



**RESOLUCIÓ DE LA CONVOCATÒRIA D'AJUTS PER A LA RECERCA 2025  
(NOUS PROJECTES DE RECERCA CIENTÍFICA EN L'ÀMBIT DE LA CATALANÍSTICA)**

La Comissió d'Investigació, reunida el propassat 3 de desembre, un cop examinades les avaluacions i les propostes de la Secretaria Científica, eleva al Consell Permanent de l'IEC del dia 12 de desembre de 2024 la proposta de finançament que es detalla a continuació, condicionada a la confirmació de la disponibilitat pressupostària prevista per a cada anualitat i, en el cas de les concessions previstes per al 2026 i 2027, condicionada també a l'informe favorable del progrés aconseguit al final de cada anualitat finançada:

*Objectiu a. Desenvolupament de recursos i models per a l'ús de la intel·ligència artificial aplicada a la llengua catalana*

Investigadora principal: Mercè Lorente (Secció Filològica)

Títol: CERCACAT: Extracció automàtica de terminologia i neologia en llengua anglesa de les publicacions dels centres CERCA de Catalunya per a la planificació terminològica en llengua catalana

Import concedit 2025: 20.000 €

Import concedit 2026: 20.000 €

Import concedit 2027: 20.000 €

Import total concedit (2025-2027): 60.000 €

Investigador principal: Antoni Oliver

Ponent: Ramon López (Secció de Ciències i Tecnologia)

Títol: Eines d'intel·ligència artificial per al foment de la comunicació científica en català

Import concedit 2025: 19.777 €

Import concedit 2026: 19.777 €

Import concedit 2027: 14.050 €

Import total concedit (2025-2027): 53.604 €

*Objectiu b. Programes d'intervenció per a la millora de l'ús social i l'ús científic del català*

(Desert)

*Objectiu c. Estudis sobre diversitat en relació amb els moviments migratoris, l'equilibri interterritorial, natural o l'ambiental*

Investigadora principal: Alba Abras

Ponent: Joaquim Gosálbez (Secció de Ciències Biològiques)

Títol: Conservació de la bagra catalana, *Squalius laietanus*, a les conques fluvials de Catalunya: integritat genètica de les poblacions

Import concedit 2025: 18.730 €

Import concedit 2026: 12.820 €

Import concedit 2027: 0 €

Import total concedit (2025-2027): 31.550 €

Investigadora principal: Elisabet Alacid

Ponent: F. Xavier Turon (Secció de Ciències Biològiques)

Títol: Avaluació de la Perkinsosi al delta de l'Ebre: una malaltia marina emergent que amenaça l'aqüicultura del musclo a Catalunya

Import concedit 2025: 14.280 €

Import concedit 2026: 9.650 €

Import concedit 2027: 0 €

Import total concedit (2025-2027): 23.930 €

Investigador principal: Andreu Domingo (Secció de Filosofia i Ciències Socials)

Títol: Migracions i diversitat a Catalunya: reproducció demogràfica, estratificació social i llengua catalana al segle XXI

Import concedit 2025: 20.000 €

Import concedit 2026: 20.000 €

Import concedit 2027: 20.000 €

Import total concedit (2025-2027): 60.000 €

*Objectiu d. Estudis sobre diversitat en relació amb els reptes comunicatius (xarxes, etc.), educatius o de salut*

(Desert)

*Objectiu e. Elaboració de recursos, preferiblement consultables en línia, relacionats amb el patrimoni natural, material i immaterial dels territoris catalanoparlants i la seva cultura al llarg de la història*

Investigador principal: Manuel Pérez Saldanya (Secció Filològica)

Títol: Corpus Gramatical del Català (CGC)

Import concedit 2025: 12.500 €

Import concedit 2026: 12.500 €

Import concedit 2027: 12.500 €

Import total concedit (2025-2027): 37.500 €

## RESUMS DELS PROJECTES CONCEDITS

**Títol:** CERCACAT: Extracció automàtica de terminologia i neologia en llengua anglesa de les publicacions dels centres CERCA de Catalunya per a la planificació terminològica en llengua catalana

**Investigadora principal:** Mercè Lorente (Secció Filològica)

**Institució:** Universitat Pompeu Fabra (UPF)

**Import total concedit:** 60.000 €

### **Resum del projecte:**

El projecte CERCACAT vol posar a disposició de la comunitat científica i de les institucions que han de vetllar per la normalització i l'increment de l'ús de la llengua catalana la terminologia emprada en anglès pels investigadors dels centres de recerca de la Generalitat de Catalunya, amb l'objectiu d'analitzar-ne les novetats conceptuals en àmbits de ciència, a través de la denominació en anglès, i la possibilitat d'establir la denominació en català, ja sigui adoptant o adaptant el manlleu, ja sigui buscant un equivalent adequat i viable.

La metodologia del projecte —basada en la desenvolupada i actualitzada en els projectes anteriors Guaita Terminològica— consisteix bàsicament en la constitució de sengles corpus textuais segons àrees de coneixement, a partir de les publicacions dels anys 2023-2024 dels investigadors dels centres CERCA (títols, resums i textos complets, quan sigui possible), la creació de sengles corpus d'exclusió lexicogràfica, l'extracció automàtica de candidats a terme, amb indicació de la freqüència i grau de terminològicitat dels candidats. Els resultats previstos són llistes d'unitats terminològiques i/o neològiques que encara no tenen equivalent en català perquè el TERMCAT, Centre de Terminologia de Catalunya, en pugui fer propostes adients amb la col·laboració dels membres de l'equip investigador del projecte i d'investigadors dels centres CERCA.

El projecte s'organitza com a projecte plurianual, per poder analitzar els materials en anglès publicats en el període 2023-2024, i poder fer propostes adients en llengua catalana. Els materials que es podran processar en el primer any del projecte corresponen als articles publicats en revistes d'impacte dels investigadors dels centres CERCA del període 2023-2024. Inicialment els responsables de CERCA han lliurat les referències de 3.824 articles de 21 centres diferents, des de l'1 de gener de 2023 fins al 30 d'abril de 2024. Per a disposar d'una referència del volum de dades a processar, els centres CERCA van publicar un total de 48.461 entre 2018 i 2021 (uns 12.000 articles per any aproximadament).

L'adscripció del projecte en el tema *a* (Desenvolupament de recursos i models per a l'ús de la intel·ligència artificial aplicada a la llengua catalana) està justificada perquè els resultats són recursos lingüístics i perquè en la fase d'extracció automàtica de terminologia està previst interactuar també amb aplicacions d'IA generativa textual, per avaluar els resultats de diferents algorismes i grans models de llenguatge, comparar-los i preveure millores en la metodologia per a anys posteriors. Amb tot i això, cal indicar que el projecte està relacionat també amb el tema *b* (Programes d'intervenció per a la millora de l'ús social i l'ús científic del català), atès que l'acció s'inclou també en el «Pla d'enfortiment de la recerca en català».

**Títol:** Eines d'intel·ligència artificial per al foment de la comunicació científica en català

**Investigador principal:** Antoni Oliver

**Ponent:** Ramon López (Secció de Ciències i Tecnologia)

**Institució:** Universitat Oberta de Catalunya (UOC)

**Import total concedit:** 53.604 €

**Resum del projecte:**

Els científics tenen dificultats per a difondre la seva recerca en català, tant de forma escrita com oral, i sovint fan servir altres llengües, com l'anglès. En els darrers anys han sorgit un gran nombre d'eines relacionades amb la intel·ligència artificial que permeten la traducció de textos i veu.

Aquest projecte té com a objectiu principal facilitar l'ús d'eines tecnològiques que permetin fer difusió científica en català, tant de forma escrita com oral, i fer-la accessible en anglès.

Aquestes eines inclourien:

- Traducció automàtica text - text, és a dir, la traducció de textos científics del català a l'anglès.
- Traducció automàtica veu - text, és a dir, oferir un text en anglès d'un discurs oral en català, que podria ser tant una locució en viu com enregistrada, en forma d'àudio o vídeo.
- Traducció automàtica veu - veu, és a dir, oferir traducció simultània d'una comunicació oral en català amb sortida d'àudio en anglès.

La facilitació d'aquestes eines es concretaria en tres nivells:

- Anàlisi i avaluació de solucions comercials disponibles.
- Anàlisi i avaluació de solucions lliures disponibles.
- Creació d'eines pròpies, o des de zero o amb la millora d'algunes de les solucions lliures disponibles.

**Títol:** Conservació de la bagra catalana, *Squalius laietanus*, a les conques fluvials de Catalunya: integritat genètica de les poblacions

**Investigadora principal:** Alba Abras

**Ponent:** Joaquim Gosálbez (Secció de Ciències Biològiques)

**Institució:** Universitat de Girona (UdG)

**Import total concedit:** 31.550€

### **Resum del projecte:**

El gènere *Squalius*, àmpliament distribuït i diversificat a la regió Mediterrània, representa un gran component de l'ictiofauna autòctona ibèrica que inclou nou espècies endèmiques. Fins a l'any 2007, es considerava que la conca de l'Ebre i els rius mediterranis al nord de la conca estaven habitats per la bagra europea, *S. cephalus*. Mitjançant anàlisis morfològiques i del DNA mitocondrial, es van evidenciar diferències entre la bagra de l'Ebre i les poblacions europees de *S. cephalus*, que van portar a descriure una nova espècie, la bagra catalana, *S. laietanus*. La bagra catalana es distribueix pels rius de Catalunya, de l'Ebre fins a Portbou, i als rius francesos Tec, Tet, Maçana i Aglí. D'acord amb la Llei 22/2009, del 23 de desembre, d'ordenació sostenible de la pesca en aigües continentals a Catalunya, *S. laietanus* es classifica com a espècie pescable. No obstant això, les poblacions de bagra catalana es troben en regressió i la Unió Internacional per a la Conservació de la Natura (UICN) va considerar, l'any 2022, que el seu estat de conservació, establert com de mínima preocupació el 2008, requeria una revisió que encara està pendent. Estudis realitzats sobre les poblacions de bagra en els rius francesos suggereixen la hibridació entre *S. laietanus* i *S. cephalus* a l'Aglí i el Tec. En conseqüència, *S. laietanus* es considera amenaçada a França des del 2019, ja que podria ser desplaçada i substituïda per la bagra europea si es mantenen els encreuaments híbrids. En aquest context, cal comprovar si aquesta amenaça és present també en els rius catalans.

En un estudi previ, el Laboratori d'Ictiologia Genètica (LIG) va analitzar la integritat genètica de *S. laietanus* a les conques fluvials catalanes, des de la Muga fins a l'Ebre, durant el període 2021-2023. Es va tipificar el llinatge mitocondrial dels exemplars mitjançant el marcador COI i, en els trams de rius amb presència comprovada de *S. laietanus*, es va determinar si existia hibridació amb *S. cephalus* mitjançant l'anàlisi del marcador nuclear RAG1. Els resultats preliminars indiquen que només en el Daró, el Francolí i el Gaià no hi ha *S. cephalus*, mentre que al Fluvià, la Muga, el Ter i l'Ebre, el grau d'hibridació entre ambdues espècies és molt elevat en els trams analitzats. El Llobregat i la Tordera mantenen poblacions pures de bagra catalana, però amb trams on també és present la bagra europea. Els objectius són (i) localitzar poblacions pures de bagra catalana en les conques en les quals fins ara només s'ha detectat la presència d'híbrids amb *S. cephalus*; (ii) confirmar la integritat de l'espècie al llarg del curs d'aquells rius on, fins ara, les prospeccions fetes només han detectat la presència de poblacions de *S. laietanus*; (iii) establir el grau de diferenciació poblacional entre les poblacions pures de *S. laietanus* dels rius estudiats mitjançant marcadors mitocondrials addicionals; i (iv) desenvolupar marcadors nuclears complementaris que permetin diferenciar *S. laietanus* de la resta d'espècies del gènere, principalment aquelles natives de la península Ibèrica. Els resultats obtinguts seran directament aplicables als plans de gestió i conservació d'aquesta espècie endèmica com a part del patrimoni natural dels països catalans.

**Títol:** Avaluació de la Perkinsosi al delta de l'Ebre: una malaltia marina emergent que amenaça l'aqüicultura del musclo a Catalunya

**Investigadora principal:** Elisabet Alacid

**Ponent:** F. Xavier Turon (Secció de Ciències Biològiques)

**Institució:** Centre d'Estudis Avançats de Blanes (CEAB-CSIC)

**Import total concedit:** 23.930 €

**Resum del projecte:**

L'objectiu d'aquest projecte és avaluar l'estat de la malaltia Perkinsosi en musclos cultivats al delta de l'Ebre per tal d'entendre el potencial impacte que té en el sector de l'aqüicultura a Catalunya.

Les malalties marines emergents causades per protists suposen una preocupació creixent, ja que afecten la salut dels ecosistemes i produeixen pèrdues econòmiques importants per a la societat, especialment en la pesca i l'aqüicultura. La malaltia marina emergent més significant en l'aqüicultura de mol·luscs és la Perkinsosi, causada per paràsits protists del gènere *Perkinsus*. Aquesta malaltia ha estat notificada arreu del món, afecta una gran varietat d'espècies de bivalves, i és responsable de casos de mortalitat massiva a causa de la seva gran capacitat invasiva i l'alta virulència. Actualment, s'han descrit set espècies dins del gènere *Perkinsus*, algunes d'aquestes espècies incloses en la llista de paràsits de declaració obligatòria pel Codi Aquàtic de l'Organització Mundial de Sanitat Animal.

El principal centre de producció de mol·luscs al Mediterrani i la segona àrea de producció a Espanya són les badies d'estuaris del delta de l'Ebre a Catalunya, on es cultiven unes 4.000 tones per any de musclos i ostres. Diversos estudis han informat de la presència d'espècies de *Perkinsus* que infecten mol·luscs al delta de l'Ebre, com ara cloïsses i escopinyes; i molt recentment s'ha observat la infecció en musclos per primera vegada. És una amenaça per al sector de l'aqüicultura a Catalunya. Malgrat l'impacte negatiu d'aquesta malaltia, no hi ha estudis que n'avaluïn l'epidemiologia (diversitat, distribució, prevalença i la intensitat de les seves infeccions) al delta de l'Ebre. A més, recentment, s'ha observat que el desenvolupament de malalties en animals es deu no només a una espècie, sinó al resultat de les interaccions entre l'hoste i els seus simbiotes. En el cas de la Perkinsosi, no hi ha cap estudi que avaluï com el microbioma afecta la progressió de la malaltia en els hostes infectats.

La gestió de la Perkinsosi en bivalves suposa un repte addicional, a causa del seu sistema immunitari adaptatiu de memòria limitada, que fa que els seus mecanismes de defensa es basin en les funcions immunitàries innates. Actualment, no existeix un tractament eficaç per a combatre aquesta malaltia, i els plans de vigilància i control esdevenen claus per a la seva prevenció. Per tant, una millor comprensió de la seva epidemiologia és crucial per a entendre el potencial impacte, per a desenvolupar protocols de vigilància efectius, programes de monitoratge per a la detecció precoç, i per a implementar estratègies de gestió eficaces per a prevenir la propagació de les infeccions i reduir, així, les pèrdues econòmiques en el sector de l'aqüicultura a Catalunya.

**Títol:** Migracions i diversitat a Catalunya: reproducció demogràfica, estratificació social i llengua catalana al segle XXI

**Investigador principal:** Andreu Domingo (Secció de Filosofia i Ciències Socials)

**Institució:** Centre d'Estudis Demogràfics (CED)

**Import total concedit:** 60.000 €

### **Resum del projecte:**

El projecte proposa testar la hipòtesi principal que diu que per a entendre la situació actual del català, cal abordar un estudi sistèmic de la relació entre la centralitat de les migracions en la reproducció demogràfica i social (i, per tant, en l'estratificació social), i la diversitat en relació amb el paper i ús de la llengua en la mobilitat social.

La demografia i, especialment, les migracions, en el millor dels casos, solen aparèixer com a element contextual per a explicar el retrocés de l'ús de la llengua catalana en termes relatius, malgrat l'augment en nombres absoluts. En el pitjor dels casos, es podria tornar a la visió vandelloosiana, en què el futur del català (i de la nació catalana) es trobaria amenaçat tant per la fecunditat diferencial entre autòctones i immigrades, com per la intensitat de la migració, en què la substitució poblacional per força avançava la substitució lingüística. Una visió sistèmica, en canvi, a més de la mesura de l'impacte de les migracions en el territori (diversitat, concentració i segregació), planteja el seu efecte en la creixent heterogeneïtat per origen de les generacions, especialment en les més recents, i en l'anàlisi del desplaçament de la llengua catalana com a element central de la mobilitat social ascendent, en un context de reconfiguració de l'estratificació social, tenint en compte la migració massiva de caràcter internacional.

El projecte s'organitza al voltant de quatre blocs corresponents als objectius concrets: 1) Anàlisi de l'impacte de les migracions i la diversitat en la població i el territori; 2) Estudi de les transformacions en l'estratificació social a partir de la integració segmentada dels immigrants i la seva descendència; 3) Relacionar els canvis en la reproducció demogràfica i social en l'ús de la llengua catalana; i 4) Realitzar una diagnosi de la situació des d'una perspectiva interseccional que contempli gènere, origen i classe social.

Volem destacar l'aproximació metodològica, amb l'ús de fonts amb sèries estadístiques de població més enllà de les enquestes d'usos lingüístics (el Cens, l'ECEPOV, el padró continu, el Cens de població anual, l'Enquesta de Condicions de Vida, entre d'altres). L'anàlisi s'ha restringit al cas de Catalunya i no al conjunt de les terres de parla catalana per dos factors principals: d'una banda, per la complexitat de situacions demogràficament i socialment diferenciades, que excediria amb molt els recursos materials dels quals es disposa per a desenvolupar l'estudi; i, per l'altra banda, l'accessibilitat i existència de dades estadístiques necessàries per a dur a termini l'anàlisi, inexistents o inaccessibles en altres territoris.

Finalment, s'ha d'assenyalar que l'adopció d'una perspectiva sistèmica des de la demografia, sense canviar els resultats dels càlculs de la mesura dels efectes de la migració i el creixement de la diversitat al territori i la població, pot modificar-ne la interpretació i, per tant, pot orientar la intervenció política des de noves perspectives.



**Títol:** Corpus Gramatical del Català (CGC)

**Investigador principal:** Manuel Pérez Saldanya (Secció Filològica)

**Import total concedit:** 37.500 €

**Resum del projecte:**

El projecte consisteix en la digitalització de la producció científica i divulgativa centrada en l'estudi de la gramàtica catalana i en l'elaboració i posada en funcionament d'una interfície de consulta tant de les metadades com del contingut de les obres. El producte final serà un Corpus Gramatical del Català (CGC) que permetrà recuperar la informació sobre qualsevol aspecte gramatical de la llengua catalana, la qual cosa implicarà:

- a) Reunir i digitalitzar les publicacions més rellevants sobre gramàtica catalana d'ençà de l'establiment de la normativa lingüística prenent com a referent la bibliografia de la *Gramàtica de la llengua catalana* (amb més de 3.500 referències bibliogràfiques), del 2019, amb la pertinent actualització fins al 2024, i planificar-ne l'actualització successiva.
- b) Gestionar l'obtenció dels documents:
  - En el cas de documents ja digitalitzats d'accés obert, obtenir-los directament o mitjançant un conveni de cessió dels repositoris que els contenen, com ARCA, DRAC, BiValDi, Dialnet, MDC, RACO, TDX, Traces o de les universitats, entre altres, i aprofitant la digitalització duta a terme per l'IEC Obert.
  - En el cas dels documents ja digitalitzats d'accés restringit, gestionar-ne la cessió gratuïta directament de la persona editora o de l'autora o l'obtenció mitjançant subscripció o compra.
  - En el cas dels documents no digitalitzats, gestionar l'accés a la versió impresa i planificar-ne i emprendre'n la digitalització.
- c) Convertir a l'estàndard Unicode, sempre que el format documental ho permeti, la diversitat de codificació de caràcters existent en la documentació (alternativament incorporar al motor d'indexació documental del corpus una funcionalitat que permeti superar l'heterogeneïtat de codificacions de caràcters dels documents incorporats al corpus).
- d) Dissenyar un sistema d'indexació, consulta, gestió i explotació amb un cercador potent i simple, adaptat a cerques multilingües i sobre documents en català prenormatiu o no estandarditzat i que compleixi els requisits d'adaptabilitat, accessibilitat i usabilitat.
- e) Proporcionar un sistema d'obtenció de la informació continguda a les fonts que compatibilitzi l'accés il·limitat amb la preservació dels drets d'autor, de la propietat intel·lectual i de còpia en el cas que els titulars no n'hagin fet la cessió a l'IEC (especialment, però no pas exclusivament, per a publicacions alienes a l'IEC o als seus membres actuals i pretèrits).
- f) Facilitar l'extracció de bibliografies a partir dels resultats de cerca, la generació de citacions bibliogràfiques d'acord amb els formats bibliogràfics amb més difusió (ISO, APA, MLA, etc.) i l'exportació a gestors de referències bibliogràfiques (Mendeley, Zotero, etc.).

El nou CGC pretén ser una clau d'accés potent i fiable a tot el coneixement que avui tenim de la llengua catalana reunint la part més significativa i influent de tota la producció de recerca, codificació i divulgació existent en aquesta àrea del coneixement.