

Portal de la Recerca de Catalunya: estat a abril de 2022 i pla de treball 2022-23

(Doc. CO22/13; Ciència Oberta\5 PRC\2204PlaMillloresPRC22_23.docx; Comissió PRC 26.04.22)

1. Antecedents i funcionament

La DGR va delegar en el CSUC la creació del Portal de la Recerca de Catalunya (PRC, <https://portalrecerca.csuc.cat/>). Aquest té per finalitats visualitzar i difondre des d'un lloc únic l'activitat investigadora que es duu a terme a Catalunya en quant a investigadors/es, publicacions, tesis, grups i projectes de recerca i fomentar l'accés obert a la producció científica. El PRC s'adreça a la comunitat científica internacional, empreses, organitzacions finançadores i públic en general¹.

Els treballs per posar-lo en funcionament van completar-se el 2015² i es va obrir al públic en dos períodes, un de consulta restringida (setembre del 2015), i ja de forma pública (setembre del 2016). El Portal va ser inaugurat oficialment el 23 de gener del 2017 en un acte que va tenir lloc al Recinte Modernista de Sant Pau i que ser presidit pel Sr. Jordi Baiget, Conseller d'Empresa i Coneixement³.

La creació del PRC es va concebre en dues fases, una primera com a Portal de la Recerca de les Universitats de Catalunya ja que llavors no hi havia encara dades de cap centre de recerca, i una segona com a Portal de la Recerca de Catalunya, iniciada a partir de la integració del primer centre de recerca (el VHIR).

Els principis de funcionament del PRC, resumidament, són basar-se en la informació proporcionada directament per les universitats i centres de recerca a través dels seus sistemes de gestió de la recerca (CRIS) i seguir bones pràctiques de projectes equivalents a nivell local i internacional⁴.

El Portal es crea a partir de càrregues periòdiques. Al llarg dels anys de funcionament, el PRC ha tingut tres línies de millora⁴:

- **Creixement**, és a dir l'ampliació de continguts incloent més informació.
- **Ampliació**, amb la inclusió de noves institucions.
- **Extensió** o ampliació de serveis, per exemple, monitoritzant l'accés obert.

¹ Doc. 14/34: "Portal de la Recerca de Catalunya: antecedents, objectius, metodologia i estat a maig de 2014".

² Doc. 15/25: "Portal de la Recerca de Catalunya: situació prevista al juliol de 2015".

³ Vídeo del Portal de la Recerca de les Universitats de Catalunya en l'acte de presentació oficial <https://www.youtube.com/watch?v=c8NQAlVliYw&feature=youtu.be>.

⁴ Doc. 15/40: "Portal de la Recerca de Catalunya: accions de futur".

2. La incorporació d'informació de recerca dels centres CERCA

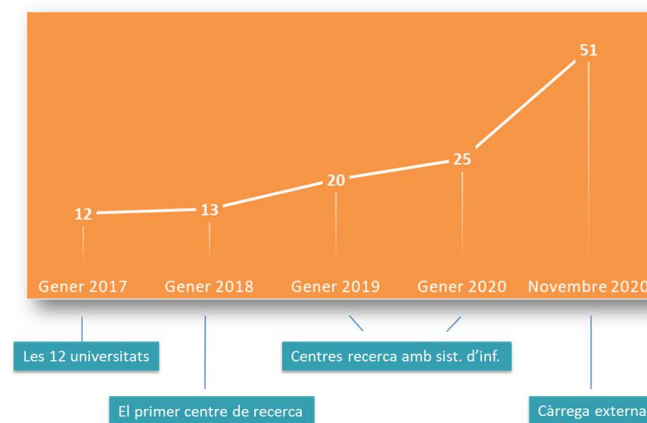
La incorporació de dades de les Universitats al PRC es va poder fer amb una relativa celeritat degut a que totes disposaven de sistemes de gestió de recerca (CRIS) consolidats. El mateix no va ser tan immediat per als centres de recerca CERCA ja que, quan es va iniciar, la majoria de centres no tractaven la informació de la seva producció de recerca amb sistemes CRIS sinó amb sistemes sense capacitat per exportar en CERIF.

De cara a fer que el Portal pogués presentar les dades de tota la recerca feta a Catalunya es va fer una acció especial consistent en carregar la producció científica dels centres de recerca de fonts externes (concretament de les bases de dades WoS i Scopus), de manera que els centres només haguessin de validar la informació recollida. Això va tenir lloc entre 2019 i 2020 i va consistir en diferents fases⁵.

Paral·lelament, al llarg d'aquest període hi ha hagut dues coses que han millorat la incorporació de dades dels centres CERCA al PRC. Per una banda, molts d'aquests han passat a usar sistemes CRIS amb capacitat d'exportar dades al Portal⁶. Per una altra s'ha habilitat al PRC una funcionalitat que permet incorporar informació de la producció científica al PRC als centres sense sistemes d'informació amb aquesta capacitat.

Aquestes actuacions permeten una alimentació normalitzada de totes les universitats i els centres CERCA i que el PRC pugui avançar cap a incorporar nova informació, sigui de les fonts actuals o d'externes, o de noves institucions.

A continuació es pot veure l'evolució d'institucions participants al PRC des de que es va inaugurar oficialment al gener del 2017 fins a la integració del darrer centre de recerca. De gener a novembre de 2020 es quan es va fer la càrrega d'informació dels centres de recerca a partir de fonts externes.



⁵ Doc. CO20/14 "Portal de la Recerca de Catalunya: estat a novembre de 2020 i pla de treball per al curs 2020-21"

⁶ Vegeu l'Annex 3

3. El PRC a abril de 2022

3.1 Continguts

A nivell de continguts, a abril de 2022, el PRC permet la consulta conjunta de la següent informació:

- 51 institucions
 - 12 universitats (UB, UAB, UPC, UPF, UdG, UdL, URV, UOC, URL, UVic-UCC, UIC i UAO)
 - 39 centres de recerca CERCA (Agrotecnio, CED, CIMNE, CRAG, CREAM, CREI, CRG, CRM, CTFC, CTTC, CVC, i2CAT, IBEC, ICAC, ICFO, ICIQ, ICN2, ICP, ICRA, ICRPC, IDIBAPS, IDIBELL, IDIBGI, IEEC, IFAE, IGTP, IIBSantPau, IISPV, IJC, IMIM, IPHES, IRB Barcelona, IRB Lleida, IREC, IrsiCaixa, IRTA, ISGlobal, VHIO i VHIR)
- 20.257 investigadors/es
- 675.912 publicacions i 38.382 tesis, el 33,1% en accés obert
- 41.754 projectes (37.367 tancats i 4.387 oberts) i 2.036 grups de recerca

A continuació es mostra el total de continguts per universitats i centres de recerca CERCA, el percentatge de la seva aportació respecte el total i el percentatge de coincidència.

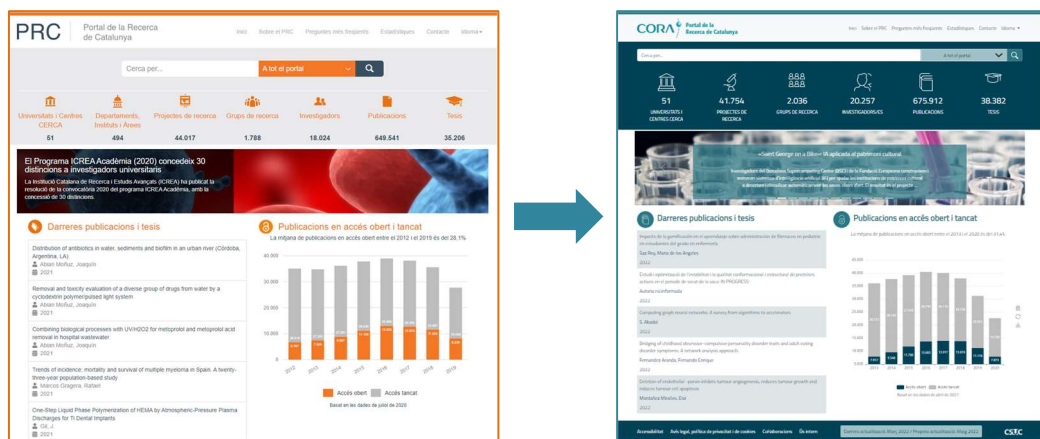
	Investigadors	Publicacions	Tesis	Projectes recerca	Grups recerca
Total universitats	11.925	543.436	33.067	34.371	1.276
Total centres recerca CERCA	8.332	132.624	5.315	7.383	760
% universitats	54%	77%	85%	81%	62%
% centres recerca CERCA	36%	17%	13%	17%	36%
% coincidència	11%	7%	2%	2%	3%

3.2 Principals accions i millores fetes el 2021

L'any 2021 s'han seguit fent les tasques de càrregues periòdiques i de manteniment del PRC. Així com actuacions de qualitat per millorar les dades a l'origen. Addicionalment hi ha hagut tres actuacions destacables referides a la renovació de la imatge i incorporació a CORA, a la publicació de les dades al catàleg de dades obertes de la Generalitat de Catalunya i a l'actualització de la informació dels centres CERCA.

Renovació de la imatge i incorporació a CORA

- S'ha renovat la imatge del PRC incorporant-lo a CORA (Catalan Open Research Area), denominació que engloba els serveis de ciència oberta que proporciona el CSUC.



Publicació de les dades al catàleg de dades obertes de la Generalitat de Catalunya

- S'han afegit cinc conjunts de dades: Investigadors, Publicacions, Projectes de recerca, Grups de recerca i Departaments.
- Les dades són acumulatives i s'actualitzen el dia 30 després de cada càrrega.
- Es poden exportar en diferents formats (csv, xml, etc.) i també es permet l'accés programàtic a través d'una API.

Actualització de la informació dels centres CERCA

- S'ha desenvolupat una funcionalitat a través del mateix programari que usa el PRC (DSpace-CRIS) per a que els centres CERCA que no disposen d'un sistema de gestió de la recerca puguin actualitzar la informació al PRC usant el format estàndard CERIF.

3.3 Possibles millores al PRC gràcies al fons REACT-EU

A finals del 2021 es va concedir al CSUC un ajut del fons REACT-EU per crear “Serveis per incrementar la competitivitat del sistema de coneixement de Catalunya”. Aquest serà un projecte per dur a terme en el període 2021-2023 i “per impulsar la competitivitat del sistema de coneixement, basat en serveis de dades (d’emmagatzematge, de preservació.... de forma molt lligada al Portal de la Recerca i al Repositori de Dades per a la Recerca) i d’administració digital.”

Pel que fa a serveis relacionats amb el Portal de la Recerca de Catalunya, es pretén posar en marxa dues noves millores significatives, una destinada a millorar la interoperabilitat del PRC amb altres fonts de dades i, l’altra, adreçada a la creació de noves funcions.

Interoperabilitat amb noves plataformes i fonts de dades

Actualment el PRC només pot captar informació dels CRIS de les universitats i centres de recerca i només la que es va acordar en el moment inicial de la creació del Portal. Aquesta restricció impedeix que el PRC pugui incorporar informació d’altres plataformes (com Dialnet), de bases de dades externes (com WoS, SCOPUS o UNEIX), o de serveis oberts (com ORCID).

La incorporació d’aquesta informació addicional enriquiria la informació proporcionada i, sobretot, permetria la creació de nous mecanismes de cerca i visualització de la informació continguda al PRC. Les limitacions actuals del PRC tampoc permeten avui importar les metadades procedents del CORA.Repositori de Dades de Recerca, cosa que és una greu limitació tant per al PRC com per al CORA.RDR.

Així doncs l’objectiu d’aquesta actuació seria analitzar l’arquitectura informàtica actual del PRC per tal de millorar la interoperabilitat actual del PRC amb els CRIS actuals, amb repositoris nous com el CORA.RDR i amb bases de dades, plataformes i serveis externs.

Desenvolupament de noves funcionalitats

Es vol reforçar el portal amb noves funcionalitats per fer que la recerca produïda a Catalunya sigui més oberta, maximitzar-ne l’impacte, i fer-la més competitiva incrementant la seva visibilitat. Això es faria incorporant indicadors bibliomètrics i elaborant un mapa d’experts que faciliti la transferència de coneixement de les universitats i els centres de recerca a les empreses (vegeu els annexos 1 i 2).

4. Pla de treball per al 2022-23

Amb els treballs fets per incorporar la informació de recerca dels centres CERCA, el Portal de la Recerca de Catalunya ha assolit l'objectiu de recollir la informació dels resultats de la recerca d'una bona part del sistema de recerca de Catalunya. Aquesta fita és important i fa que també ens puguem plantejar nous objectius alhora que es continua mantenint i millorant el funcionament del PRC.

Alguns d'aquests objectius ja s'havien formulat en el Pla de treball anterior però no s'han pogut realitzar en la seva totalitat. Els objectius que es plantegen per al 2022-2023 s'agrupen sota aquests apartats:

- Actualització i manteniment de la informació
- Augment de la quantitat i qualitat de la informació
- Explorar noves millores i serveis Millora de la visualització i de la cerca

Actualització i manteniment de la informació

- Continuar fent les tasques de manteniment del PRC que consisteixen bàsicament en resoldre les incidències i preguntes que genera el funcionament del PRC i preparar i executar les càrregues de dades que es fan cada dos mesos.
- Resoldre l'actualització de la informació de totes aquelles institucions participants que no l'aporten de forma automàtica a través dels seus sistemes d'informació o que no ho fan amb l'estàndard CERIF (identificats a l'Annex 3 amb **).
- Estendre l'ús de les dades en obert i concretament el procediment usat amb l'AQU d'aprofitar la informació del PRC per formalitzar processos administratius.

Augment de la quantitat i qualitat de la informació

- Continuar fent actuacions de qualitat per millorar les dades als sistemes de gestió de la recerca (CRIS) o similars de les institucions.
- Incloure més informació al PRC, concretament la que actualment ja està a alguns sistemes d'informació però no a tots, informació completa des d'UNEIX dels projectes de recerca (entitat financadora, finançament, programa, convocatòria i ID del projecte), informació dels conjunts de dades de recerca publicats al CORA.RDR i d'altres.
- Estendre l'ús dels DOIs al PRC (més accés obert, indicadors bibliomètrics, mapa d'experts, menys duplicats...).
- Ampliar, dins el que sigui possible, les institucions participants al PRC.

Millora de la visualització i de la cerca

- Revisar l'arquitectura actual del PRC de cara a actualitzar-la, si cal, i a facilitar la interoperabilitat del Portal amb altres fonts de dades.
- Explorar i desenvolupar noves funcionalitats, concretament
 - Incorporar indicadors bibliomètrics
 - Elaborar un mapa d'experts que faciliti la transferència de coneixement de les universitats i els centres de recerca a les empreses.

- En funció de la informació del portal i de les millores que es puguin introduir, presentar els resultats de forma visual i de gràfiques i millorar i ampliar el sistema de les cerques.
- Mantenir i millorar l'indicador de l'accés obert donant dades tant de forma global com desagregada per institució i tipologies de documents i a nivell d'investigador, més transparència i distingir la tipologia d'accés obert.

Annex 1

Creació d'un indicador d'accés obert a la producció científica de Catalunya

El Dr. Ángel Borrego, a proposta de l'Àrea de Ciència Oberta del CSUC, va elaborar l'informe: "Creació d'un indicador d'accés obert a la producció científica de Catalunya". L'informe, que pot ser consultat a RECERCAT⁷, revisa la bibliografia publicada en l'última dècada sobre l'evolució de l'accés obert a la bibliografia científica, analitza des d'un punt de vista metodològic les característiques de diversos indicadors nacionals d'accés obert i posa de manifest la necessitat de prendre un seguit de decisions metodològiques sobre la col·lecció documental sobre la qual calcular l'indicador, l'eina utilitzada per determinar la disponibilitat en accés obert, la definició d'accés obert emprada i el procediment de recollida de dades.

Les recomanacions finals de l'informe es copien a continuació.

"La revisió de les iniciatives internacionals de monitorització de l'accés obert mostra que la decisió d'establir un indicador d'accés obert a la producció científica de Catalunya és oportuna i útil. Es tracta d'una iniciativa viable, que no requereix de grans recursos, i que està en sintonia amb el que s'està fent a la Unió Europea per mesurar la concreció de les polítiques de ciència oberta.

L'anàlisi de diversos models ha posat de manifest la necessitat de prendre un seguit de decisions que s'haurien de consensuar entre les institucions responsables de la producció bibliogràfica i els agents que utilitzaran l'indicador.

Decisions metodològiques:

- El Portal de la Recerca de Catalunya seria la font més adient per identificar la producció científica de Catalunya, tot i que caldria garantir la compleció de les dades per part de les institucions participants.
- Unpaywall sembla la millor eina per determinar la disponibilitat en accés obert d'aquesta producció. En ser l'instrument emprat a la major part d'iniciatives internacionals, facilitaria la comparació de les dades.
- Per bé que alguns indicadors es completen amb dades d'altres fonts per determinar, per exemple, el cost del trànsit a l'accés obert, en aquesta primera etapa segurament seria més adient limitar l'abast de la iniciativa al mesurament de la disponibilitat de les publicacions en accés obert.
- L'indicador podria publicar-se amb freqüència anual, amb un calendari públic que facilités la revisió de les dades per part de les institucions participants i tingués en compte els possibles períodes d'embargament.
- Seria convenient publicar la informació metodològica i les dades recollides que permetessin reproduir el càlcul de l'indicador."

⁷ <http://hdl.handle.net/2072/520604>

Annex 2

Indicadors bibliomètrics al Portal de la Recerca de Catalunya

El Dr. Llorenç Arguimbau, a proposta de l'Àrea de Ciència Oberta del CSUC, va elaborar: "Informe sobre els indicadors bibliomètrics al Portal de la Recerca de Catalunya (PRC)". L'informe, que pot ser consultat a RECERCAT⁸, fa una introducció general als indicadors bibliomètrics, revisa els indicadors bibliomètrics que ofereixen els portals científics de diversos països europeus, els d'algunes universitats catalanes, i el PRC, i examina alguns aspectes de la informació que dona el Portal de cara a explorar les potencialitats bibliomètriques del PRC.

L'apartat de conclusions de l'informe diu:

"La revisió d'altres portals de recerca europeus i catalans (en estats diversos d'informació, evolució tecnològica i serveis) demostra la viabilitat, l'interès i les potencialitats dels indicadors bibliomètrics (actuals i futurs) del PRC en el marc del sistema d'R+D+I català.

La recopilació periòdica d'informació a partir dels sistemes de les universitats i dels centres CERCA garanteix una cobertura més exhaustiva que les bases de dades internacionals de referència, però també comporta una qualitat desigual de la informació i, de retruc, una dificultat a l'hora de construir indicadors bibliomètrics sòlids.

A causa de l'heterogeneïtat dels sistemes d'origen, el portal s'ha anat desenvolupant a partir d'acords consensuats amb les institucions participants per garantir l'homogeneïtat de la informació. En aquest context, recomanem alimentar el PRC per una doble via:

- l'actual, que extreu informació dels sistemes interns de les institucions i en garanteix així l'exhaustivitat, i 2 2 / 2 9 Informe sobre els indicadors bibliomètrics al Portal de la Recerca de Catalunya (PRC)
- la nova, que incorporaria informació de bases de dades comercials, o d'altres fonts rellevants, que afegirien valor i qualitat a la informació que ja hi ha al portal."

⁸ <http://hdl.handle.net/2072/521698>

L'informe acaba amb un decàleg de recomanacions. Resumidament, aquestes són:

1. Incorporar de forma progressiva la informació de publicacions dels centres sanitaris, els instituts i centres catalans del CSIC, les grans infraestructures científiques i els centres tecnològics.
2. La implementació de nous indicadors es pot realitzar per fases, amb el corresponent càlcul de costos de personal, serveis externs (proveïdors de dades) i tecnologia (eina analítica, gràfics i visualitzacions, etc.).
3. Els sistemes d'informació de les institucions participants han de subministrar l'identificador de cada publicació (sobretot, el DOI) i el gènere dels investigadors.
4. Les dades a sol·licitar a proveïdors externs.
5. Revisar el tractament actual dels canvis d'institució, jubilacions o trasllats a l'estranger, per presentar els indicadors adequats d'una part considerable de la producció científica generada des de les institucions catalanes.
6. Estudiar la possibilitat d'incloure al PRC les citacions d'altres proveïdors (per exemple, Dialnet) per ampliar la cobertura de les citacions de, principalment, les àrees de ciències socials i humanitats.
7. Implementar la classificació de revistes de Web of Science
8. Establir criteris uniformes sobre cada tipus de document dels inclosos (per exemple, quines subtipologies inclouen) i comprovar-ne l'aplicació per part de les institucions, sobretot en el cas dels articles de revista.
9. En la mesura que sigui possible, els indicadors bibliomètrics del PRC s'han de construir a partir de les dades de cada publicació.
10. Donar més visibilitat als indicadors bibliomètrics disponibles i implantar directament al PRC algunes opcions de consulta (exportar; descarregar; compartir; informar errada) i tecnològiques (accés al PRC via API), per facilitar la tasca bibliomètrica i la interoperabilitat amb altres sistemes d'informació.

Annex 3

Sistemes usats pels integrants del PRC per aportar informació

Actualment la informació de les universitats i centres de recerca es carrega al Portal a través de 3 sistemes diferents: CERIF-XML, Excel i càrrega externa. L'objectiu a mig termini hauria de ser que tots els participants aportessin els continguts a través del seu propi sistema de gestió de la recerca (CRIS) en format CERIF-XML. Dels CRIS que es fan servir a Catalunya hem treballat amb tots per modernitzar-los i aconseguir que implementin l'ús de l'estàndard internacional CERIF per afavorir la interoperabilitat dels continguts, i per tant també la participació al Portal.

A continuació es detalla la font des d'on es carreguen les dades per les 51 institucions participants actualment al PRC.

Universitat/Centre de recerca	Font	Format
UB	GREC	CERIF-XML
UAB	PURE	CERIF-XML
UPC	DRAC	CERIF-XML
UPF	SIGMA	CERIF-XML
UdG	GREC	CERIF-XML
UdL	GREC	CERIF-XML
URV	UXXI (projectes i grups) / iMarina (producció científica)	CERIF-XML
UOC	UXXI	CERIF-XML
URL	GREC	CERIF-XML
UVic-UCC	SIGMA	CERIF-XML
UIC	DRAC	CERIF-XML
UAO	iMarina	CERIF-XML
Agrotecnio	GREC-UdL	CERIF-XML
CED**	No tenen cap sistema	Excel
CIMNE**	No tenen cap sistema	Excel
CRAG**	Càrrega externa	CERIF-XML
CREAF**	Càrrega externa	CERIF-XML
CREI**	Càrrega externa	CERIF-XML
CRG	iMarina	CERIF-XML
CRM**	Càrrega externa	CERIF-XML
CTFC*	Càrrega externa	CERIF-XML
CTTC*	Càrrega externa	CERIF-XML
CVC	Sistema inf. propi	CERIF-XML
i2CAT	Sistema inf. propi	CERIF-XML
IBEC	iMarina	CERIF-XML
ICAC	Sistema inf. propi	CERIF-XML
ICFO	iMarina	CERIF-XML

ICIQ	iMarina	CERIF-XML
ICN2	iMarina	CERIF-XML
ICP**	No tenen cap sistema	Excel
ICRA	Sistema inf. propi	CERIF-XML
ICRPC**	No tenen cap sistema	Excel
IDIBAPS	iMarina	CERIF-XML
IDIBELL	Fundanet	CERIF-XML
IDIBGI*	Càrrega externa	CERIF-XML
IEEC**	Càrrega externa	CERIF-XML
IFAE	iMarina	CERIF-XML
IGTP**	Càrrega externa	CERIF-XML
IIBSantPau	Fundanet	CERIF-XML
IISPV*	Càrrega externa	CERIF-XML
IJC	Sistema inf. propi	CERIF-XML
IMIM	Sistema inf. propi	CERIF-XML
IPHES**	No tenen cap sistema	Excel
IRB Barcelona	iMarina	CERIF-XML
IRB Lleida	Fundanet	CERIF-XML
IREC**	Càrrega externa	CERIF-XML
IrsiCaixa	Fontventa	CERIF-XML
IRTA	iMarina	CERIF-XML
ISGLOBAL	iMarina	CERIF-XML
VHIO	Fundanet	CERIF-XML
VHIR	Sistema inf. propi (Hipòcrates)	CERIF-XML

Notes:

- Les institucions marcades amb * (x) properament disposaran d'un CRIS que exportarà les dades en CERIF-XML.
- Les institucions marcades amb ** són institucions per les que s'està definint el procediment per poder usar un entorn DSpace-CRIS per proporcionar des d'allà les dades actualitzades al PRC en CERIF-XML.