convertite în noul sistem c tre sfâr itul anului 2010. Pe lâng modific rile tehnice (inclusiv pr o- ducerea tabelor în format XML), este modificat i designul con inutului publica i- iilor (textele, tabelele, structura comun a publica iei), este revizuit politica de pu- blicare a institu iei (taxa pentru publicare, dac publica iile vor fi sau nu disponibile în format de tipar, etc.), de asemenea, a fost modificat conceptul de divizare a seri i- lor statisticilor oficiale.

2.4. Modificarea designului tabelor

Rolul tabelor într-o publica ie de baz dintr-un set de statistici este de a descrie principalele puncte în mod cât mai concis. De i versiunea PDF a publica iilor este, de asemenea, creat în scopul distribuirii prin mai multe canale , principiul este ca versiunea electronic s fie creat în primul rând. Datele con inute într-o publica ie

¹¹ Sondajul For ei de Munc , martie 2007, (publica ia statistic este în toate limbile, tabelele din an e- x sunt doar în finlandez): <http://tilastokeskus.fi/til/tyti/2007/03/index.html>

¹² Sondajul For ei de Munc , iunie 2008 : http://tilastokeskus.fi/til/tyti/2008/06/tyti_2008_06_2008-07-22_tie_001_en.html

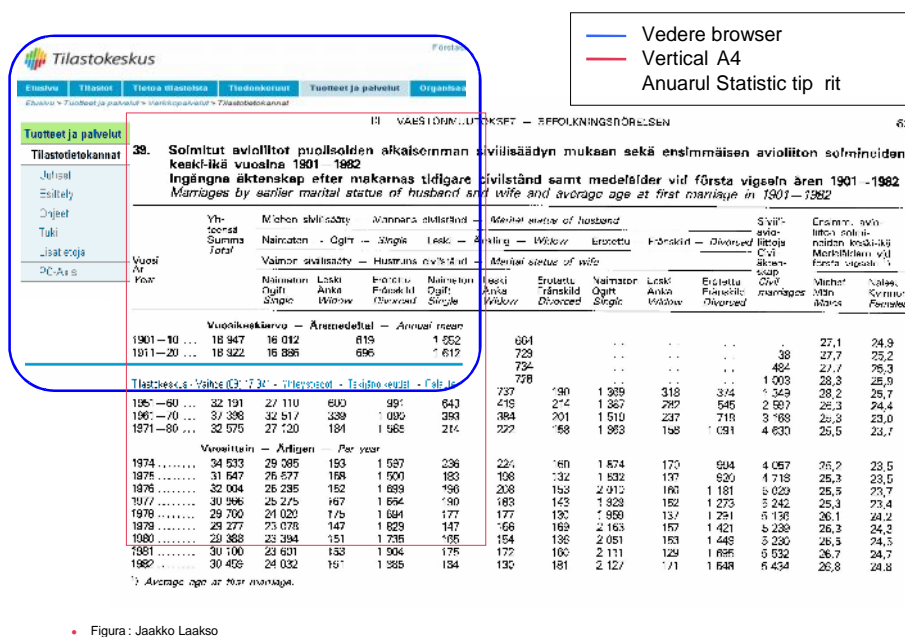
¹³ Conturi financiare: http://tilastokeskus.fi/til/rtp/tie_en.html

sunt optimizate în primul rând pentru a putea fi utilizate cu un browser Internet. Deoarece tabelele publicate trebuie să funcționeze adecvat în format HTML, acestea trebuie să fie mici și să se refere la un singur subiect.¹⁴

Deși argumentul pentru tabelele mici îl constituie restricțiile de spațiu și posibilitatea de utilizare, folosirea tabelelor mici poate fi justificată și de serviciul informațional. Mesajul experților în utilizare este clar: utilizatorii serviciilor Internet doresc să acceseze informații pe web rapid și într-un format simplu de înțeles. Tabelele mari și complicate nu satisfac aceste necesități informaționale. Rolul figurilor de asemenea este de a prezenta fenomenele cheie într-un mod compact.

Figurile au reprezentat mereu o formă convenabilă de prezentare a unor aspecte concrete, de aceea la modificarea design-ului publicatelor, nu sunt necesare modificări majore ale figurilor. Oricum, numărul figurilor trebuie sporit considerabil, deoarece proprietățile acestora le fac potrivite pentru utilizarea online. Până acum acest lucru nu era posibil din cauza lipsei unor instrumente avansate de creație grafică. Astăzi însă se încearcă aplicarea acestor instrumente.

Figura 3. Trei forme de tabele: tip rite, A4, browser



Rolul tabelor din bazele de date este semnificativ în cazul transferului de la publicațiile vechi tip rite la publicațiile electronice. Tabelele cu un conținut amplu de date trebuie să fie divizate în tabele mici (care pot fi utilizate în browser) ce tratează un aspect, să fie transferate în baza de date pentru publicare. În practică, majoritatea tabelor din publicații, care sunt mai mari decât o pagină de format A4, sunt transferate în baza de date.

Astfel, publicația de bază dintr-un set de statistici poate fi divizată în conținutul electronic difuzat online în două formate diferite și interfețele utilizatorilor (HTML și PDF), și tabelele publicate într-o bază de date statistică. Pentru producătorul unei publicații, compilarea publicației astfel încât aceasta să funcționeze normal în toate formatele și în diferite situații de operare, constituie o provocare importantă. Acesta este un caz de distribuție prin mai multe canale la scară deplină.

Disponibilitatea statisticilor oficiale în timp se reduce enorm dacă tabelele publicate în baza de date nu sunt arhivate și stocate pentru o perioadă îndelungată de timp,

¹⁴ Telecommunicații 2008; http://www.stat.fi/til/tvie/2008/index_en.html

într-un loc unde utilizatorii le-ar putea accesa. Pentru baza de date tabelar , Biroul de Statistic al Finlandei utilizează aplicația bazei de date PX-Web, care nu are legături cu nici o bază de date generală; baza de date PX-Web se bazează pe fișiere statice PC-Axis publicate pe server. Biroul de statistic al Finlandei a început arhivarea diferitelor versiuni de tabele din baze de date, publicate în format PC-Axis, fapt ce va permite inter-conexarea permanentă între publicațiile ce reprezintă un set de statistici într-o perioadă anumită de timp și tabelele bazei de date care îi corespund, astfel încât legăturile să nu fie distruse la arhivarea publicațiilor și a tabelor sale.¹⁵ Acest fapt va permite, de asemenea, crearea legăturilor permanente între tabelele compacte ale publicațiilor și tabelele surse voluminoase publicate în baza de date, precum și între versiunile arhivate ale acestora.

2.5. Interpretarea datelor – metadatele

Metadatele – variabile, concepte și definiții, descrieri ale calității, etc. – au o importanță deosebită pentru interpretarea și utilizarea corectă a datelor. În trecut, publicarea metadatelor era limitată în publicațiile scrise în funcție de spațiul disponibil. Metadatele publicate ca parte a publicațiilor tipice erau disponibile și rămăneau la fel și în arhive. În arhivarea diverselor publicații mai vaste de metadate care serveau ca ghid al utilizatorului era simplă. În formatul electronic spațiul disponibil pentru publicarea documentelor cu metadate nu mai este limitat; toate documentele cu metadatele relevante acum pot fi publicate. Documentele electronice cu metadate trebuie dotate cu ID-urile necesare, la fel ca și în publicațiile, pentru a face mai simplu și sigură informarea ilor.

Un mod simplu de a asigura că metadatele necesare pentru interpretarea datelor însoțesc mereu datele propriu-zise este de a le publica pe ambele în același pachet. Acest fapt este valabil și pentru publicațiile electronice. Din acest motiv, între altele, publicațiile produse prin intermediul sistemului de publicare XML al Biroului de Statistic al Finlandei, vor avea întotdeauna în anexă o descriere a calității. Potrivirea fiecărei versiuni de descriere a calității la publicația propriu-zisă este simplă dacă acestea două sunt combinate într-un singur pachet.¹⁶ Lista conceptelor și a definițiilor este publicată prin intermediul bazei de date a conceptelor și corespunde statutului ultimei publicații; arhiva nu este disponibilă utilizatorilor informației.¹⁷ Prin anexarea secunzii concepte și definiții la fiecare publicație, ar fi ușor de asigurat că metadatele adecvate însoțesc permanent publicația. Pe de altă parte, clasificările sunt deseori atât de extensive, încât nu este practic ca acestea să fie atașate integral la publicație. În astfel de cazuri este important ca versiunile anterioare ale clasificărilor să fie stocate pentru a putea fi accesate de către utilizatori.¹⁸ Evident, ulterior și acestea trebuie publicate astfel încât clasificarea integrală să poată fi stocată pe adresa sa originală URL și să fie accesibilă utilizatorilor și în viitor.

3. Surse:

[1] Heikki Rouhuvirta, Markku Huttunen. How NSOs can respond to changing user needs in the Internet era. Statistical Journal of the United Nations Economic Commission for Europe, Volume 20, Number 1 (2003), pp. 55-69.

[2] Heikki Rouhuvirta, Lehtinen Harri, Common Structure of Statistical Information (CoSSI) <<http://www.stat.fi/rossi>>.

¹⁵ Populația după genuri și zone 31.12.2007 și sporul populației (fișier arhivat PC-Axis): http://www.stat.fi/til/vaerak/2007/vaerak_2007_2008-03-28_tau_101_en.px

¹⁶ Migrația, 2009 : http://tilastokeskus.fi/til/muutl/2009/muutl_2009_2010-04-22_en.pdf

¹⁷ Migrația, Concepte și definiții : http://tilastokeskus.fi/til/muutl/kas_en.html

¹⁸ Clasificarea industrială – versiuni : http://tilastokeskus.fi/meta/luokitukset/toimiala/versio_en.html