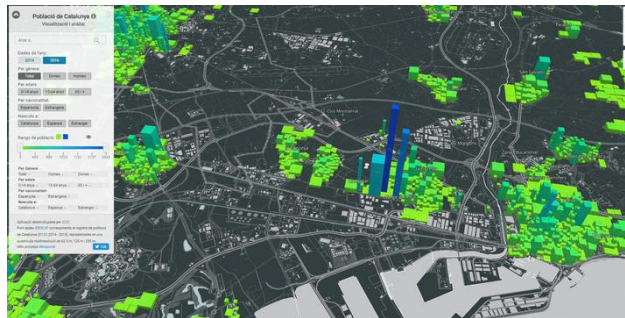


POBLACIÓ DE CATALUNYA GEOREFERENCIADA A 1 DE GENER DE 2016 v1.0 (RP2016)



Especificacions tècniques

versió 1.0

2018-12-10

Índex

1. Introducció	5
2. Identificació del producte	5
3. Sistemes de referència.....	5
4. Àmbit.....	6
5. Estructura i contingut.....	6
6. Qualitat	9
7. Metadades	9
8. Distribució	10
9. Captació.....	10
10. Manteniment	11
11. Representació.....	11
Annex A. Referències normatives	13
Annex B. Termes i definicions.....	15
Annex C. Glossari de sigles i abreviatures.....	17

1. Introducció

L'objectiu del present plec és definir el contingut, estructura i característiques del conjunt d'informació geogràfica Població de Catalunya georeferenciada a 1 de gener de 2016 (RP2016).

El plec descriu les especificacions generals del producte, així com la seva implementació en el format SHP.

Aquestes especificacions han estat elaborades per l'Institut d'Estadística de Catalunya (IDESCAT), amb el suport i la col·laboració de l'Institut Cartogràfic i Geològic de Catalunya.

2. Identificació del producte

L'RP2016 proporciona informació relativa a la distribució territorial de les principals variables demogràfiques en un grid multiresolució (de 62,5 m, 125 m i 250 m), que és compatible amb el grid estàndar europeu. Les variables publicades són el sexe, l'edat, la nacionalitat i el lloc de naixement. La font d'informació és el Registre de població de Catalunya, amb data 1 de gener de 2016.

El propòsit del conjunt d'informació geogràfica RP2016 és que la informació demogràfica contribueixi a:

- Facilitar la màxima eficiència en la planificació, gestió, execució i avaluació de les polítiques públiques.
- Afavorir la transparència en la gestió pública i, per tant, millorar el control democràtic dels poders públics, i vetllar per la confiança dels ciutadans en els governants.
- Impulsar el desenvolupament social i econòmic.
- Permetre als investigadors de les ciències socials l'accés a informació rellevant (garantint la confidencialitat d'aquesta) i així afavorir que les seves aportacions millorin les polítiques públiques.
- Possibilitar la realització d'operacions i consultes d'anàlisi o incidència territorial amb altres conjunts de dades d'informació geogràfica.

És un conjunt d'informació geogràfica vectorial en 2D.

L'entitat responsable és l'Institut d'Estadística de Catalunya.

La classificació d'aquest conjunt d'informació geogràfica dins del Pla Cartogràfic de Catalunya és la següent:

- Grup PCC- INSPIRE: III-10 DISTRIBUCIÓ DE LA POBLACIÓ - DEMOGRAFIA.
- ID conjunt PCC: 31001 Distribució de la població per seccions censals, municipis o comarques.
- idLocalProducte: poblacio-seccions-censals-municipis-comarques

3. Sistemes de referència

El sistema geodèsic de referència és l'EPSG:25831 - ETRS89 / UTM31N.

L'ordre de les coordenades és: coordenada horitzontal, coordenada vertical (X, Y), Easting, Northing.

4. Àmbit

La cobertura territorial del conjunt d'informació geogràfica RP2016 és el territori de Catalunya.

5. Estructura i contingut

5.1 Característiques del model de dades

En base a la informació continguda al Registre de població de Catalunya, el conjunt d'informació RP2016 proporciona dades sobre la distribució territorial de les principals variables demogràfiques de Catalunya (edat, sexe, nacionalitat i lloc de naixement).

En la secció 5.2 s'inclou la descripció conceptual (esquema conceptual) del model de dades, que descriu de forma genèrica el contingut d'informació del conjunt de dades, així com la seva concreció o implementació en el format SHP (esquema físic en format SHP).

Es tracta d'un únic objecte (*Poblacio*) amb una geometria i uns atributs associats.

Els trets més rellevants del model de dades implementat en SHP són:

- El tipus d'objecte geogràfic *Població* té un identificador únic (atribut *ID*).
NOTA: Quan els fitxers SHP s'importen a ArcGis, s'afegeix automàticament un atribut identificador únic addicional, propi d'aquest format, amb nom *FID* i de tipus *OBJECTID*.
- La geometria es concreta amb l'atribut *SHAPE*. És de tipus *Polygon*. Aquesta geometria es correspon al tipus *GM_Surface* previst a la norma ISO19107:2003.
- La unitat de mesura és el metre. Les coordenades estan emmagatzemades com a números reals de doble precisió, d'acord amb l'estàndard d'aquest format. La seva precisió és consistent amb la resolució de captura.

5.2 Model de dades o esquema d'aplicació

La descripció conceptual del model de dades RP2016 queda definida per l'esquema conceptual descrit seguidament. Consta d'una única classe, corresponent a l'únic objecte del model: *Poblacio*.

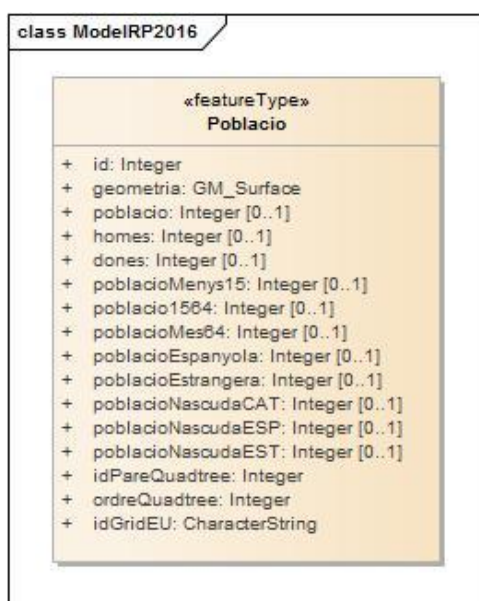


Figura 1. Esquema conceptual del model de dades de l'RP2016 (data creació: 10/12/2018).

La implementació del model de dades en format SHP queda definida per l'esquema físic descrit a continuació. L'únic objecte del model de dades s'implementa en una única taula (*Poblacio*) que conté el conjunt d'informació RP2016.

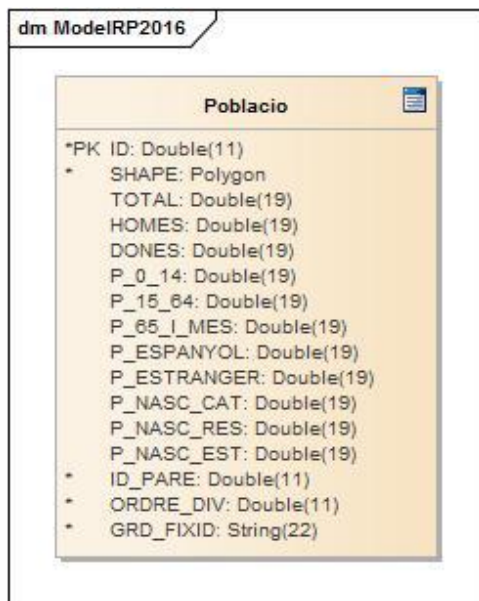


Figura 1. Esquema físic del model de dades de l'RP2016 en el format SHP (data creació: 10/12/2018).

En el format SHP, cada element d'un tipus d'objecte té associada una representació geomètrica i uns atributs alfanumèrics registrats en una taula adjunta (fitxer .dbf). Els objectes s'implementen vinculant la informació alfanumèrica emmagatzemada en camps amb els elements gràfics.

A l'apartat 5.3 es descriu l'equivalència entre l'esquema conceptual i l'esquema físic SHP, així com les característiques dels diferents atributs de l'objecte *Poblacio*.

5.3 Catàleg d'objectes geogràfics

La següent taula descriu la implementació en format SHP de l'objecte geogràfic inclòs a l'RP2016. El format de cada camp es mostra en la forma L,T{D} - on L indica la longitud del camp en número de caràcters o de dígit (xifres significatives) – segons correspongui, T el tipus (C=caràcter, N=numèric), i D és el nombre de decimals, si s'escau.

OBJECTE		Població				
NOM		Població				
DEFINICIÓ		Polígon que aporta informació de les variables demogràfiques sexe, edat, nacionalitat i lloc de naixement de la població de Catalunya.				
DESCRIPCIÓ CONCEPTUAL						
ATRIBUT	FORMAT DEL CAMP L, T, [D]	ATRIBUT	NOM	DEFINICIÓ	TIPUS DE VALOR	MULTIPLICITAT
ID	Double 11,N,0	id	Identificador del element	Identificador únic de l'objecte.	Integer	1
SHAPE	Polygon	geometria	Geometria	Representació espacial de l'objecte.	GM_Surface	1
TOTAL	Double 19,N,0	poblacio	Població	Població total de Catalunya.	Integer	0..1
HOMES	Double 19,N,0	homes	Homes	Població total d'homes.	Integer	0..1
DONES	Double 19,N,0	dones	Dones	Població total de dones.	Integer	0..1
P_0_14	Double 19,N,0	poblacioMenys15	Població <15 anys	Població amb edats menors a 15 anys.	Integer	0..1
P_15_64	Double 19,N,0	poblacio1564	Població de 15 a 64 anys	Població amb edats compreses entre 15 i 64 anys.	Integer	0..1
P_65_I_MES	Double 19,N,0	poblacioMes64	Població de 65 anys i més	Població amb edats de 65 anys i superiors.	Integer	0..1
P_ESPANYOL	Double 19,N,0	poblacioEspanyola	Població espanyola	Població de nacionalitat espanyola.	Integer	0..1
P_ESTRANGER	Double 19,N,0	poblacioEstrangera	Població estrangera	Població de nacionalitat estrangera.	Integer	0..1
P_NASC_CAT	Double 19,N,0	poblacioNascudaCAT	Població nascuda a Catalunya	Població nascuda a Catalunya.	Integer	0..1
P_NASC_RES	Double 19,N,0	poblacioNascudaESP	Població nascuda a Espanya	Població nascuda a la resta d'Espanya.	Integer	0..1
P_NASC_EST	Double 19,N,0	poblacioNascudaEST	Població nascuda a l'estranger	Població nascuda a l'estranger.	Integer	0..1
ID_PARE	Double 11,N,0	idPareQuadtrees	Identificador del pare al quadtree	Identificador únic de l'element pare d'aquest objecte en la jerarquia del quadtree. Amb aquest atribut es poden fer agregacions per passar a un grid de 500 m o d'1 km. Agregant per GR_FIXID es pot arribar directament fins a 1km.	Integer	1
ORDRE_DIV	Double 11,N,0	ordreQuadtrees	Ordre del quadtree	Número d'ordre de la divisió que correspon a l'objecte, partint del grid a 1 km. Aquests grid es subdivideixen en 64 unitats de 125 m de cantó que se numeren de l'1 al 64 començant per la cantonada inferior esquerra.	Integer	1
GRD_FIXID	String 22,C	idGridEU	Identificador del grid europeu	Correspon a la clau primària original del grid estàndard europeu 1 km (Grid_ETRS89_LAEA_ES_1K) publicat en format shape per l'European Forum for Geography and Statistics (EFGS).	CharacterString	1
RESTRICCIONS						
A les cel·les de resolució mínima (250 m), si el valor de població és menor de 17, els valors informats de tots els atributs són nuls. S'ha realitzat una supressió primària si apareix un valor inferior a 7 i les corresponents supressions secundàries en cas necessari a qualsevol resolució.						

6. Qualitat

A la següent taula es descriuen les mesures de qualitat classificades per elements de qualitat segons ISO 19157:2013 Geographic information - Data quality, i s'especifica quins són els resultats esperats de l'avaluació de la qualitat.

REQUISIT DE QUALITAT	DESCRIPCIÓ DE LA MESURA
COMPLETESA	
Conté la totalitat de les cel·les a on hi ha població. Si aquesta es inferior a 17, malgrat que els valor associats a la geometria són nuls, aquesta geometria es proporciona. NOTA: en els lloc a on no hi ha població no es proporciona cap mena de geometria.	Detecció automàtica d'elements absents del conjunt de dades, calculant el percentatge total d'elements omesos i que han de ser presents. En cas que el percentatge no sigui del 0% es revisen les dades de forma sistemàtica fins que s'assoleix aquest objectiu.
CONSISTÈNCIA CONCEPTUAL	
No hi ha dos objectes geogràfics amb el mateix identificador.	Detecció automàtica de cadascuna de les regles de consistència conceptual estipulades com a requisit de qualitat, calculant el percentatge total d'elements que les incompleixen. En cas que el percentatge no sigui del 0% es revisen les dades de forma sistemàtica fins que s'assoleix aquest objectiu.
Els valors dels atributs es corresponen al tipus de dades estipulat en el model de dades.	
EXACTITUD TEMÀTICA	
Globalment la població és exacte sense cap error. Malgrat això, l'exactitud dels valors totals de població informats per a cada cel·la no ho són estrictament (vegeu descripció de la mesura). En les simulacions fetes per mètodes de Monte Carlo s'ha obtingut una mediana de l'error del 5% quan es calculen totals de poblacions en polígons de interès per l'usuari.	L'exactitud dels valors totals de població informats per a cada cel·la no són estrictament exactes, atès que ha calgut fer una imputació de coordenades per a un 4,42% del total de la població. Addicionalment, per a evitar agregacions en els casos en que la variància de la població en cel·les germanes de la jerarquia és gran, s'han realitzat translacions que minimitzen els errors en el càlcul de la població dins de polígons aleatoris. L'estimació de la cota d'error (mediana) s'ha obtingut aplicant mètodes de Monte Carlo.

7. Metadades

Les metadades es generen a nivell de conjunt de dades i es creen d'acord amb el perfil IDEC de l'estàndard ISO 19115:2003 vigent en el moment de la seva generació. La metadada està calogada a la IDEC <http://www.ide.cat/>

Les metadades donen informació sobre les dades, la seva qualitat, el sistema de referència i les pròpies metadades.

8. Distribució

L'RP2016 es distribueix segons s'indica a la taula següent.

DISTRIBUCIÓ	ADREÇA
DADES	
Format SHP	http://www.idescat.cat/catalog/?tc=c&idp=172
SERVEIS	
Servei WMS	http://geoserverdev.icgc.local:8080/geoserver/idescat/wms?
APLICACIONS	
Betaportal	http://betaportal.icgc.cat/wordpress/poblacio/

La distribució territorial de les variables demogràfiques es distribueix en un [fitxer .zip](#) anomenat *gridpoblacio01012016*, que conté una carpeta *rp2016* amb un arxiu .pdf amb la metodologia i els següent arxius en format Shapefile:

- rp2016_qtree_level2_ofus_allvar.dbf
- rp2016_qtree_level2_ofus_allvar.prj
- rp2016_qtree_level2_ofus_allvar.shp
- rp2016_qtree_level2_ofus_allvar.shx

9. Captació

La font d'informació és el Registre de població de Catalunya, amb data de referència l'1 de gener de 2016. Aquest registre conté, per a cada habitant, una sèrie de variables demogràfiques (edat, sexe, nacionalitat, etc.) i altres territorials (província, municipi, districte, secció censal i adreça postal).

La informació de cada habitant ha estat sotmesa a la geocodificació de l'adreça postal utilitzant el servei de geocodificació de l'Institut Cartogràfic i Geològic de Catalunya (ICGC).

El resultat del procés de geocodificació ha permès obtenir les coordenades de les adreces postals en un 93,2% dels casos, que apleguen un 95,3% del total de població. La resta d'adreces postals han estat imputades fent servir com a base de dades el cadastre urbà.

Els casos no geocodificats, que representen un 4,7% de la població, s'han imputat amb un conjunt de mètodes utilitzant com a informació els centroides de les finques urbanes del cadastre.

S'ha construït una taula d'equivalències entre carrers INE i carrers cadastre utilitzant dos criteris:

- a) Semblança entre els noms dels carrers utilitzant la mètrica Jaro-Winkler.
- b) La relació "més proper" entre els punts geocodificats i els centroides de les finques.

Aquesta taula d'equivalència ha permès assignar posicions en els casos en què el nom del carrer INE existia a la taula, cercant la finca amb la mateixa numeració (1,59% de la població) o la finca amb numeració més propera (1,32%).

Per a la resta de casos, i atès que a priori es coneixia la secció censal però no es disposava de la seva geometria, s'ha construït una aproximació, mitjançant el convex hull dels punts geocodificats, per a cada secció censal amb la intersecció de les zones urbanes definides en el SIGPAC.

Finalment, la imputació s'ha realitzat construint una taula que relacionava el nombre de béns immobles en cada finca amb el nombre de llars de cada portal geocodificat¹, a partir de la qual s'han obtingut uns pesos que s'han utilitzat en la imputació aleatòria a les finques dins les zones del convex..²

Amb l'objectiu de presentar les dades d'una manera gràfica fàcilment interpretable i al mateix temps preservar les dades personals, s'ha optat per fer servir una estructura quadtree per emmagatzemar la informació per a la seva distribució.

Agregació en quadtrees

Un quadtree és una estructura jeràrquica on els elements d'un nivell Grid poden tenir com a màxim quatre descendents. Aquesta estructura es pot obtenir de la divisió en dos dels elements del grid estàndard europeu a una determinada resolució, obtenint quatre elements a la resolució immediatament superior i així successivament. El procés de subdivisió està determinat per la població i així permet optimitzar la resolució en funció de la densitat de població, i oferir més resolució en zones on hi ha una densitat de població elevada i una resolució més baixa en zones on la densitat de població és menor.³

La capa de punts derivada de la geocodificació de les adreces ha estat agregada espacialment utilitzant un quadtree de resolució màxima (62,5 m) amb un llindar mínim de 17 habitants. En aquells casos en què la població és inferior al valor del llindar es produeix una agregació als nivells superiors (125 m o 250 m).

En la creació del quadtree s'han realitzat translacions entre elements comuns en la jerarquia a resolució 125 m en zones on hi ha valors elevats de la variància entre els elements, per tal d'evitar l'agregació inherent al mètode de creació dels quadtrees.

Amb aquestes translacions es minimitzen els errors deguts a les agregacions espacials en el còmput d'efectius de població

A resolució mínima (250 m) i en els casos en què la població és inferior al llindar (17 habitants) els valors associats són nuls.

Si el valor d'una categoria d'una variable és inferior a set habitants es procedeix a la supressió primària de les dades i, en cas necessari, a la supressió secundària d'altres categories, deixant els valors a nul.

10. Manteniment

Es tracta d'un conjunt d'informació que presenta una imatge fixa de la població de Catalunya en un moment determinat, 1 de gener de 2016. Per tant no es fa cap actualització de les dades.

Bianualment es genera un nou conjunt de dades amb les dades de població de Catalunya corresponents a aquell any.

11. Representació

No s'estableix una simbolització específica per a aquest conjunt d'informació, són els usuaris qui decideixen quina simbolització apliquen a les dades.

¹ [EEA referencia grid – European Environment Agency.](#)

² [Hacia un registro estadístico de territorio | SEIO2015.](#)

³ [Using quadtree representations in building stock visualization and analysis.](#)

Annex A. Referències normatives

- Llei 23/1998, de 30 de desembre, d'estadística de Catalunya.
- Decret 62/2010, de 18 de maig, pel qual s'aprova el Pla cartogràfic de Catalunya.
- Directiva 2007/2/EC del Parlament Europeu i del Consell de 14 de març de 2007 per la qual s'estableix una Infraestructura d'informació espacial a la Comunitat Europea (INSPIRE).
- ISO 19101:2002, Geographic information - Reference model
- ISO/TS 19103:2005, Geographic information - Conceptual schema language
- ISO 19107:2003, Geographic information - Spatial schema
- ISO 19108:2002, Geographic information - Temporal schema
- ISO 19108:2002/Cor 1:2006, Geographic information - Temporal schema, Technical corrigendum 1
- ISO 19109:2005, Geographic information - Rules for application schema
- ISO 19110:2005, Geographic information - Methodology for feature cataloguin
- ISO 19110:2005/Amd 1:2011, Geographic information - Methodology for feature cataloguing, Amendment 1
- ISO 19111:2007 Geographic information - Spatial referencing by coordinates
- ISO 19115:2003, Geographic information – Metadata
- ISO 19115:2003/Cor 1:2006, Geographic information – Metadata
- ISO 19131:2007, Geographic Information - Data product specifications
- ISO 19131:2007/Amd 1:2011, Geographic Information - Data product specifications, Amendment 1
- ISO 19157:2013, Geographic information - Data quality

Annex B. Termes i definicions

(1) Registre de Població

És un registre administratiu on figuren les dades actualitzades de caràcter obligatori dels veïns inscrits en els padrons municipals d'habitants de tots els ajuntaments de Catalunya. [Llei 23/1998, de 30 de desembre, d'estadística de Catalunya].

(2) convex hull

Superfície convexa que tanca un conjunt de punts.

Notes:

Una forma intuïtiva de veure el seu significat és imaginar una banda elàstica estirada que els tanca a tots. Quan s'allibera la banda elàstica, aquesta pren la forma de convex hull.

(3) Quadtree

Estructura jeràrquica corresponent a la divisió recursiva del espai en que cada element es divideix en quatre elements iguals.

Notes:

Si es parteix de un quadrat de 1Km de costat, en el primer nivell obtenim quatre de 500m, en el segon quatre de 250m i així successivament.

Anàlogament la construcció es pot fer en sentit invers, agregant, amb els mateixos resultats. Si el criteri de divisió o agregació s'associa a un cert llindar de un fenomen (població en aquest cas) obtenim una divisió del territori en el que la seva resolució depèn del llindar. En el cas que ens ocupa la resolució espacial serà més gran en zones a on hi ha més població.

(4) Variables demogràfiques

Característiques que descriuen a la població en estudi.

Notes:

Exemples de població en estudi són: per sexe, edat, lloc de naixement, etc.

(5) Conjunt d'informació geogràfica

Col·lecció de dades, en qualsevol estructura de dades espacials (a tall d'exemple, vectorial o *ràster*), que formen una unitat operativa i d'emmagatzematge i que representen una o més classes d'entitats geogràfiques, relacionades o simplement reunides per afinitat temàtica, per coincidència geogràfica o per conveniència. Segons l'estructura de dades i el format pot ser, entre d'altres, un fitxer, una part d'un fitxer o una col·lecció de fitxers. Sovint s'anomena també base cartogràfica digital [PCC].

(6) Element de qualitat de les dades

Component que descriu un cert aspecte de la qualitat de les dades geogràfiques [ISO 19157].

Notes:

Aquest terme es descriu a la secció 7.3.1 de ISO 19157.

L'aplicabilitat d'un element de qualitat a un conjunt de dades depèn tant del contingut del conjunt com de les seves especificacions de producte, és per aquest motiu que no tots els elements de qualitat poden ser aplicables a tots els conjunts de dades.

(7) Esquema d'aplicació

Esquema conceptual de les dades requerides per una o més aplicacions [ISO 19101].

(8) Mesura bàsica de la qualitat de les dades

Mesura genèrica de qualitat de les dades utilitzada com a base per a la creació de mesures específiques de qualitat de les dades [ISO 19157]

Notes:

El concepte de mesura de la qualitat de les dades equival al de l'avaluació d'un element de qualitat de les dades, segons ISO 19157.

(9) Model de dades

Conjunt d'estructures i regles per mitjà de les quals s'organitzen i operen en un sistema d'informació les dades corresponents a la informació que constitueix la representació d'un univers determinat (per exemple, el model de dades relacional o, en el context del Sistema d'Informació Geogràfica, el model de dades ràster). Proporciona les estructures a partir de les quals es construeix el programari i, alhora, constitueix el patró de disseny de bases de dades, alfanumèriques o espacials, per organitzar la informació [PCC].

(10) Objecte geogràfic

Terme usat per a descriure l'abstracció d'elements del món real [ISO 19101]. És la unitat fonamental d'informació geogràfica [ISO 19109].

Notes:

El tipus d'objecte geogràfic es representa amb una classe al diagrama UML.

El terme "*objecte geogràfic*" substitueix a "*fenomen*" com a traducció oficial i normalitzada del terme anglès "*feature*" des del 2 de desembre de 2013, segons acord del Comitè Tècnic de Normalització 148 d'AENOR a la seva reunió plenària número 43.

Annex C. Glossari de sigles i abreviatures

BT-5M	Base Topogràfica 1:5 000 de Catalunya
EPSG	European Petroleum Survey Group (Grup europeu d'estudi del Petrolí)
ETRS89	European Terrestrial Reference System 1989 (Sistema de referència terrestre europeu 1989)
ETRS-UTM31	Projecció ETRS89 Universal Transversa de Mercator Fus 31
ID	Identificador
ICGC	Institut Cartogràfic i Geològic de Catalunya
IDEC	Infraestructura de Dades Espacials de Catalunya
IDESCAT	Institut d'Estadística de Catalunya
INE	Instituto Nacional de Estadística
INSPIRE	Infrastructure for Spatial Information in Europe (Infraestructura de la informació espacial a Europa)
ISO	International Organization for Standardization (Organització internacional per a l'estandardització)
PCC	Pla Cartogràfic de Catalunya
SIGPAC	Sistema d'Informació Geogràfica de Parcel·les Agrícoles
UML	Unified Modelling Language (Llenguatge unificat de modelatge)
XML	Extensible Markup Language (Llenguatge de marques extensible)