

INTEGRACIÓ DE DADES PER ESTUDIS DE SOSTENIBILITAT

Jordi Bofill*, José Juan de Felipe**

Càtedra UNESCO de Sostenibilitat
carrer Colom, 1
08222 - Terrassa, Spain
Phone: +34 93 739 86 84

<http://www.catunesco.upc.edu>

jordi.bofill@upc.edu*, felipe@mmt.upc.edu**

Tema del congrés: Temes transversals

Tòpics transversals: Tecnologies per a la sostenibilitat

Paraules clau: Mesura de la sostenibilitat, Territori, Integració de dades, Web semàntica, Software lliure

Resum

Els estudis de sostenibilitat es poden realitzar en àmbits concrets com poden ser el cicle de l'aigua, les migracions de població o la mobilitat territorial; si l'àmbit d'estudi és més global o general, llavors els models conceptuals aplicats (com poden ser els de Gallopin, Bossel, etc.) són importants. En aquests tipus d'estudis calen un nombre de dades considerables: dades econòmiques, socials, territorials, ambientals, etc. Les fonts d'informació per obtenir les dades són també nombroses i diverses. Aquest fet ha motivat la concepció de diferents indicadors i índex per facilitar aquestes visions més amplies. No obstant el nombre de dades necessàries continua creixent, més encara si tenim en compte que l'anàlisi evolutiu requereix de la repetició de mesures per obtenir sèries temporals.

Una part important de les dades provenen de fonts públiques (Idescat, Ine, Eurostat, Conselleries, Ministeris, etc.), que tenen temporalitats i cobertures diferents segons els organismes. Per altre banda, amb l'evolució de la web estan emergent noves fonts de dades, com la web semàntica o web de dades. La web semàntica enllaça dades de la mateixa manera que ara s'enllacen documents. Aquests lligams entre dades de diferents fonts facilita als rastrejador d'informació l'accés a dades en funció de la definició semàntica de la pròpia dada.

Per realitzar els estudis de sostenibilitat és habitual recolzar-se en fulls de càlcul, per la seva agilitat i facilitat d'ús. Aquests fulls fan diverses funcions: eina matemàtica, eina de presentació de resultats i base de dades. El "tot en un", el "swiss army knife", és una eina útil, però és important evolucionar i incorporar noves estratègies d'anàlisi.

En aquest treball presentem l'estat actual d'una innovadora eina d'estudis de sostenibilitat territorial. Aquesta eina incorpora visions territorials, conceptuals i temporals de les dades, integrant fins i tot fonts de dades semàntiques; s'ha desenvolupat totalment usant software lliure, el llenguatge de programació ha estat Python i el framework de desenvolupament Django.

En una següent fase de la recerca incorporarà un model d'eines informàtiques d'anàlisi, de forma que es puguin escollir diferents software per l'anàlisi i la presentació de resultats (estadístiques, gràfiques, modelització, etc.)

A nivell pràctic es presenta la implementació per l'estudi de sostenibilitat per municipis i comarques

catalanes.

1 Introducció

L'estudi de la sostenibilitat implica l'estudi d'un sistema complex. Com es pot desenvolupar una eina de suport a l'estudi d'un sistema complex? Si aquesta eina ha de facilitar l'anàlisi d'un sistema complex, com es pot fer per que la concepció d'aquesta eina introdueixi el mínim de condicionants al sistema? A una eina d'aquest tipus li demanaríem:

- Que els límits del sistema els introdueix l'investigador. Deixar en mans de l'investigador el major nombre de decisions possible.
- Que l'eina imposi el mínim nombre de requisits, i que siguin deguts a la tecnologia emprada.
- Que la lògica de l'eina es limiti als conceptes clau de la mesura de la sostenibilitat.
- Que l'eina sigui el més genèrica possible.

A continuació s'expliquen les característiques d'aquesta eina de suport als investigadors que realitzen estudis relacionats amb la sostenibilitat territorial. L'eina es pot aplicar a territoris que tinguin organitzacions territorials diferents. El prototipus experimental s'ha aplicat a Catalunya, particularment a les comarques de Girona.

1.1 Mesura de la sostenibilitat. Models Conceptuals.

Per mesurar la sostenibilitat s'utilitzen variables, indicadors i índexs:

- Variable: element mesurable canviant en el temps i/o espai emprat per a descriure fenòmens o magnituds. Les variables de mesura de la sostenibilitat són totes les dades d'interès que es poden trobar i que són rellevants per la finalitat.
- Indicador: és una variable. L'objectiu d'un indicador és poder avaluar l'estat d'un objecte o d'un fenomen en un moment determinat i poder comparar-lo amb el seu estat passat o comparar-lo amb objectes similars.
- Índexs: s'obtenen mitjançant l'agregació i ponderació d'indicadors i/o de variables

El Desenvolupament Sostenible és un àrea d'estudi amb diversitat de principis, que ha generat molts sistemes d'indicadors i que, també, ha propiciat el desenvolupament de diferents marcs conceptuals. Les diferències principals entre els marcs són la manera en què conceptualitzen les dimensions principals del DS, les interconnexions entre aquestes dimensions, la manera com s'agrupen, i els conceptes que justifiquen la selecció i agregació d'indicadors.

2 Unitats lògiques de l'eina

2.1 Mesures

L'investigador pot definir les mesures que necessiti per la seva recerca. Les mesures poden ser numèriques (variables, indicadors o índexs), texts o lògics. Cada mesura s'identifica amb un nom i porta associat la seva descripció, les unitats emprades i altres observacions. Es pot definir que una mesura sigui el resultat del càlcul senzill d'altres mesures.

Tipus	Valors
Numèric	Qualsevol valor numèric
Text	Text lliure fins a 125 caràcters
Lògic	Veritat/Fals
Enllaç	Enllaç d'internet (http, ftp, ...)
Imatge	Enllaç a un arxiu gràfic local

2.2 Marcs conceptuals

En la seva concepció l'eina no està limitat ni a les dimensions principals de desenvolupament sostenible (Societat-Economia-Ambient) ni a un model conceptual determinat, sinó que permet definir diferents marcs conceptuals¹, de forma que l'investigador és qui ah de definir el seu propi model o utilitzar un definit prèviament.

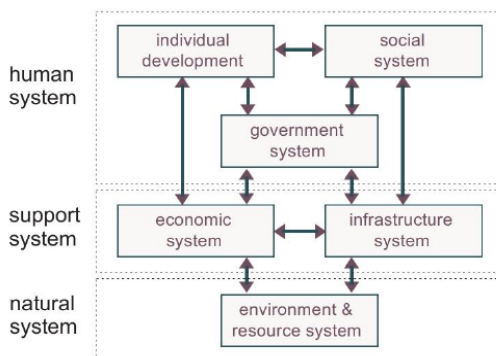


Figura 1: Model conceptual de Bossel

Per integrar els marcs conceptuals en l'eina, aquests es defineixen com un conjunt d'àrees conceptuals, a on cada àrea agrupa les mesures que intervenen. Per exemple, a la figura 1 tenim l'esquema del model conceptual de Bossel (Bossel 1999) i la figura 2 les àrees conceptuals tal com es veuen en el prototipus experimental. Les associacions entre les àrees d'un marc conceptual i les mesures, és una decisió de l'investigador.

L'associació d'una mesura a una àrea conceptual permet introduir variables de seguiment d'aquella mesura. Per exemple, es pot introduir per cada mesura una llegenda de llindars o semàfors. Una mateixa mesura pot estar associada a més d'un model conceptual amb diferents valors per les variables de seguiment (figura 3).

Area marc	Nom	Descripció
Bossel->des. individual	des. individual	Desenvolupament individual: dret humans, salut, participació, ...
Bossel-> sistema social	sistema social	població, seguretat social, distribució renda, ...
Bossel->govern	govern	polítiques de finances, impostos, democràcia, immigració, sistema lega, ...
Bossel->infraestructura	infraestructura	transport, serveis de salut, sistemes de distribució aigua, aliments, etc ...
Bossel->sist. economic	sist. economic	producció i consum, moneda, comerç, ...
Bossel->recursos i medi	recursos i medi	Medi natural, ecosistemes, pol·lució, ..

Figura 2: Model conceptual de bossel en el prototipus experimental

¹ En el desenvolupament de l'eina s'utilitza el nom de 'marc conceptual' en lloc de 'model conceptual' i 'camps' en lloc de 'mesures'.

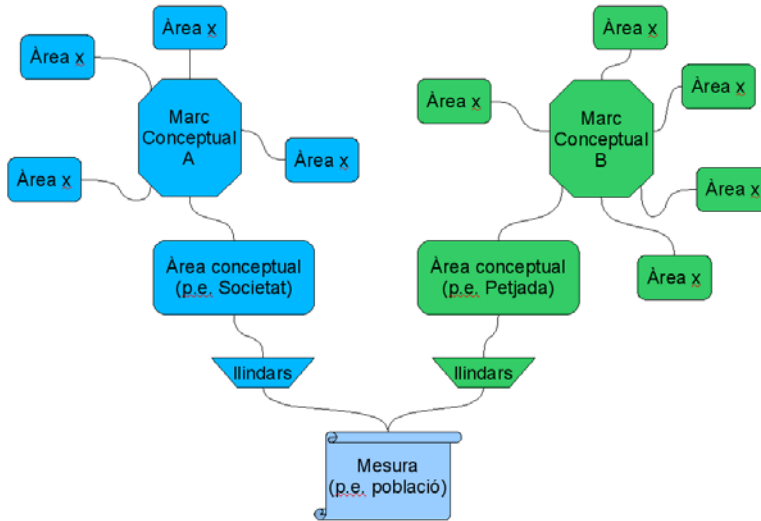


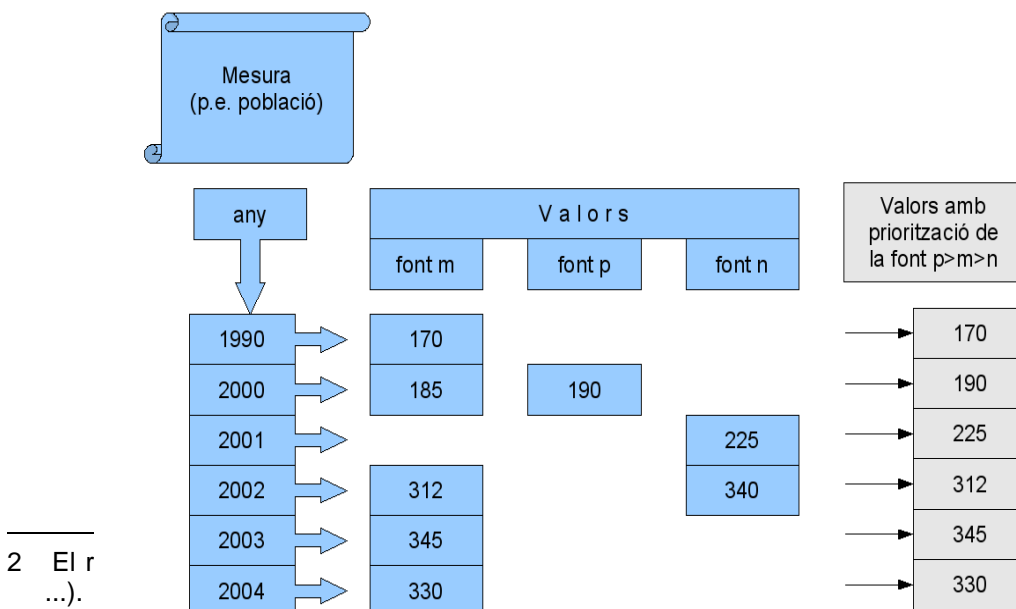
Figura 3:

Dos marcs conceptuals utilitzant una mateixa mesura.

2.3 Valors

Per cada mesura es guarda el valor i la data de la mesura, podent-se obtenir fcilment sries temporals. El rang temporal per defecte s l'annual².

Com a característica innovadora, l'eina permet guardar ms d'un valor d'una mesura en una mateixa data, es a dir, valors repetits. Aquests valors amb repetico han de ser provets per fonts de dades diferents. Es pot establir una prioritziacio de les fonts de dades pel cas de que existeixin valors repetits. Aquesta facilitat permet realitzar simulacions del tipus 'que passaria' si el valor de unes dades fos diferent, sense necessitat de variar el valor de la dada original. Per aixo, l'investigador crearia una font de dades prpia amb una prioritat diferent a les altres. A la figura 4 es presenta un exemple de com seria una seleccio d'una mesura que tingues valors repetits de fonts de dades amb prioritats diferents.



2.4 Font s de dade s

Es pot gestionar les diferents fonts de

ensual,

² El r
...).

Figura 4:

Exemple seleccio de dades en el cas de valors repetits en una mateixa data, amb prioritziacio de les fonts de dades.

dades utilitzades, de forma que es pot saber cada mesura de quina font de dades prové. Actualment no es pot gestionar automàticament la descàrrega i l'obtenció de dades, encara que esta previst facilitar el coneixement del flux de treball a utilitzar per les descàrreges.

2.4.1 Web de dades o web semàntica.

Les especificacions dels enllaços entre les dades de la web semàntica es realitzen mitjançant Resource Description Framework (RDF). Aquest model de metadades es base en transformar les descripcions de dades a un format 'subjecte-predicat-objecte'. El subjecte seria el recurs i el predicat (o propietat) especificaria la relació entre el recurs i l'objecte. Per exemple, "El municipi de Terrassa té una superfície de 70,2 km²", en RDF seria un triplet de text estructurat a on el subjecte seria "El municipi de Terrassa", que té una propietat "superfície" i l'objecte seria el valor de la superfície "70,2 km²".

En el prototipus experimental s'ha introduït la consulta a dades de la web semàntica, concretament la base de dades semàntica de Wikipedia, <http://dbpedia.org>. En les consultes territorials, el sistema comprova si existeixen dades a dbpedia per aquell territori i, en funció del tipus de dada, les configura per la seva visualització. De la informació disponible de wikipedia s'han seleccionat les propietats³ de la taula 1.

Propietat	Descripció
http://dbpedia.org/property/population	Habitants
http://dbpedia.org/property/datePopulation	Data del recompte del habitants
http://xmlns.com/foaf/0.1/page	Enllaç a la pàgina de wikipedia
http://dbpedia.org/property/locationMap	Imatge per localitzar el territori.
http://www.w3.org/2003/01/geo/wgs84_pos#lat http://www.w3.org/2003/01/geo/wgs84_pos#long	Coordenades geogràfiques
http://dbpedia.org/property/area	Superfície del territori
http://xmlns.com/foaf/0.1/img	Imatge del territorial

Taula 1: Propietats de wikipedia.org seleccionades.

2.5 Organització Territorial

Alhora de definir l'organització territorial més convenient hem de tenir en compte tres coses:

1. Un mateix territori pot tenir alhora diferents subdivisions. Per exemple, els Estats tenen subdivisions polítiques, administratives, militars, etc. (circumscripcions territorials)⁴. En el cas d'Espanya, hi ha una primera divisió per autonomies i la comunitat autònom de Catalunya té com a divisions les províncies, les comarques i els municipis.
2. Les subdivisions, encara que molt estables, poden canviar amb el temps. Per exemple, hi ha sectors de la societat catalana que demanen la reintroducció d'una antiga subdivisió territorial intermunicipal: la vegueria.
3. El territori objecte de l'estudi pot ser diferent al d'una subdivisió territorial política o administrativa.

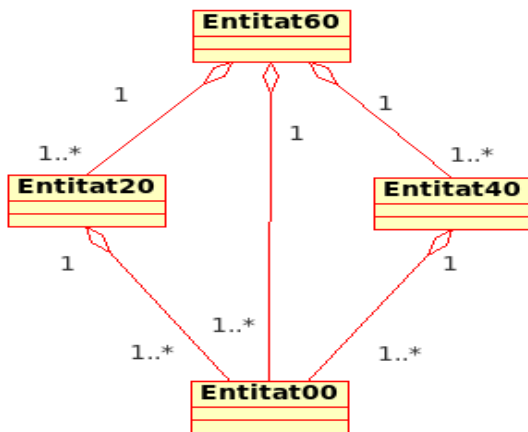
Aquestes tres raons han motivat modelitzar l'eina de forma que no tingui un àmbit territorial

³ Una llista bastant completa de propietats es pot trobar a <http://dbpedia.org/page/Barcelona>

⁴ http://en.wikipedia.org/wiki/Table_of_administrative_country_subdivisions_by_country

específic, prèviament definit. És per això que l'eina proporciona unes agrupacions territorials genèriques, anomenades Entitats Territorials, que es defineixen durant la implementació pràctica de l'eina.

El sistema parteix de la base que existeix una entitat territorial bàsica: són les entitats més petites (o atòmiques) que, sumades totes, configuren la totalitat del territori objecte d'estudi. En el cas de Catalunya, es podria escollir com entitat territorial bàsica el municipi, la comarca o la província, doncs qualsevol de les tres subdivisions territorials configuren la totalitat del territori de la comunitat autònoma. Les organitzacions territorials que defineixen un territori més gran que l'entitat territorial bàsica, s'han de poder 'construir' mitjançant agregacions d'aquestes. Seria el cas, per exemple, de les comarques ja que són una agregació de municipis.



L'organització territorial bàsica s'anomena Entitat00. Les altres organitzacions segueixen la nomenclatura EntitatXX a on XX és un número que indica, per exemple, si l'entitat cobreix una superfície del territori superior a les altres, encara que això no és un requisit. A la figura 5 es presenta el diagrama de classe UML⁵ del prototipus experimental, amb les següents correspondències d'entitats territorials:

- Entitat00 -> Municipi
- Entitat20 -> Comarca
- Entitat40 -> Província
- Entitat60 -> Catalunya

Figura 5: Model UML de la organització territorial.

3 Característiques de l'eina

Les característiques rellevants són:

- Eina instal·lada en un servidor a internet. Ús mitjançant navegador web, sense instal·lació local al PC.
- Definició dels models conceptuals.
- Definició de les mesures (variables, indicadors, etc.).
- Definició de l'organització territorial.
- Valors amb temporalitat anual. Valors d'una mesura repetits per una mateixa data.
- Gestió de les dades mitjançant full de càlcul.
- Generació automàtica de gràfics i informes predefinits.
- Scripts definits per l'usuari.
- Selecció de dades persistent. Filtres per facilitar la selecció.
- Enllaç a fonts de la web semàntica (actualment, wikipedia).
- Enllaç a fonts d'informació cartogràfica.

Característiques tècniques

- Llenguatge programació: python <http://www.python.org/>.
- Base de dades relacional: Sqlite, Postgres o Mysql. Prototipus experimenta basat en Sqlite <http://www.sqlite.org/>.
- Framework web: django <http://www.djangoproject.com/>.
- Servidor virtual del prototipus: GNU/Linux Gentoo, 256Mb ram, 10Gb disc.

5 UML: Unified Modeling Language. http://en.wikipedia.org/wiki/Unified_Modeling_Language

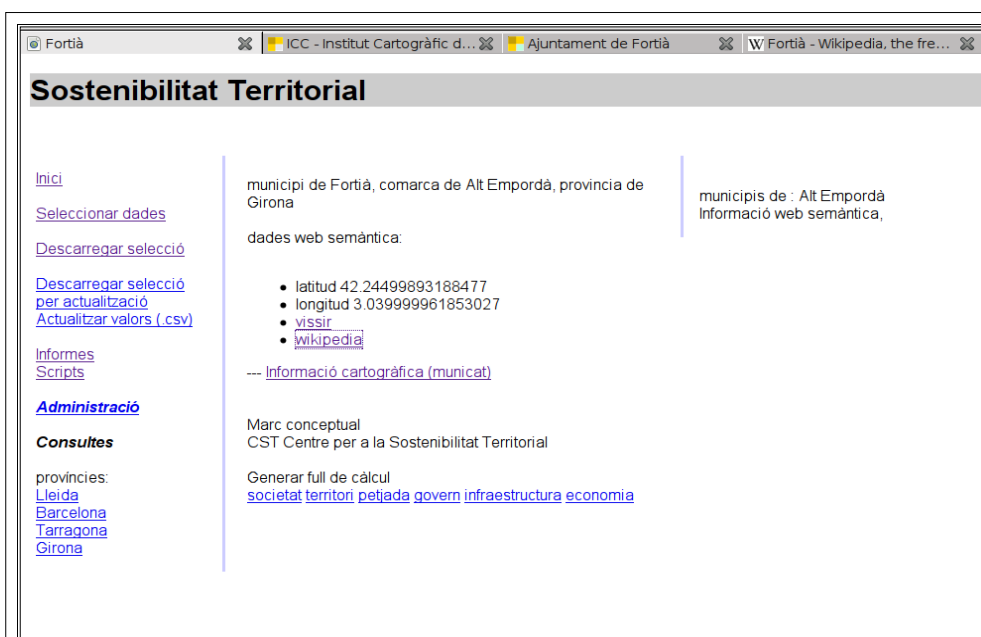
- Espai de disc ocupat per l'aplicació, inclòs la base de dades prototipus: 16Mb.

4 Prototipus experimental. Implementació a Girona.

Amb el suport de l'Observatori de Sostenibilitat de les Comarques Gironines, que ha facilitat les dades, s'ha realitzat una versió de l'eina per l'àmbit territorial de Girona. El model de dades té les següents magnituds: 9 comarques, 221 municipis, 216 mesures, 142.628 valors, 1 font de dades, 1 font semàntica i 1 marc conceptual amb 6 àrees conceptuais.

Com exemples pràctics, es mostra a la figura 6 la consulta de dades del municipi de Fortià. A partir de les dades semàntiques de latitud i longitud obtingudes de dbpedia.org s'haa generat automàticament un enllaç al visualitzador cartogràfic de Institut Cartogràfic de Catalunya i, mitjançant el codi de municipi l'enllaç a la informació cartogràfica municipal de Fortià.

A la figura 7, veiem la selecció de dades (marc, font, municipis i mesures) i, a la figura 8, la seva visualització gràfica. Els municipis seleccionats són Palafrugell, Palamós, Sant Feliu de Guíxols i Torroella de Montgrí i les mesures d'atur d'homes i de dones.



Sostenibilitat Territorial

[Inici](#)

[Seleccionar dades](#)

[Descarregar selecció](#)

[Descarregar selecció per actualització](#)

[Actualitzar valors \(.csv\)](#)

[Informes](#)

[Scripts](#)

SELECCIONAR

[Marc Conceptual](#)

[provincies](#)

[comarques](#)

[municipis](#)

[font](#)

[campes](#)

Filtrar per font:

[CST](#)

[Dèpèdia](#)

[ib](#)

[ib2](#)

[vidra](#)

[no filtrar](#)

Filtrar per ...

[economia](#)

[govern](#)

[infraestructura](#)

SELECCIONAR

Area: **economia** Font: **CST**

- ATURATS. Font: Idescat. [PBL_ATU]
- COMTES DE COTITZACIÓ. Font: Observat... [CC]
- TREBALLADORS TOTALS. Font: Observat... [TRB]
- TREBALLADORS SECTOR AGRARI. Font: O... [TRB_AGR]
- TREBALLADORS SECTOR INDUSTRIAL. Fon... [TRB_IND]
- TREBALLADORS SECTOR CONSTRUCCIÓ. Fo... [TRB_CON]
- TREBALLADORS SECTOR SERVEIS. Font: ... [TRB_SER]
- TREBALLADORS SECTOR NO CONSTA. Font... [TRB_NOC]
- NOMBRE DE TREBALLADORS EN COMTES DE... [TRBCC_GR]
- LOCALITZACIÓ DE L'OCCUPACIÓ. PERCENT... [TRBCC_GR_%]
- COMTES DE COTITZACIÓ TOTALS. Font: ... [EMP_TOT]
- COMTES DE COTITZACIÓ PER SECTORS. F... [EMP_AGR]
- COMTES DE COTITZACIÓ PER SECTORS. F... [EMP_IND]
- COMTES DE COTITZACIÓ PER SECTORS. F... [EMP_CON]
- COMTES DE COTITZACIÓ PER SECTORS. F... [EMP_SER]
- COMTES DE COTITZACIÓ PER SECTORS. F... [EMP_ALT]
- EMPRESSES COOPERATIVES [EMP_COOP]
- EMPRESSES AMB CERTIFICAT ISO 14.001 [EMP_ISO]
- EMPRESSES AMB CERTIFICAT EMAS [EMP_EMAS]
- EMPRESSES AMB CERTIFICAT CCPAE [EMP_CCPAE]

SELECCIONAT

- PBL_ATU_H
- PBL_ATU_D

Usuari: **AnonymousUser**

SELECCIÓ

Marc: **CST**

municipis: **Palafrugell, Palamós, Sant Feliu de Guíxols, Torroella de Montgrí.**

Fonts: **CST.**

Camps: **PBL_ATU_H, PBL_ATU_D.**

Figura 7: Selecció de Marcs, Municipis, Fonts i Mesures

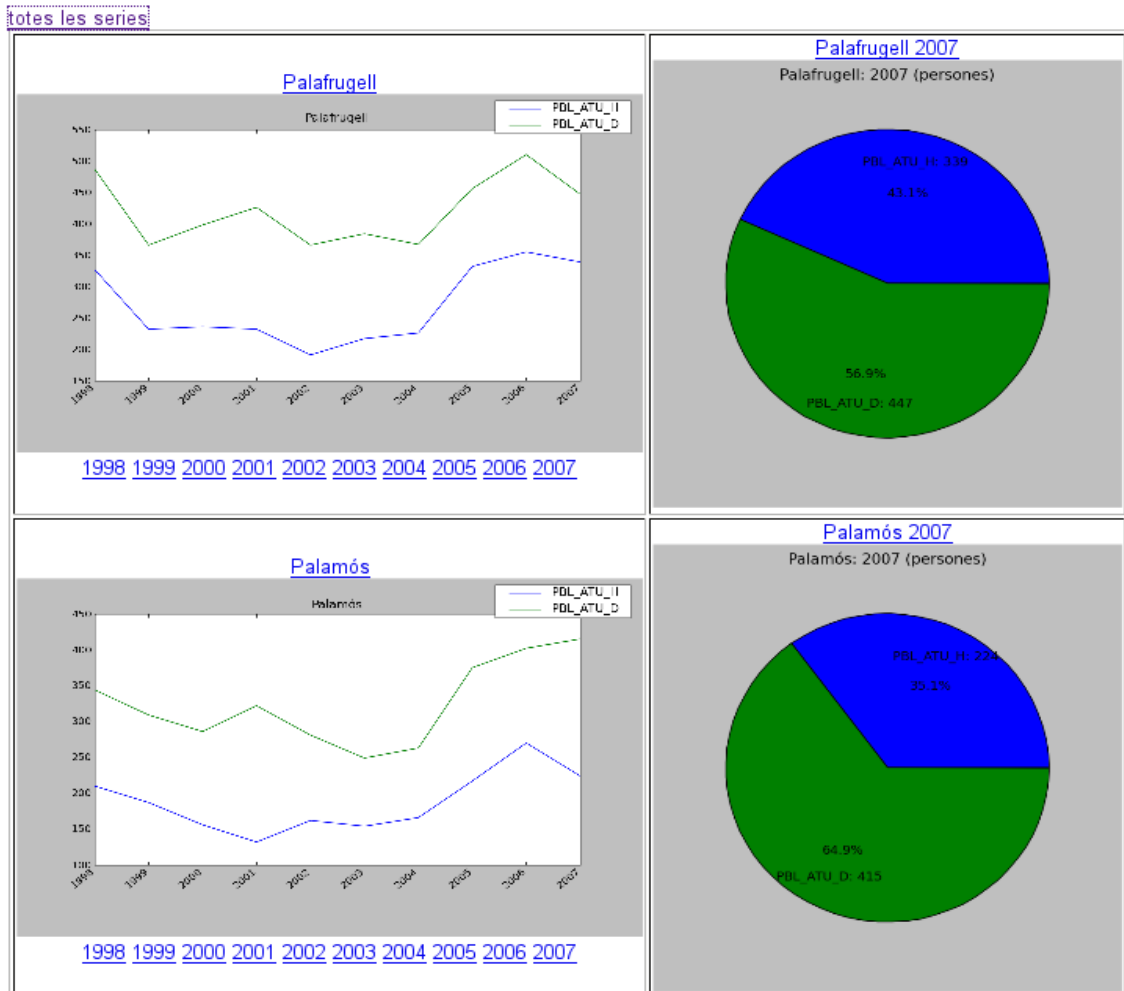


Figura 8: Visualització gràfica de les dades

5 Bibliografia

Bossel, Hartmut. *Indicators for Sustainable Development: Theory, Method, Applications*. International Institute for Sustainable Development, 1999. <http://www.iisd.org/pdf/balatonreport.pdf>

Breitman, Karin K., Casanova, Marco Antonio, i Truszkowski, Walter. *Semantic Web. Concepts, Technologies and Applications*, Springer, 2007.

Gallopin, Gilberto. *Los indicadores de desarrollo sostenible aspectos conceptuales*, Seminario de Expertos sobre Indicadores de Sostenibilidad en la Formulación y Seguimiento de Políticas, FODEPAL, 2006. <http://www.fodepal.org/Bibvirtual/semex/indicadores/homedoc.html>

García, Rolando. *Sistemas Complejos. Conceptos, métodos y fundamentación epistemológica de la investigación interdisciplinaria*. Gedisa, 2006.

Holt, Jon: *UML for Systems Engineering: watching the wheels*, The Institution of Electrical Engineers, 2004.

Holovaty, Adrian, and Kaplan-Moss, Jacob. *The Django Book*, 2008. Recurs electrònic. <http://www.djangobook.com>

Langtangen, Hans Peter. *Python Scripting for Computational Science*, Springer, 2006.

Olivé, Antoni. *Conceptual Modeling of Information Systems*, Springer, 2007.

Sureda, Bàrbara. *Proposta metodològica per a l'anàlisi de la sostenibilitat, utilitzant indicadors i índexs, implementats i analitzats amb una eina de suport al raonament. Cas d'estudi: municipi de Terrassa*, Universitat Politècnica de Catalunya, 2007.