

Projecte Dades Obertes de Catalunya

<http://dadesobertes.gencat.cat>

Dades obertes gencat

Projecte d'obertura de dades públiques
(open data) de la Generalitat de Catalunya



Què és Dades Obertes? (*Open data*)

L'*open data* és una filosofia i una pràctica que requereix que certes dades siguin de lliure accés per a tothom, reduint les limitacions tècniques, les restriccions legals o altres mecanismes de control. Es considera que, de la mateixa manera que obrir el codi font del programari permet que altres el millorin en benefici de la comunitat, **les dades obertes beneficien al conjunt de la ciutadania**.

En el sector públic, tenir accés a les dades de l'administració garanteix la transparència, l'eficiència i la igualtat d'oportunitats a l'hora de crear valor. La **transparència** perquè es poden consultar i tractar dades que provenen directament de les fonts oficials. L'**eficiència** perquè ciutadans i organitzacions poden crear serveis de forma més ajustada en col·laboració amb l'administració; i la **igualtat d'oportunitats** perquè l'accés és el mateix per a tothom a l'hora de crear aplicacions o fer negocis. Per totes aquestes raons, molts [governos arreu del món](#) estan impulsant projectes similars al que Catalunya presenta ara amb el portal <http://dadesobertes.gencat.cat>

El World Wide Consortium (W3C), dirigit pel britànic Tim Berners Lee, és un dels principals impulsors del moviment internacional a favor de l'obertura de dades públiques. El [grup d'eGovernment del W3C](#), liderat per [Jose Manuel Alonso](#) de la Fundació CTIC a Astúries, ha assessorat a la Generalitat en el projecte.

Les [licències i els termes d'ús](#) de les dades obertes estan sotmeses a les lleis de reutilització de la informació del sector públic, i en alguns casos poden tenir llicències de propietat intel·lectual encara que es tendeix a obrir-les sense condicions, sempre que es mantinguin sense manipular i amb l'obligatòria citació de la font i de la seva darrera actualització.

Tant la Unió Europea com el govern espanyol consideren que les dades públiques han de ser reutilitzades per la ciutadania i per les empreses, ja que això suposa un enfortiment democràtic i un motor per al desenvolupament de la societat de la informació i el coneixement, sobretot per al sector dels continguts digitals. Per aquest motiu, han legislat a favor amb la [Llei 37/2007, de 16 de novembre, sobre reutilització de la informació del sector públic](#), que transposa la [Directiva 2003/98/CE del Parlament Europeu i del Consell](#).

Més informació sobre les dades públiques com a motor del desenvolupament de la societat de la informació: Guia del [Proyecto Aporta](#) del Ministeri de Ciència i Tecnologia (Plan Avanza).

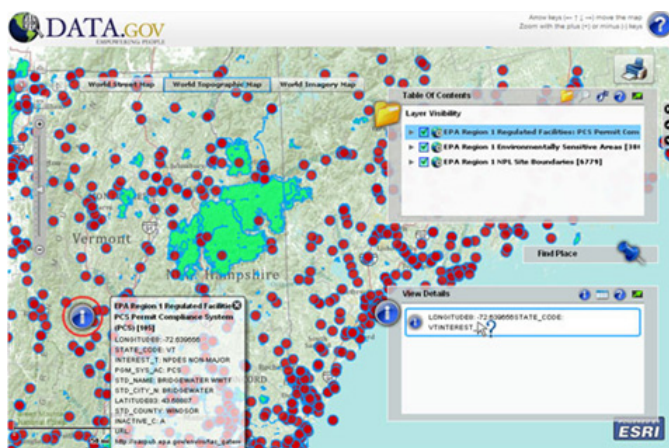
Què es pot fer amb les dades obertes?

Les dades obertes permeten la consulta i la construcció d'aplicacions, en especial programari i formes de visualització, que usen la informació alliberada com a font, en molts casos de forma automatitzada. D'aplicacions en pot fer la mateixa administració o altres (particulars, organitzacions i empreses) en el moment que se n'allibera l'accés en un portal públic, com el de gencat, i se'n fa difusió.

Les aplicacions que es poden construir amb dades alliberades són molt diverses, ja que responen a objectius i finalitats molt diferents. Van des de les dels ciutadans que incorporen al seu mòbil -per exemple, els avisos per RSS de les incidències del transport o les previsions meteorològiques-, fins als d'associacions de malalts respiratoris que controlen els índexs de contaminació o actualitzacions automàtiques amb la informació d'equipaments oficials que poden alimentar, entre d'altres, els navegadors GPS per a vehicles.

Altres exemples de la riquesa i la varietat de les aplicacions que es poden crear a partir de dades són mapes i gràfics interactius que s'utilitzen en l'anomenat [periodisme de bases de dades](#) i en nombrosos serveis de cerca, com és el cas del [cercador d'equipaments de gencat](#) que permet localitzar escoles, hospitals o centres esportius, entre d'altres. També formen part de les aplicacions generades amb dades els titulars incorporats via RSS a blogs, webs o aplicacions mòbils (com les [pròpies de gencat](#)).

Algunes eines de visualització de dades són:



- Gràfics interactius publicats a mitjans digitals com els de [l'informació.com](#) o el [New York Times](#).
- El visor desenvolupat pel govern nord-americà al web [data.gov](#) de dades obertes.

Dades Obertes gencat

Amb la llibertat d'accés a la informació pública, hi ha la **voluntat del Govern d'avançar cap a la governança oberta**.

La Generalitat de Catalunya ja fa temps que treballa en la línia de garantir la reutilització de la informació pública i l'obertura de dades. En són exemples:

- La publicació d'un [avís legal](#) que contempla la reutilització de la informació publicada al web gencat.
- L'emissió de [nombrosos RSS](#) d'informació i serveis, que ja són utilitzats per altres serveis digitals públics i privats.
- La possibilitat de descàrrega de bases de dades, com les del [Termcat](#) de terminologia oberta, o l'alliberament d'aplicacions (anomenades [API](#)) sobre les dades, com les que ofereix [l'Institut d'Estadística de Catalunya](#).

Amb el portal dadesobertes.gencat.cat es fa un pas més en aquesta direcció, agrupant i ampliant totes les iniciatives de dades obertes de la Generalitat en un únic catàleg on s'afegeix la informació associada més rellevant perquè sigui reutilitzada.



The screenshot shows the homepage of the 'Dades obertes gencat' portal. At the top, there's a navigation bar with links like 'Webs A-Z', 'Mapa web', 'Contacte', and 'Castellano'. Below this is a header with the 'Dades obertes gencat' logo and a list of data formats: TMX, ZIP, PDF, CSV, KML-KMZ, DOC, XLS, XML, JSON, RDF/XML, and SHP. The main content area is titled 'Catàleg de dades obertes' and lists several categories: 'Equipaments de Catalunya' (Museus, escoles, hospitals, instal·lacions esportives... Més de 26.000 registres oberts), 'Tràmits de la Generalitat de Catalunya per a la ciutadania' (Ajuts, subvencions, permisos i llicències, concursos i licitacions... Un total de més de 1.000 tràmits), and 'Predicció meteorològica a Catalunya' (Informació actualitzada tres cops al dia sobre l'estat del cel, la precipitació, la temperatura, la visibilitat, el vent i l'estat de la mar). On the right side, there's a sidebar with links to various services like 'Tràmits Ajut Estadístiques Ind', 'Mapes Equipam', 'Ortofotografies Ima', 'Multimèdia Da', 'lingüístiques Term', 'Meteorologia Pre', 'Trànsit Transport', and 'Actualitat'. At the bottom, there's a section for 'Aplicacions fete' with links to 'Cercador d'equip de Catalunya' and 'Oficina Virtual de'.

Per primer cop s'obren, a més, un conjunt significatiu de dades ben estructurades en termes de la seva reutilització. La successiva publicació de dades, segons l'acord de Govern aprovat avui, es farà "prioritzant els interessos de servei públic, garantint la màxima qualitat en el contingut i el format, fent que siguin al màxim d'accessibles i intel·ligibles per a la ciutadania i usant llicències de difusió oberta quan escaigui".

Dades disponibles al portal gencat (primera fase)

Tràmits i procediments administratius

Tràmits de la Generalitat de Catalunya sobre ajuts, subvencions, permisos i llicències, concursos i licitacions...

1. Tràmits per a la ciutadania
2. Tràmits per a les empreses

** En format estructurat i reutilitzable (RDF-XML) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Dades estadístiques

3. Indicadors al dia de l'economia catalana
4. Indicadors econòmics de Catalunya
5. Estadístiques bàsiques dels 947 municipis i 41 comarques de Catalunya.

** En format estructurat i reutilitzable (API-RSS) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Dades geogràfiques i cartogràfiques

6. Equipaments de Catalunya: museus, escoles, hospitals, instal·lacions esportives (més de 26.000 equipaments d'interès ciutadà).

** En format estructurat i reutilitzable (RDF-XML) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

7. Cartografia de Catalunya: bases topogràfiques, mapes, ortofotografies

** Còpia permesa amb finalitat d'estudi o recerca, citant la font Institut Cartogràfic de Catalunya. Per a qualsevol altre ús cal demanar autorització.*

8. Serveis de posicionament geodèsic: Caster NTRIP, CATNET WEB, GEOFONS...

** En diversos formats sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Dades visuals

9. Àlbums fotogràfics de les restauracions
10. Planimetria dels monuments catalans

** En formats reutilitzables (API) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització*

11. Cartoteca digital: fotografies històriques

** Còpia permesa amb finalitat d'estudi o recerca, citant la font Institut Cartogràfic de Catalunya. Per a qualsevol altre ús cal demanar autorització.*

Dades lingüístiques

12. Terminologia Oberta: repertoris terminològics d'interès general, procedents de treballs de recerca terminològica duts a terme pel TERMCAT.

** En formats reutilitzables (XML) amb llicències Creative Commons.*

Dades meteorològiques

13. Prediccions de Catalunya i dels Països Catalans, comarcal, marítimes i per al Pirineu.

** En formats reutilitzables (XML) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Dades d'actualitat

14. RSS de titulars de notícies i altres continguts d'actualització freqüent.

** En formats reutilitzables (RSS-XML) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Trànsit i mobilitat

15. Estat del servei de Rodalies

** En formats reutilitzables (RSS-XML) sense condicions especials excepte la no desnaturalització, només citant la font i la data d'actualització.*

Catàleg de dades obertes gencat

Tipus de dades obertes



Tràmits i procediments administratius

Ajuts i subvencions, permisos i llicències, concursos, pagaments...



Dades geogràfiques i cartogràfiques

Mapes topogràfics, ortofotos, instal·lacions, museus, escoles, hospitals...



Dades estadístiques

Registres demogràfics, econòmics, socials, per divisions territorials...



Dades lingüístiques

Terminologia d'interès general, nous vocables, diccionaris generals i específics...



Dades visuals

Fotografies i imatges, originals sense manipular o postproduïts, productes gràfics...



Dades meteorològiques

Mapes isobàrics, imatges de satèl·lit, radars, precipitacions, pluges, temperatura, llamps...



Trànsit i mobilitat

Càmeres de vigilància, carreteres, retencions, obres, incidències, mapes...



Dades d'actualitat

Darreres notícies, novetats i avisos de temes com salut, turisme, treball, família, tecnologia...

Projectes en curs (propera fase)

Adreces i telèfons de la Generalitat de Catalunya

Informació bàsica i de contacte dels organismes que componen la Generalitat de Catalunya i dels seus responsables.

Exemple d'aplicació: [Qui és qui de la Generalitat](#)

Informació sobre l'estat del transport públic:

Metro, autobusos, tramvia, Ferrocarrils de la Generalitat, etc.

Exemple d'aplicació: [Giny incidències en el transport públic](#)

Informació meteorològica

Temperatura, humitat relativa, precipitació diària acumulada i la velocitat i direcció del vent expressades en mapes.

Exemple d'aplicació: [Dades d'estacions automàtiques](#) (XEMA)

Dades en temps real de l'aigua disponible a Catalunya amb informació recollida de les estacions automàtiques, dades de pluges, mapa d'avisos i control del reg.

Exemple d'aplicació: [L'aigua en temps real](#)

Radiacions electromagnètiques

Mapa amb la ubicació de les estacions que mesuren les emissions electromagnètiques de les antenes de telefonia mòbil i consulta de resultats en línia.

Exemple d'aplicació: [Mapa de les estacions](#)

Cobertura serveis TIC

Cercador de cobertures de TDT, banda ampla i telefonia mòbil en funció de la comarca i el municipi. També ofereix resultats per població o polígon industrial.

Exemple d'aplicació: [Cercador de cobertures](#)

Itineraris en transport públic

Consulta d'itineraris des d'una adreça, punt d'interès, parada o estació, amb la possibilitat de seleccionar el mitjà de transport, els minuts que es vol caminar i el tipus de ruta (ràpida, amb transbordaments, a peu, etc.)

Exemple d'aplicació: [Mou-te amb transport públic per Catalunya](#)

Imatges

Selecció d'imatges i vídeos del banc iconogràfic de la Generalitat de Catalunya (BIG), que il·lustren els paisatges i el patrimoni de Catalunya.

Monuments arqueològics i arquitectònics de Catalunya

Exemple d'aplicació: [CliCat, galeria multimèdia de Catalunya](#)

Informació dels observatoris sectorials d'Agricultura, Alimentació i Acció Rural

Informació estadística dels observatoris sectorials del Departament d'Agricultura, Alimentació i Acció Rural: observatoris de la llet, de la vinya, el vi i el cava, del porcí, de la fruita, del sector cunícola, de l'agricultura ecològica, i del d'agroalimentari de preus.

Exemple d'aplicació: [Pàgina d'observatoris sectorials del Departament](#)

Capes gràfiques i cartogràfiques del SIG (Sistema d'Informació Gràfica)

Informació de capes gràfiques SIG del Departament d'Agricultura, Alimentació i Acció Rural: parcel·les agrícoles, capa de vinya (varietats, superfícies, anys de plantació...), capa fructícola (producte), explotacions ramaderes, mapa de sols, cartografia marina...

Exemple d'aplicació: [Pàgina de Cartografia i SIG del Departament](#)

Guia d'allotjaments turístics de Catalunya

Llistat d'allotjaments turístics del Registre de Turisme elaborat pel Departament d'Innovació, Universitats i Empresa: càmpings, hotels i establiments de turisme rural.

Exemple d'aplicació: [Cercador d'allotjaments turístics](#)

Observatori d'Emergències i Seguretat

El Departament d'Interior, Relacions Institucionals i Participació, a través del Programa d'Anàlisi i Seguretat de la Informació, anirà obrint de forma progressiva dades i/o informació estadística –sota la premissa de l'acompliment normatiu en matèria de seguretat de la informació i de la protecció de dades sensibles– relacionades amb les volumetries i tipologia de trucades que arriben al 112, les operacions de protecció civil, les actuacions dels bombers i de la policia, agregades estadísticament.

Com s'han de publicar les dades?

Segons diu el W3C en la seva [darrera guia de publicació](#), s'han d'alliberar dades útils i en formats reutilitzables. Qualsevol [format](#) és benvingut, però quan més ben estructurades i enriquides estiguin les dades més fàcil serà reutilitzar-les i construir aplicacions que les tractin automàticament.

Per això hi ha formats i projectes de dades obertes que es consideren de menys qualitat que d'altres. Segons la [classificació proposada per Tim Berners Lee](#), els formats més adequats són l'[RDF](#), susceptible de passar fàcilment al següent nivell anomenat "dades enllaçades" ([linked data](#)), que al seu torn és la base tècnica per alimentar l'anomenada [web semàntica](#), un estàndard on cada dada conté informació associada que la relaciona automàticament amb d'altres.

L'W3C recomana crear un catàleg amb la descripció acurada de tots i cadascun dels conjunts de dades d'una institució (*data sets*) i, alhora, alliberar també el propi catàleg en format RDF. La Generalitat així ho ha fet en aquest [portal](#), des d'on no només es pot consultar sinó també descarregar, el catàleg amb totes les iniciatives de dades obertes dins de la institució. Es fa, a més, un salt qualitatiu ja que s'obren, per primer cop, un conjunt significatiu de dades en els formats que es consideren d'alta qualitat ([RDF](#)). Aquest primer conjunt de dades alliberades en la categoria de quatre estrelles està compostat per la base de dades dels 26.000 equipaments oficials de Catalunya, els 1.400 tràmits de la Generalitat, una part dels arxius de patrimoni cultural i dades estadístiques. S'està treballant, a més, per assajar prototipus en el nivell de dades enllaçades o *linked data*, considerat cinc estrelles, ja que, en molts casos, les bases de dades corporatives estan prou estructurades com per poder ser relacionades amb d'altres fàcilment.

Nivells de Qualitat en l'alliberament de les dades, segons el W3C

Publicat ★

Alguna dada ha estat publicada en un format no estructurat, és a dir, que es requereixen tècniques per tractar el document. Poden ser:

- Formats "caixa negra": imatges, vídeo, música...
- Formats binaris: pdf, ps...

Dades estructurades ★★

Dades publicades en formats estructurats, per exemple, fulls de càlcul Excel.

Problema: Els formats propietaris requereixen eines que no són públiques.

Formats oberts ★★★

Dades publicades en formats estructurats, oberts y no propietaris, per exemple, CSV o XML.

Problema: El contingut depèn del context.

URI per identificar una dada ★★★★★

Hi ha diferents representacions per a cada tipus de dada (XML)

- Negociació del contingut: El protocol de comunicació permet que el navegador informi al servidor del tipus de contingut que prefereix i el servidor torna una representació diferent segons les preferències del client.

El format XML permet representar informació estructurada, els documents es poden validar i el significat de les etiquetes depèn de l'aplicació.

Enllaçar amb altres dades ★★★★★

Enllaçar dades significatives mitjançant URIs per facilitar la identificació de nous conjunts de dades.

Format RDF:

- Model de representació en forma de graf.
- Les propietats tenen valors globals.
- Dos grafs independents poden compondre's.
- El llenguatge SPARQL permet realitzar consultes a grafs.
- Negociació de contingut HTML només dona informació per navegadors web i RDF dona informació per a sistemes automàtics.