

## CONVENIO PLURIANUAL DE COLABORACIÓN ENTRE A CONSELLERÍA DE CULTURA, EDUCACIÓN E UNIVERSIDADE E A UNIVERSIDADE DE SANTIAGO PARA A REALIZACIÓN DUN PROGRAMA DE ACCIÓNS EN MATERIA DE INVESTIGACIÓN A TRAVÉS DO CENTRO RAMÓN PIÑEIRO PARA A INVESTIGACIÓN EN HUMANIDADES

En Santiago de Compostela,

### REUNIDOS:

Dunha parte, Román Rodríguez Rodríguez González, conselleiro de Cultura, Educación e Universidade, que actúa en virtude das facultades atribuídas no artigo 34 da Lei 1/1983, do 22 de febreiro, reguladora da Xunta e da súa presidencia, modificada pola Lei 12/2007, do 27 de xullo, de acordo co seu nomeamento polo Decreto 112/2020, de 6 de setembro, e no uso das atribucións que lle confire o Decreto 130/2020, de 17 de setembro (DOG nº 190, do 18 de setembro), polo que se establece a estrutura orgánica da Consellería de Cultura, Educación e Universidade.

**Doutra, Don Antonio López Díaz**, reitor magnífico da Universidade de Santiago de Compostela, en virtude do establecido nos seus estatutos e do nomeamento como reitor polo Decreto 57/2018, do 31 de maio, (DOG nº 106 de 5 de xuño), actuando en nome e representación da devandita entidade.

As partes actúan no exercicio dos seus respectivos cargos e coa representación competencial que ostentan, e recoñecéndose mutuamente capacidade para se obrigaren nos termos do presente convenio,

### EXPOÑEN:

1. Que para potenciar o emprego do galego, segundo establecen os artigos 5.3 do Estatuto de Autonomía de Galicia e 6.3 da Lei de normalización lingüística, cómpre dispor de estudos científicos sobre diversos eidos da lingua galega, que contribúan a protexela, promovela e modernizala para que sexa un medio de comunicación eficiente en calquera ámbito.
2. Que no Decreto 198/2020, do 20 de novembro (DOG do 1 de decembro) establécese a estrutura orgánica da Consellería de Cultura, Educación e Universidade, e



atribúeselle á Secretaría Xeral de Política Lingüística (en diante, SXPL), entre outras, as funcións de promover, coordinar e desenvolver a política de investigación que favoreza a normalización lingüística nas súas diferentes manifestacións, así como, promover a cooperación e a colaboración con outras institucións competentes en materia de política lingüística, contribuíndo así á súa difusión.

**3.** Que a Universidade de Santiago de Compostela (en diante USC) é unha entidade que leva a cabo actividades de investigación, docencia e desenvolvemento científico e tecnolóxico, e ten como un dos seus fins o de procurar o seu enraizamento na sociedade galega que a sustenta e, en xeral, contribuír ao seu desenvolvemento social, económico e cultural.

**4.** Que estes fins son apoiados e compartidos pola Consellería de Cultura, Educación e Universidade (en diante consellería). Así, co fin de xestionar o desenvolvemento de proxectos e programas de investigación e conservación das fontes lingüísticas, literarias e antropolóxicas que permitan un estudo profundo da nosa historia e que garantan a utilidade da lingua galega na sociedade actual, con data 11 de febreiro de 1993, foi creado o Centro de Investigacións lingüísticas e literarias Ramón Piñeiro, posteriormente denominado Centro Ramón Piñeiro para a Investigación en Humanidades (en diante CRPIH), polo Decreto 330/1997, do 13 de novembro, dependente organicamente da SXPL.

**5.** Que a dixitalización do coñecemento e a busca da innovación a través da tecnoloxía e da apertura de datos están volvéndose eixos claves da axenda dixital europea e que un dos piares da súa proposta estratéxica de investigación é a transferencia á sociedade e que, neste sentido, o CRPIH xa hai máis dunha década que vén impulsando o acceso aberto aos contidos de investigación, a través de presentacións sinxelas, accesibles, gratuítas e actualizadas.

**6.** Que as partes son conscientes da necesidade actual de promover as Humanidades dixitais (HD), co propósito de acadar unha mellor infraestrutura para a xestión de datos, recursos dixitais e ferramentas, que permitan a súa compartición e reutilización, aumentando a súa canalización, difusión, proxección e recoñecemento. Así, consideran que o acceso e intercambio de recursos entre elas -toda vez que existen complementariedades e sinerxías en proxectos de interese recíproco- pode ser explotado para acadar unha maior dimensión, mediante a interacción e colaboración científica entre os axentes participantes, o que supón unha optimización dos recursos destinados para o mesmo fin e redundante nun beneficio de ben común a prol do interese público.



7. Que, posto que os seus intereses son concorrentes e perseguen a obtención dunha finalidade común, a USC e o CRPIH veñen colaborando ininterrompidamente desde o ano 1995 (sinatura do primeiro convenio) no desenvolvemento conxunto de actuacións de investigación relacionadas coas áreas de coñecemento mencionadas. Isto amosa un destacado compromiso en continuar cooperando nestas materias de estudo e a USC -en clave de visión xeral sobre o nivel actual do estado das investigacións- é quen dispón dos coñecementos necesarios para propiciar as contornas de colaboración que favorezan unha axeitada continuidade dos proxectos de investigación en curso que se desenvolven no centro.

8. Que, co propósito de que as tarefas encomendadas ao centro acaden os resultados axeitados, a consellería desexa seguir contando coa colaboración dos investigadores de maior reputación en cada un destes ámbitos e, dadas as particularidades e características específicas dos proxectos de investigación de que se trata, tendo en conta que os departamentos de Filoloxía da USC contan con recursos tecnolóxicos axeitados e con destacados investigadores de recoñecida capacidade científica -en relación á natureza dos proxectos incluídos no anexo deste convenio- os obxectivos perseguidos non poderían acadarse co mesmo nivel de excelencia nun escenario de concorrencia pública.

9. Que é de marcado interese xeral e das partes asinantes, no marco da colaboración establecida, incorporar novas perspectivas -na liña da *Estratexia Galicia Dixital 2030*, (particularmente no relacionado co capítulo 9.1 "Proxección dos sinais de identidade de Galicia")- que potencien a innovación dixital aplicada á dinamización e salvagarda do patrimonio e a produción cultural da Comunidade, como recurso accesible e compoñente fundamental da nosa tradición, identidade e recoñecemento colectivo, incluíndo a lingua galega como creadora de cultura e vehículo de posibilidades representativas, co fin de ampliar o alcance e a accesibilidade dos traballos e resultados da execución destes proxectos, facendo deles unhas ferramentas modernas e eficaces de referencia científica mediante a introdución e mellora no manexo de novos usos asociados ás tecnoloxías da información e da comunicación (TIC).

10. Que a USC dispón de investigadores de valía profesional e prestixio recoñecido que se dedican sectorialmente á realización deste tipo de actividades e cos que se pode establecer un espazo común de intereses no eido da investigación, que permiten a utilización da figura do convenio de colaboración como fórmula comercial apropiada para articular as súas relacións, dado que non se manifesta nunha contraposición de intereses, senón que se trata de establecer unha colaboración institucional para levar a cabo unha actuación en resposta de obxectivos



compartidos e os intereses de ambas as dúas partes teñen por causa a consecución dun fin común en beneficio da comunidade.

Así mesmo, ao abeiro do establecido no artigo 6 –“*Convenios e encomendas de xestión*”–, da Lei 9/2017 de 8 de novembro de contratos do sector público, este caso considérase excluído e plenamente aplicable á figura do convenio de colaboración debido ás prestacións de investigación e desenvolvemento que, con respecto aos proxectos incluídos no anexo do presente convenio, van ser levadas a cabo por parte da Universidade e que son susceptibles de se incluír no punto 1, apartados b) e c) do citado artigo.

Por todo isto, a consellería e mais a USC establecen o presente convenio de colaboración que se rexerá polas seguintes

## ESTIPULACIÓNS:

### **Primeira. Obxecto do convenio**

Mediante este convenio establécense as condicións polas que se rexerá a colaboración entre a consellería e a USC, a través da SXPL, para proseguir desenvolvendo o programa de accións en materia de investigación, especialmente no ámbito lingüístico, que se está levando a cabo no CRPIH e que se describe no anexo citado na estipulación segunda.

O fin último desta cooperación é desenvolver proxectos de investigación, fundamentais para a comunidade científica e para toda a sociedade, capaces de garantir a utilidade da lingua galega en todos os ámbitos da vida actual e de contribuír, asemade, ao acceso igualitario do coñecemento no ámbito das Humanidades, en xeral, e das HD, en particular.

### **Segunda. Obrigas da USC**

As actividades que ha desenvolver a USC en colaboración co centro, recóllense no ANEXO I ao presente convenio onde se especifican os traballos que se realizarán en cada unha das accións que conforman o programa de investigación.

A USC cooperará coa consellería e coa SXPL na realización das actividades do programa de actuacións en materia de investigación científica do centro e comprométese a realizar -a través dos seus investigadores- os distintos proxectos que nel se desenvolven:



PROXECTO TERMINOLOXÍA CIENTÍFICO-TÉCNICA (TERMIGAL).  
PROXECTO BASE DE DATOS DO ATLAS LINGÜÍSTICO ROMÁNICO (ALIR)  
PROXECTO BIBLIOGRAFÍA INFORMATIZADA DA LINGUA GALEGA (BILEGA)  
PROXECTO CORPUS DE REFERENCIA DO GALEGO ACTUAL (CORGA/XIADA)  
PROXECTO CORPUS DOCUMENTALE LATINUM GALLAECIAE (CODOLGA)  
PROXECTO ARQUIVO GALICIA MEDIEVAL (ARGAMED)  
PROXECTO LÍRICA GALEGO-PORTUGUESA  
PROXECTO PROSA LITERARIA  
PROXECTO CANTIGAS DE SANTA MARÍA  
BIBLIOGRAFÍA DE REFERENCIA DA LITERATURA GALEGA MEDIEVAL (BIRMED)  
PROXECTO DICIONARIOS DE LITERATURA  
PROXECTO BIBLIOGRAFÍA DA LITERATURA GALEGA  
PROXECTO INFORME ANUAL DE LITERATURA GALEGA  
PROXECTO LITERATURA INFANTIL E XUVENIL  
PROXECTO DIMENSIÓNS DA IDENTIDADE COLECTIVA DE GALICIA

A USC colaborará na realización das distintas accións que integran este programa de investigación cunha contribución en especie, consistente en horas de dedicación á investigación, que serán as estimadas para a consecución dunha axeitada execución das actuacións.

A estes efectos, a valoración económica da devandita contribución equivalería ó resultado do cálculo –en horas- sobre a retribución dun docente investigador –catedrático de universidade (T.C.)-. En consecuencia, a Universidade asumirá, pola súa conta, a cantidade de **noventa e dous mil cincocentos sesenta (92.560,00 €)** derivada dos custos do seu persoal docente investigador, distribuída en catro anualidades (2022, 2023, 2024 e 2025), por un importe de **23.140 ,00 €** cada unha.

Ademais, a USC adoptará as actuacións de soporte técnico necesarias que permitan un intercambio de datos, co fin de contribuír a un mellor desenvolvemento das investigacións, en termos de alcance.



A USC deberá cumprir coas estipulacións xurdidas do vínculo xerado coa celebración do presente convenio no xeito e termos establecidos.

### **Terceira. Obrigas da Consellería de Cultura, Educación e Universidade**

A Consellería ha de propiciar a participación dos profesionais especialistas do CRPIH e a súa contribución activa nos proxectos obxecto de investigacións conxuntas.

Contribuirá á difusión e divulgación dos resultados das investigacións, co fin de canalizar e proxectar os estudos resultantes, con efectos sobre o conxunto da sociedade en xeral e dos profesionais especializados, en particular, non só no ámbito autonómico, senón tamén no nacional e mesmo no internacional.

Facilitará a transferencia de coñecemento científico propio do CRPIH nas áreas temáticas obxecto das investigacións.

Velará polo cumprimento das disposicións, que han de ser executadas no xeito e termos establecidos, en favor do patrimonio público.

A consellería, a través da Secretaría Xeral de Política Lingüística colaborará na realización destas accións cunha achega económica de seiscentos corenta e seis mil seiscentos corenta euros (646.640,00 €), repartidos en catro anualidades coa seguinte distribución:

*Exercicio 2022:* 161.660,00 € con cargo á aplicación orzamentaria 10.50.151A.640.1

*Exercicio 2023:* 161.660,00 € con cargo á aplicación orzamentaria 10.50.151A.640.1

*Exercicio 2024:* 161.660,00 € con cargo á aplicación orzamentaria 10.50.151A.640.1

*Exercicio 2025:* 161.660,00 € con cargo á aplicación orzamentaria 10.50.151A.640.1

Estas cantidades só cubrirán o custo da realización das actividades, polo que non levan asociado ningún beneficio económico para USC.

Para atender os compromisos de gasto derivados das accións establecidas existe crédito axeitado e suficiente na aplicación orzamentaria sinalada, dos presupostos xerais da comunidade autónoma para os citados exercicios orzamentarios.

O pagamento, en recibindo a certificación correspondente dos traballos de investigación realizados, efectuarase en catro entregas repartidas do seguinte xeito:



- 1ª.- 80.830,00 € (1 de xuño de 2022)
- 2ª.- 80.830,00 € (10 de novembro de 2022)
- 3ª.- 80.830,00 € (1 de xuño de 2023)
- 4ª.- 80.830,00 € (10 de novembro de 2023)
- 5ª.- 80.830,00 € (1 de xuño de 2024)
- 6ª.- 80.830,00 € (10 de novembro de 2024)
- 7ª.- 80.830,00 € (10 de xuño de 2025)
- 8ª.- 80.830,00 € (10 de novembro de 2025)

#### **Cuarta. Custo total do convenio**

Tendo en conta as obrigas da USC, relacionadas na cláusula 2ª, e as da consellería, estipuladas na cláusula 3ª, o custo total do convenio ascende a unha cantidade de setecentos trinta e nove mil douscentos (739.200,00 €), distribuídos anualmente tal e como se recolle no Anexo II

#### **Quinta. Libramento**

Para recibir as achegas da Consellería de Cultura, Educación e Universidade que se indican anteriormente, a USC entregará na Secretaría Xeral de Política Lingüística, antes das datas indicadas, a seguinte documentación:

- Certificación da persoa xerente da universidade que acredite a execución das actividades do convenio ao longo do exercicio correspondente, xunto coas pertinentes notas de cargo emitidas polo servizo que lle corresponda á USC para a súa tramitación, así como unha relación valorada das horas dedicadas a cada unha das accións do programa de investigación.
- Unha vez executadas as actividades do convenio, e antes do segundo libramento de cada ano, os investigadores responsables destas emitirán unha memoria detallada das actuacións realizadas.

No caso de que a acreditación dos custos sexa inferior ao importe do convenio, a achega da consellería diminuírase na mesma proporción.

A USC facilitará toda a información que lle sexa requirida pola consellería, así como pola Intervención Xeral da Comunidade Autónoma, Consello de Contas e o Tribunal de Contas no exercicio das súas funcións de fiscalización e control.



Antes da súa sinatura e do seu pagamento, a USC deberá presentar un certificado de non ter contraído débedas, por ningún concepto, coa Administración pública da Comunidade autónoma galega, así como de estar ao día das obrigas tributarias e ao corrente das obrigas sociais.

#### **Sexta. Propiedade ou titularidade dos traballos e publicidade**

A propiedade, titularidade e uso dos traballos será común polas partes e só poderán usalos para os fins de cada una delas sempre de xeito grauíto.

Nas actuacións informativas e divulgativas, nas publicacións que se fagan relacionadas coas actividades deste convenio ou cando a USC ou calquera dos investigadores que participe nas actividades obxecto deste convenio desexe utilizar os resultados acadados para a súa publicación, deberán facer constar, nun lugar destacado, a colaboración da Consellería de Cultura, Educación e Universidade, a través do CRPIH, e da USC, para o que se usarán os anagramas aprobados pola normativa de imaxe corporativa de ambas as institucións.

Sexa cal for o medio de difusión, cada unha das partes comprométese a facer mención da outra parte e deste convenio.

#### **Sétima. Regulación da participación dos investigadores**

A participación dos investigadores da USC nas actividades recollidas no presente convenio, efectuarase conforme á regulamentación propia da USC recollida nos seus estatutos e no Regulamento de cursos de especialización e convenios, e de acordo coa Lei de incompatibilidades do persoal ao servizo das administracións públicas e coa Lei orgánica 6/2001, do 21 de decembro de universidades.

#### **Oitava. Utilización da lingua galega**

Todas as accións e actividades que se realicen ao abeiro deste convenio deberán levar a cabo de acordo co contemplado nas vixentes normas ortográficas e morfolóxicas do idioma galego, aprobadas pola Real Academia Galega no ano 2003, tal e como se establece na disposición adicional da Lei 3/1983, do 15 de xuño, de normalización lingüística. Así mesmo, coidarase especialmente do respecto rigoroso da toponimia oficial nos termos previstos no artigo 10 da citada lei.

#### **Novena. Vixencia**

Este convenio entrará en vigor o mesmo día da súa sinatura e terá vixencia durante os anos **2022, 2023, 2024 e 2025**, sen posibilidade de prórroga, e poderá ser



denunciado por calquera das partes asinantes por incumprimento do que nel se establece, cunha antelación dun mes polo menos, e a consellería poderá, se fose o caso, esixir o reintegro das cantidades achegadas para a súa realización.

#### **Décima. Comisión de seguimento**

Crearase unha comisión de seguimento das accións obxecto do presente convenio-facultada para proporlles ás partes as resolucións de cantas dúbidas xurdan na súa execución- presidida polo secretario xeral de Política Lingüística e integrada polo xerente da USC, polo titular da coordinación científica (membro dos órganos colexiados do CRPIH) e investigador principal e coordinador deste convenio, así como polo secretario do Consello científico e executivo do CRPIH, que actuará como secretario da mesma.

#### **Décimo primeira. Causas de resolución do convenio**

Serán causa de resolución do convenio:

- a) O mutuo acordo das partes.
- b) O incumprimento das estipulacións do mesmo, previa reunión da Comisión de Seguimento.

A denuncia do convenio poderá ser realizada por calquera das partes, debendo comunicalo por escrito coa antelación establecida na cláusula novena.

#### **Décimo segunda. Normas de aplicación supletorias**

Al presente convenio seralle aplicable subsidiariamente a seguinte normativa:

- I. Lei 40/2015, do 1 de outubro, de Réxime Xurídico do Sector Público (artigos 47 e seguintes), na medida que queda excluído do ámbito de aplicación da Lei 9/2017, de 8 de novembro, de Contratos do Sector Público (artigo 6).

#### **Décimo terceira. Protección de datos de carácter persoal e garantía dos dereitos dixitais.**

As partes comprométense a cumprir o establecido no *Regulamento (UE) 679/2016 do Parlamento Europeo e do Consello, de 27 de abril de 2016, relativo á protección de datos das persoas físicas no que respecta ao tratamento de datos persoais e á libre circulación destes datos e polo que se derroga a Directiva 95/46/CE (Regulamento xeral de protección de datos)*, así como na *Lei orgánica 3/2018, de 5 de decembro, de*



*protección de datos persoais e garantía dos dereitos dixitais* e no resto da lexislación española aplicable nesta materia.

O tratamento de datos faise de acordo co regulamento e coa coa lei orgánica citados, así como co resto da lexislación española aplicable, e este tratamento responde aos principios de licitude, lealdade, transparencia, limitación da finalidade, minimización dos datos, exactitude, limitación do prazo de conservación, integridade, confidencialidade e responsabilidade proactiva.

A información que as partes puidesen revelar para a consecución deste convenio e referida ás súas actividades, así como as que se revelen como consecuencia da súa execución terá a consideración de confidencial, debendo gardar segredo sobre toda a información á que poidan ter acceso, coa salvidade de que a mesma sexa de dominio público ou se coñecese legal ou lexitimamente pola outra

#### **Décimo cuarta. Responsabilidades das actuacións**

O asinamento deste convenio non implica relación laboral, contractual ou de calquera outro tipo entre os profesionais que vaian desenvolver as actividades previstas e a Xunta, de tal xeito que a esta non se lle pode esixir responsabilidade ningunha, nin directa, nin indirecta, nin subsidiaria, polos actos ou feitos acaecidos no desenvolvemento do convenio.

#### **Décimo quinta. Rexistro de convenios. Transparencia e bo goberno**

A sinatura do presente convenio suporá o consentimento expreso do interesado á Administración para incluír e facer públicos os datos referidos ao convenio, de conformidade co artigo 15 da Lei 1/2016, do 18 de xaneiro, de transparencia e bo goberno de Galicia e co Decreto 126/2006, do 20 de xullo, polo que se regula o rexistro de convenios da Xunta de Galicia.

Así mesmo, coa súa firma, as partes manifestan o seu consentimento para que os datos persoais que aparecen nel e máis o resto das especificacións contidas no mesmo poidan ser publicados no Portal da Transparencia e Goberno Aberto.

#### **Décimo sexta. Notificación electrónica**

Segundo o artigo 14.2 da Lei 39/2015, do 1 de outubro, do procedemento administrativo común das administracións públicas, os colectivos que están obrigados a relacionarse a través de medios electrónicos coas administracións públicas, e polo tanto, a recibir notificacións por canles electrónicas son: as persoas xurídicas, as entidades sen personalidade xurídica, os colexios profesionais, os que



representen un interesado que estea obrigado a relacionarse electronicamente coa Administración e, o persoal empregado das administracións públicas para os trámites que realicen con elas por razón da súa condición de empregado público.

As notificacións realizaranse só por medios electrónicos a través do Sistema de Notificación Electrónica de Galicia Notific@, <https://notifica.xunta.gal>. Para poder acceder a elas, o interesado deberá contar cun certificado electrónico asociado ao NIF que figure como destinatario da notificación (persoa física ou xurídica).

### **Décimo sétima. Administración electrónica**

As partes asinantes respectarán cantas esixencias establece a Lei 4/2019, do 17 de xullo, de administración dixital de Galicia (DOG 26 de xullo de 2019), que será aplicable ao sector público autonómico, integrado, de acordo co artigo 3 da Lei 16/2010, do 17 de decembro, de organización e funcionamento da Administración xeral do sector público autonómico de Galicia, pola Administración xeral e polas entidades instrumentais e tamén será aplicable á cidadanía que ten o deber de utilizar os servizos públicos dixitais que se poñan á súa disposición de xeito adecuado, co respecto das normas de uso, xerais ou específicas, que se establezan para cada un dos servizos e, en particular, os cidadáns e as cidadás deberán utilizar os sistemas de identificación e de sinatura electrónicas de que sexan lexítimos titulares, usar a información dispoñible conforme a política de privacidade publicada e repectar a normativa en materia de protección de datos de carácter persoal (artigos 2 a) e b) e 7 apartados 1 e 2 respectivamente da citada lei).

### **Décimo oitava. Normativa aplicable e xurisdición**

Este convenio ten carácter administrativo e os seus efectos rexeranse polo establecido nas súas estipulacións, e quedará suxeito ao disposto na disposición adicional segunda da Lei 1/2016, do 18 de xaneiro, de transparencia e bo goberno de Galicia.

Aínda que é un dos excluídos do ámbito de aplicación da Lei 9/2017, do 8 de novembro, de contratos do sector público (BOE 9-11-17) segundo especifica o seu artigo 6, ao non previsto, e de acordo co artigo 4 da citada lei, aplicaráselle os principios recollidos nela, co fin de resolver dúbidas e omisións que puidesen presentarse. Así mesmo, supletoriamente, terase en conta o Decreto legislativo 1/1999, do 7 de outubro, polo que se aproba o texto refundido da Lei de réxime financeiro e orzamentario de Galicia e demais normativa vixente na materia



A xurisdición contencioso administrativa será a competente para dirimir as cuestións que puidesen xurdir da aplicación deste convenio.

En proba de conformidade cos termos deste convenio, as partes comparecentes asinan por triplicado, no lugar e data mencionados.

Pola consellería de Cultura, Educación e  
Universidade

D. Román Rodríguez González  
Conselleiro de Cultura, Educación e Universidade

Pola Universidade de Santiago de  
Compostela

D. Antonio López Díaz  
Reitor da USC



## ANEXO I

As actividades de investigación que ha desenvolver a USC, ás que se fai referencia na cláusula segunda do presente convenio de colaboración, relaciónanse a continuación:

### 1. ÁREA DE LINGÜÍSTICA

#### 1.1. LINGÜÍSTICA COMPUTACIONAL

##### ACCIÓN 1.1.1

#### **PROXECTO TERMIGAL (Servicio Galego de Terminoloxía Científico-Técnica).**

As áreas de traballo nas que se incidirá no período 2022-2025 artículanse arredor de tres eixes principais:

1. **Estudo da terminoloxía puntual.** O Servizo de Terminoloxía Galega dispón dun servizo de consultas a través do cal calquera persoa ou entidade pode formular as súas dúbidas acerca dos equivalentes galegos de distintas voces especializadas. O estudo destas consultas permite detectar os baleiros denominativos existentes e a súa pronta resolución contribúe a potenciar o uso das linguas de especialidade en lingua galega. No convenio suscrito darase continuidade á atención das consultas puntuais, tanto as recibidas no propio servizo como as xestionadas a través da Área de asesoramento lingüístico da Secretaría Xeral de Política Lingüística ou os servizos de normalización lingüística de concellos e determinados organismos. As propostas terminolóxicas serán enviadas periodicamente ao Seminario de Terminoloxía da Real Academia Galega para o seu estudo e, unha vez aprobadas, serán publicadas no banco terminolóxico TERGAL, accesible no enderezo [www.cirp.gal/tergal](http://www.cirp.gal/tergal). Paralelamente, enviaranse ao Seminario de Lexicografía listaxes con aquelas voces de carácter semiespecializado ou pertencente á lingua xeral para a súa incorporación ao Dicionario da Real Academia Galega (DRAG), contribuíndo deste xeito á actualización do repertorio de referencia do galego coas novas voces detectadas.

2. **Estudo da terminoloxía sistemática.** No Servizo de Terminoloxía Galega desenvólvense distintos recursos destinados á recompilación dos termos empregados nas distintas áreas de coñecemento. Anualmente farase unha programación dos proxectos de terminoloxía sistemática nos que se traballará. Con carácter inmediato traballarase sobre a terminoloxía da COVID-19 para a publicación, en colaboración co centro de terminoloxía de Cataluña (Termcat) dun dicionario de termos da COVID-19; continuarase cos traballos de redacción dun "Dicionario do libro", que recolla a terminoloxía relacionada co libro; e seguirá a colaboración coa Academia de Farmacia de Galicia para a elaboración dun Vocabulario galego do ámbito farmacéutico. Darase



continuidade aos acordos de colaboración asinados con distintos organismos para a realización de tarefas de creación ou revisión terminolóxica, así como de avaliación da calidade lingüística dos proxectos que se analicen.

**3. Difusión da terminoloxía.** Entre as actividades relativas á difusión da terminoloxía, é preciso destacar o banco terminolóxico TERGAL, alimentado de maneira continuada cos rexistros procedentes dos traballos puntuais ou sistemáticos e que permite aos usuarios atopar os equivalentes galegos para un amplo volume de termos. Esta ferramenta que crece cada mes é xa un recurso imprescindible para resolver as dúbidas terminolóxicas en galego dos usuarios que traballan con textos orais ou escritos de linguas de especialidade. Proseguirase regularmente co fornecemento para a Comisión de Validación da Real Academia Galega do estudo previo dos termos sobre as linguas de especialidade que se estean a tratar, e continuarase co labor de revisión e correccións de vocabularios de especialidade elaborados fóra de Termigal.

### PROXECTO BASE DE DATOS DO ATLAS LINGÜÍSTICO ROMÁNICO (ALIR)

Continuación cos labores de investigación asociados á elaboración do Atlas Lingüístico Románico e de elaboración, revisión e enriquecemento da base de datos dialectais que reúna o léxico deste proxecto.

O labor do proxecto da Base de datos do ALiR está estreitamente vinculado ao desenvolvemento do proxecto europeo de elaboración do *Atlas Linguistique Roman*. Os labores fundamentais que se levarán a cabo no bienio 2022-2025 serán os seguintes:

A introdución na Base de datos do ALiR de todos os datos relacionados cos sucesivos volumes que se vaian publicando do Atlas Linguistique Roman.

En relación con estes mapas levarase a cabo o estudo e sistematización de todas as denominacións, coa incorporación de:

- denominación en cada punto
- transcripción fonética
- elementos constitutivos da denominación (radical + prefixos + sufixos)
- categoría gramatical
- estudo etimolóxico
- clasificación motivacional, etc.



Creación dunha base de datos externa, que permita a consulta pública e a explotación da información introducida. Esta base servirá non só para os dialectólogos, lexicógrafos, expertos en historia da lingua e etimoloxía, e lingüistas en xeral, senón tamén para todo aquel que desexe obter información relativa ás denominacións nas linguas románicas de conceptos diversos, coñecer as motivacións polas cales a un determinado concepto se lle aplicou unha denominación concreta, coñecer a orixe etimolóxica de cada unha das denominacións, analizar a frecuencia de determinados sufixos etc.

Seguirase tamén contribuíndo á elaboración do Atlas Lingüístico Románico, mediante a redacción das síntesis nacionais referidas ao galego dos conceptos que se traballen dentro deste proxecto europeo durante o tempo de duración do convenio.

Participar activamente nas reunións anuais do Atlas Linguistique Roman e na revisión das síntesis románicas para garantir a corrección e calidade dos datos referidos á lingua galega.

### ACCIÓN 1.1.2

#### **PROXECTO CORPUS DE REFERENCIA DO GALEGO ACTUAL E ETIQUETADOR /LEMATIZADOR DO GALEGO ACTUAL (CORGA/XIADA)**

Estes recursos computacionais abertos e dispoñibles en liña son sumamente importantes para o desenvolvemento de estudos do galego dende múltiples perspectivas: morfolóxica, léxica, sintáctica, discursiva, terminolóxica, fraseolóxica, incorporación de neoloxismos, fixación da normativa etc. Así mesmo, con este proxecto a lingua galega consolídase no mundo das novas tecnoloxías e no do procesamento da linguaxe natural e sitúase ao mesmo nivel das linguas peninsulares e na vangarda da lingüística computacional.

No período 2022-2025 as tarefas que temos previsto levar a cabo artéllanse en tres bloques diferentes, pero complementarios. Un deles consiste na compleción do corpus con novo material para cubrir cronoloxicamente os últimos períodos; outro, que se prolonga ao longo do tempo, ten que ver coa mellora da etiquetaxe automática do corpus e, por último, o terceiro atinxe á aplicación de recuperación de información e á mellora das funcionalidades existentes ou creación de novas ferramentas.

A continuación desenvolvemos máis polo miúdo cada unha destas áreas.

1) Por unha banda, ocuparémonos da introdución de novos documentos para completar o lustro 2015-2019 e inicia-lo 2020-2024. A selección farase tendo en conta os criterios clasificatorios do corpus (tipo de documento, área temática, xénero etc.) e prestarase



especial atención á variedade, tanto no referente a tipos de documentos e autores como a editoriais. Á parte de xornais, revistas, novelas e todos os restantes tipos de documentos recollidos ata o de agora, prevese a incorporación para o rexistro oral de material extraído de vídeos da plataforma YouTube.

2) Por outra banda, a mellora da etiquetaxe automática debe ser unha constante no desenvolvemento do proxecto, tanto para corrixir análises incorrectas como para facilitar o recoñecemento de unidades presentes no corpus pero que, por ser neoloxismos, variantes dialectais, formas contaminadas por interferencias, acentuación anómala etc., non se recollían nas fontes léxicas (VOLG e DRAG, principalmente) incorporadas ó sistema XIADA.

Neste sentido, cómpre atender varias fronteas:

- Adapta-lo etiquetador XIADA para facer un adestramento específico para o corpus CORGA, o que permitirá incrementar a relevancia do corpus de adestramento e reducir a influencia negativa do lexicón na estatística. Con isto preténdese mellorar a taxa de acerto xeral do etiquetador.
- Desenvolver un avaliador para o etiquetador XIADA, que permitirá, por un lado, saber a taxa de acerto real do etiquetador e, polo outro, detectar prematuramente os cambios que se introducen na etiquetaxe cada vez que se desenvolven melloras. Así mesmo, preténdese desambiguar manualmente unha selección de 10.000 elementos gramaticais extraídos aleatoriamente de textos escritos despois do 2010 para avaliala taxa de acerto e, polo outro, doutra cantidade semellante de textos de calquera período, incluído o prenормativo, para ver a taxa real en todo o corpus.
- Ampliar o lexicón coas entradas máis frecuentes procedentes do arquivo de formas sen lema extractado da versión 4.0 do CORGA.
- Modificar as entradas e grupos de derivación na base de datos do lexicón para dar cabida á variedade acentual nos casos de ditongo decrecente (industriáis, locais, papéis, nivéis, españoís...) ou en sufixos prototípicos (contrário, reaccionário...)
- Reasignar os elementos clasificados inicialmente por motivos computacionais nalgún dos subtipos de Locución (conxuntiva, adverbial e prepositiva) á clase de palabra correspondente (Conxunción, Adverbio e Preposición).
- Lematizar as formas con gheada e intentar aplicalo tamén ás formas con seseo.
- Lematizar automaticamente aqueles prefixos (*extra-*, *inter-*, *intra-*, *meta-*, *post-*, *ultra...*) e sufixos que permitan realizala con garantías (*-dade*, *-ismo*, *-ista*, *-mento...*)
- Mellora-lo recoñecemento de numerais (expresados moitas veces en unidades contractas xunto cun símbolo -euro, porcentaxe, grao etc.-) e dos propios, tanto a inicio



de secuencia como daqueles propios non prototípicos que só amosan a inicial en maiúscula (títulos de libros, programas etc.).

- Crear regras lingüísticas que corrixan análises sistemáticas erróneas fundamentadas na estatística.

3) Por último, no que atangue ás novas funcionalidades, pretendemos:

- Amplia-las posibilidades do selector *Buscar en* para recoller elementos que afectan ao nivel da palabra: alongamento, texto riscado etc.
- Elimina-los espazos introducidos na tokenización cando se amosa o texto orixinal dunha secuencia no contexto da aplicación de buscas do CORGA, de xeito que se poida ve-lo texto orixinal real no contexto.
- Procurar tendo en conta o sexo do interlocutor/falante. Este filtro aplicaríase ao atributo *Interlocutor*, equivalente nas transcricións a *Falante*, e obriga a engadir un valor máis, respecto ao sexo, para o valor do *Interlocutor* na parte escrita do corpus. Igual que pasou co valor *sexo* para os autores, esta mellora implica volver sobre todos os textos que presentan *interlocutor*, para caracterizalo con respecto ao sexo: entrevistas, coloquios, obras de teatro, guións...
- Procurar por *Título*. *Título* corresponderíase con cada arquivo dos que temos nas distintas carpetas de *Corga\_xml\_corrixidos*: 1 novela, 1 colección de ensaios ou 1 colección de relatos, sexan do mesmo autor ou de varios, 1 xornal, 1 revista, 1 blog, 1 guión etc.
- Incorporar o índice de información mútua nas buscas de coaparacións e estudar posibilidades para habilitar a obtención de coaparicións sen ter en conta a posición do colocativo e o colocado.
- Estudar a posibilidade de habilitar un sistema de carga diferida de autores e títulos no formulario do CORGA para que a experiencia de usuario sexa mellor.
- Dar cabida ao campo semántico. Trátase de implementar un campo máis no lexicón que proporcione información semántica clasificando algunhas das entradas segundo os centros de interese recollidos no léxico dispoñible para o galego ([https://gramatica.usc.es/proxectos/dispongale/?lang=es&id=lexico\\_galego](https://gramatica.usc.es/proxectos/dispongale/?lang=es&id=lexico_galego)). O obxectivo final é aplicar esta información na plataforma de consulta e facilitar que un usuario extraia todas as ocorrencias dos lemas clasificados dentro dun mesmo campo semántico.

Desenvolver un flexionador. Existe un flexionador nominal dispoñible para o galego (<http://ilg.usc.gal/flexionador/>), porén non dá conta do uso senón só da norma. Preténdese proporcionar esta funcionalidade para calquera lema segundo a información que achega o CORGA e maila que proporciona o lexicón, de xeito que o



usuario poida ver, por exemplo, os 2 paradigmas existentes para o lema *xerente*, segundo presente ou non formas con desinencia específica para o feminino e cales desas formas están presentes no corpus. Ademais do paradigma flexivo, achegaríase información sobre a normatividade do lema e das formas e sinalaríase a fonte da que se toma o lema.

## PROXECTO BIBLIOGRAFÍA INFORMATIZADA DA LINGUA GALEGA (BILEGA)

BILEGA é unha base de datos que contén información bibliográfica sobre os traballos de investigación, divulgación ou opinión dedicados ao idioma galego nas diferentes fases do seu desenvolvemento histórico. Os rexistros da base de datos ofrecen a historia editorial completa dos documentos referidos, así como información verbo do seu contido e sobre as recensións e notas de publicación a que deron lugar. Ademais de edicións impresas, as referencias bibliográficas inclúen edicións dixitais, nomeadamente as difundidas a través da Internet para as que se achega, sempre que for posible, a URL na que atopar o documento en cuestión -tanto do texto coma das recensións-. Está constituída por dúas táboas: BILEGA (que contén as entradas bibliográficas) e REVISTAS (que contén os datos das publicacións periódicas citadas na primeira).

A versión 3.0, presentada o 30 de agosto de 2019, está dispoñible para a súa consulta en liña no enderezo <https://www.cirp.es/w3/bdo/bdo-bilega3.html>. No momento actual (3 de decembro de 2021) conta con 21.916 rexistros bibliográficos -localizados na táboa BILEGA- e 1486 rexistros de revistas -que se poden atopar na táboa REVISTAS-.

No cuadrenio 2022-2025 realizaranse algunhas melloras na páxina utilizada para xestionar o recurso:

- 1- Introducir a opción de busca por Datos adicionais
- 2- Vincular as dúas táboas (REVISTAS e BILEGA).
- 3- Permitir buscas por Actualización.

Canto á base que está a disposición dos usuarios, están pendentes de implementar na interface as seguintes melloras:

- 1- Implementación de operadores que permitan arrequer as buscas pola ID do rexistro.
- 2- Implementar a posibilidade de combinar os valores de busca nos campos Enfoque e Período.



- 3- Implementar a posibilidade de facer buscas pola lingua de textos traducidos.
- 4- Activación de operadores de comparación para as buscas complexas.
- 5- Elaboración dunha guía de uso
- 6- Revisión e actualización constante dos documentos que conforman a sección de Info(rmación) adicional.

En canto ao contido está previsto:

1- Continuar coa actualización e revisión dos rexistros, xa que constitúe unha das chaves para que BILEGA siga na vangarda das bases de datos bibliográficas. É fundamental o repaso continuado para garantir a corrección da ficha e da información ofrecida aos usuarios, asegurando, por exemplo, que as URLs, a norma empregada, os datos achegados etc. son correctos e están totalmente actualizados. Ademais desta supervisión constante, tamén está programada a revisión dos rexistros e a elaboración do resumo daqueles que aínda non o teñan ou nos que se considere que poden ser máis completos.

2- Incrementar o volume textual da base: isto implica coñecer, localizar e seleccionar os diferentes traballos que van ser incorporados á base de datos no próximo cuadrienio. Estes serán sometidos a un proceso de lectura a partir do que se elabora a síntese do seu contido e a ficha correspondente, creando o arquivo que logo se incorpora como novo rexistro á BILEGA. Estes textos a incrementar -actas de congresos, artigos, monografías, teses de doutoramento etc.- corresponden ás publicacións que se produzan nos anos 2022 ao 2025.

Arestora, a base de datos alberga case 22.000 rexistros, que conteñen 57.483 asentos bibliográficos para os que se implementaron máis de 180.395 valores de busca nos campos tema, enfoque, período, variedade de lingua, ámbito, tipo de traballo, lingua do documento e na sección de identificadores do campo documento. Na inclusión de novos traballos ou de novas edicións de traballos xa referenciados na BILEGA, considérase fundamental seguir a explorar a rede Internet para identificar novas edicións dixitais en liña, das que arestora xa están incluídas as referencias dunhas 25.000.

### **ACCIÓN 1.1.3**

#### **PROXECTO CORPUS DOCUMENTALE LATINUM GALLAECIAE (CODOLGA)**

O *Corpus Documentale Latinum Gallaeciae* é un corpus textual con documentación latina medieval relacionada con Galicia producida entre os séculos VI e XV, ambos inclusive. O obxectivo principal do proxecto é ofrecer ao persoal investigador e ao



público xeral unha aplicación informática (<http://corpus.cirp.gal/codolga>) na que se van recompilando os documentos en latín referidos a Galicia do período cronolóxico mencionado.

Na última versión publicada de CODOLGA, a número 17, están incluídos 18.624 documentos

Está ferramenta informática permite buscas de léxico de acordo con varios parámetros de selección e ofrece información varia sobre cada un dos documentos. A aplicación conta tamén cunha sección na que se publican outros textos latinos medievais interesantes polas súas referencias a Galicia e estudos singulares do ámbito do proxecto. Resulta, polo tanto, un instrumento moi útil e imprescindible para o estudo e o coñecemento da lingua e da historia social, política, económica e cultural de Galicia na Idade Media.

Os obxectivos do proxecto e as tarefas correspondentes previstas para o período 2022-2025 son os seguintes:

1.- Continuar coa preparación e a incorporación á base de datos de novos documentos que xa estean ou vaian sendo publicados, especialmente de coleccións diplomáticas importantes aínda non incluídas no corpus. Un elemento de referencia continúa sendo o catálogo CODIPHIS (J. A. García de Cortázar, J. A. Munita e L. J. Fortún, *Codiphis. Catálogo de coleccións diplomáticas hispano-lusas de época medieval*, 2 vols. Santander: Fundación Marcelino Botín, 1999). A previsión é continuar coa incorporación de ao redor de 1.200 documentos por ano.

2.- Seguir coa progresiva renovación e mellora que se vén facendo da aplicación informática. Nas últimas versións foise renovando a súa interface e ampliando as posibilidades de busca de información e de presentación de resultados. Continuarase nesa mesma liña, para permitir consultar datos aínda non dispoñibles, como os relativos a arquivos e bibliotecas depositarios dos documentos, ou para mellorar a información de campos xa dispoñibles, como a clara identificación das distintas edicións dalgúns documentos.

3.- Continuar desenvolvendo a colaboración de CODOLGA cos proxectos similares creados a imitación súa, CODOLCAT (*Corpus Documentale Latinum Cataloniae*), CODOLLEG (*Corpus Documentale Latinum Legionis*), CODOLVA (*Corpus Documentale Latinum Valencie*), CODOLPOR (*Corpus Documentale Latinum Portucalense*), e fomentar na medida do posible a creación doutros proxectos que cubran áreas de Hispania aínda non atendidas. Resultado desa colaboración é a rede CODOLHisp (*Corpus Documentale Latinum Hispaniarum*) e a aplicación informática resultante, que permite consultar de forma simultánea e combinada os distintos *corpora* ([codolhisp.imf.csic.es/codolhisp/](http://codolhisp.imf.csic.es/codolhisp/)). Será necesario completar e mellorar tamén esa ferramenta.



4.- Crear un sistema de busca conxunta de CODOLGA e do *Corpus do Galego-Portugués Antigo* (CGPA: <http://ilq.usc.gal/cgpa/index.php>), que permita unir ambos recursos e presentar ás persoas interesadas a posibilidade de consultar de forma transversal toda a documentación relativa á Galicia da Idade Media tanto en lingua latina como en lingua romance.

5.- Ampliar a biblioteca de CODOLGA con copias dixitais de documentos e outras fontes medievais obxecto do proxecto, para facilitar a revisión da documentación publicada e posibilitar incluso a edición propia dalgúns importantes documentos aínda inéditos.

## 2. ÁREA LITERATURA

### 2.1. LITERATURA MEDIEVAL

#### ACCIÓN 2.1.1

#### **PROXECTO ARQUIVO GALICIA MEDIEVAL (ARGAMED)**

Este Arquivo constitúe unha referencia internacional para os investigadores da literatura galega medieval, principalmente polas prestacións que ofrecen as súas bases de datos, *MedDB* (Base de datos da Lírica Profana Galego-Portuguesa), *BiRMED* (Bibliografía de Referencia do Arquivo Galicia Medieval) e *PalMed* (Base de datos paleográfica da lírica galego-portuguesa), que seguen a ser un dos obxectivos primordiais do proxecto.

Para os próximos anos, continuarase coa serie de publicacións dixitais en aberto *ARGAMED*, que xa puxo en rede tres números (<http://www.cirp.gal/w3/argamed/>). Este recurso non exclúe que algúns traballos derivados do macroproxecto sigan vendo a luz en forma impresa (como acontece, por exemplo, co volume *Alfonso X e Galicia*, publicado en 2021 para conmemorar o oitavo centenario do nacemento do Rei Sabio), pero a finalidade fundamental desta serie é, ademais de axilizar os procesos de publicación e optimizar os recursos económicos, dar a maior visibilidade e capacidade de acceso aos estudosos a todos aqueles recursos e contidos que constitúen o núcleo das investigacións que se están a realizar no macroproxecto.

Para o período 2022-2025, ademais de manter as liñas de traballo habituais, está prevista a realización dalgún congreso ou seminario relacionado coas liñas de investigación que se integran en ARGAMED.

Forman parte deste macroproxecto (**Argamed**) catro proxectos diferentes, que teñen como eixo de estudo as fontes primarias e secundarias da literatura galega medieval.



## PROXECTO LÍRICA GALEGO-PORTUGUESA

Nacido co obxectivo de proporcionar instrumentos de traballo fiables e de fácil acceso para os estudosos da lírica galega-portuguesa, na súa primeira fase recompilouse todo o corpus (ca. 1700 cantigas), e, posteriormente, procedeuse a preparar o material para convertelo nunha base de datos que se puxo en rede en 1998 e que, actualmente, vai pola súa terceira versión: (MedDB v.3, 2016; última actualización: Versión 3.11, 24/09/2021). Esta base de datos está en proceso permanente de revisión, actualización (con particular atención aos traballos e edicións publicados polos especialistas) e mellora no seu rendemento, para que siga sendo unha referencia internacional no seu ámbito de estudo. Unha das últimas novidades importantes foi a incorporación dun LEMARIO, que axiliza as buscas lexicográficas ao reducir as diferentes formas e grafías rexistradas nas edicións a un lema común, así como a integración de documentación adicional de carácter relevante (en formato pdf) como complemento ás utilidades de MedDB.

Ao mesmo tempo, deseñouse unha base de datos para incorporar a bibliografía sobre a nosa lírica trobadoresca, coa posibilidade de realizar buscas temáticas mediante o recurso a palabras chave e coa inclusión dun índice de citas, que deu lugar á actual BIRMED (véxase máis abaixo).

Pero, sen dúbida, a principal achega recente deste proxecto foi a posta en marcha, en decembro de 2020, dunha nova ferramenta, *PalMed*, totalmente innovadora nos ambientes especializados, e que pretende converterse nun referente de prestixio internacional nos estudos de lírica galego-portuguesa: trátase dunha base de datos que recolle de forma conservadora as transcripcións paleográficas de todos os testemuños da lírica trobadoresca (uns 3.100 documentos en XML, con máis de 62.000 versos), con todo un elenco de marcas que permiten levar a cabo buscas combinadas de moitos tipos, non só do texto das cantigas en si mesmo, senón tamén de todos os paratextos, marcas e escrituras anexas que se poden ver nos cancioneiros.

Nos próximos catro anos, o proxecto centrará os seus esforzos en dúas liñas principais:

1.- Por unha banda, **MedDB** ([www.cirp.gal/meddeb](http://www.cirp.gal/meddeb)) experimentará unha modificación completa dos contidos correspondentes a “trobadores”, porque cómpre dar conta de todos aqueles datos inéditos que os investigadores sacaron (e sacarán) á luz nestes últimos anos. Ademais, tamén actualizará os textos, xa que están aparecendo novas edicións críticas que é imprescindible revisar e contrastar coas que se ofrecen como de referencia na base de datos, así como engadir a información relativa a todas elas. Neste sentido, o equipo do proxecto traballará para introducir enlaces que conecten directamente a información resumida que se ofrece sobre a bibliografía coas entradas de BIRMED e optimizar os contidos de varios campos para que o recurso se manteña permanentemente actualizado.



2.- *PalMed* ([www.cirp.gal/palmed](http://www.cirp.gal/palmed)) será sometida a unha revisión exhaustiva que permita corrixir pequenos erros que se deslizaron involuntariamente no seu complexo proceso de elaboración, pero os esforzos nesta ferramenta estarán centrados, sobre todo, no seu enriquecemento a nivel lingüístico coa lematización e etiquetaxe morfolóxica do corpus. Para abordar tal obxectivo, é necesario descender na estrutura organizativa dos XML do nivel verso ao nivel de la palabra, un proceso que supón un complexo e arduo traballo, posto que resulta imprescindible realizar unha serie de tarefas para regularizar o contido lingüístico dos manuscritos.

A primeira delas consistirá na revisión de todos os documentos que configuran *PalMed* para marcar adecuadamente as unións e separacións de palabras, coa finalidade de xerar un texto sobre o que poder "tokenizar" o corpus, é dicir, dividir os versos (ou as liñas) en unidades categoriais. Unha vez finalizada esta fase, elaborárase o elenco das "formas paleográficas" existentes e aplicaráselles unha serie de normas que as regularicen ortograficamente, pero que manteñan as propiedades fonéticas da lingua medieval.

O seguinte paso será o deseño dunha codificación que estableza as etiquetas e os atributos oportunos para clasificar gramaticalmente cada unha das palabras individualizadas na fase de traballo previa. Para realizar este labor, seguiranse as recomendacións do grupo EAGLES e das gramáticas históricas e terase como referente o protocolo de marcado fixado por TEI. A etiquetaxe morfosintáctica levarase a cabo de xeito "semiautomático", é dicir, empregárase un lematizador automático (XIADA), que deberá ser adaptado ás características do corpus do proxecto e que será adestrado mediante unha serie de documentos etiquetados manualmente.

A base de datos resultante enriquecerá *PalMed* para abordar ambiciosas investigacións lingüísticas, que resultarían moi complexas sen unha ferramenta deste tipo. Así, o recurso permitirá afondar no coñecemento do sistema gramatical da lingua da produción lírica nas diversas fases da tradición manuscrita; detectar de xeito automático os fenómenos de variación lingüística que se producen na tradición escrita; verificar a presenza de certos fenómenos gramaticais en sectores concretos dos cancioneros, e vincular a execución dos manuscritos existentes (ou a dos seus modelos) a zonas territoriais concretas mediante a comparación das variantes detectadas coas presentes nos sistemas dialectais do galego e do portugués.

## PROXECTO BIBLIOGRAFÍA DE REFERENCIA DA LITERATURA GALEGA MEDIEVAL

BiRMED (Bibliografía de Referencia do Arquivo Galicia Medieval) é unha ferramenta dixital de acceso aberto destinada a facilitar a todo investigador en literatura medieval galego-portuguesa a localización de referencias bibliográficas especializadas nesta



área de coñecemento. Simultaneamente, os investigadores encontrarán tamén en BiRMED unha plataforma que permitirá dar a coñecer as súas publicacións entre un público especializado ou en formación e que lle poderá proporcionar información valiosa sobre o seu impacto e número de citas. Atendendo a isto e coa vontade de facer avanzar o proxecto, as tarefas planeadas para os próximos catro anos son as seguintes:

- Actualización da versión publicada: cara á versión 3.0.
- Serán incorporadas novas opcións de consulta concibidas para que esta poida ser moito máis intuitiva e áxil para o usuario.
- Avanzar cara a unha consulta dinámica e fluída da información a través da introdución de hipervínculos que permitan conectar informacións dentro ou fóra da base.
- Revisión e actualización do elenco das palabras clave utilizadas ata o momento coa finalidade de responder máis adecuadamente a novas tendencias e intereses na investigación actual.
- Incorporación de novas tipoloxías de traballo non incluídas ata o momento, como bases de datos especializadas na literatura medieval galego-portuguesa e recensións de edicións críticas e estudos monográficos. Estas últimas aparecerán directamente ligadas ao traballo obxecto de revisión, de tal maneira que se facilite aos investigadores-autores unha fonte de datos sobre o impacto e calidade das súas publicacións, sumándose ao servizo de control de citas que na actualidade xa ofrece BiRMED.
- Incorporación de novos elementos no interior de tipoloxías novas e outras xa existentes coa finalidade de actualizar o tipo de informacións das publicacións en formato dixital. A conveniencia de incluír o DOI e/ou outro tipo de identificadores dixitais de carácter estable –hoxe en día tan útiles como imprescindibles nunha referencia bibliográfica dixital completa– fará precisa, ademais, unha importante tarefa de revisión e actualización de rexistros xa existentes, pero co que se espera mellorar substancialmente as posibilidades da base de datos.
- Renovación do deseño gráfico web, se ben respectando as pautas estéticas marcadas polo CRPIH, e da imaxe do proxecto, o que eventualmente poderá contribuír a estimular o seu uso entre novos públicos.

## PROXECTO PROSA LITERARIA MEDIEVAL

Este proxecto pretende conservar e difundir a través de edicións impresas e electrónicas os textos en prosa escritos en galego durante a Idade Media. Ten como obxectivo prioritario a edición, estudo, caracterización e anotación das obras medievais que integran o xénero que dá título ao proxecto.



No ano 2020 levouse a cabo o manual "Guía para o estudo da prosa galega medieval" como resultado directo do proxecto (número 3 de ArGaMed). O libro (en acceso aberto) recolle información fundamental do corpus da prosa galega medieval dando a coñecer as obras literarias que o conforman dunha forma seria e sinxela. A pesar de formar parte do noso patrimonio cultural, este conxunto textual está bastante agachado para o público pola diversidade de tipoloxías das obras e polos problemas de edición que algunhas delas presentan. O manual consolidouse como bibliografía de referencia nos estudos da prosa medieval.

Nos próximos dous anos o proxecto traballará no deseño e posta en marcha dun novo recurso en rede que se engadirá aos existentes na web oficial do centro e que será de grande axuda para os investigadores iniciados ou que se acheguen por primeira vez a algunha obra (ou a todo o repertorio) prosístico galego.

A base de datos será de tipo conceptual e recollerá todas as obras (*Crónica de Castela; Crónica de Santa María de Iria; Crónica Troiana; Crónica Xeral; Crónica Xeral de 1404; Foro Real; Flores de Dereito; Historia Troiana; Libro de Tristán; Miragres de Santiago; Tratado de Albeitería e Xeral Estoria*) facilitando buscas por diferentes ámbitos: antropónimos (diferenciando personaxe de ficción e personaxe reais), topónimos (maiores e menores), grupos sociais varios, institución, acontecementos sociais, acontecementos históricos, descrición e comportamento do individuo (enfermidade, hábitos, etc.), dos animais, dos elementos da natureza, das construción arquitectónicas.

Tralo baleirado dos textos en linguaxe XML etiquetarse o corpus para referenciar os ítems indicados e outros que, a medida que se analizan e traballan as obras, se consideren relevantes. Este labor resultará bastante extenso no tempo, por ter que facer de xeito manual os etiquetados e por non ter os textos en arquivos dixitalizados. Ademais, na cabeceira de cada texto incluírase información bibliográfica sobre a obra, dando conta das edicións máis pertinentes e dos principais estudos.

A base de datos como ferramenta de traballo será amosada en reunións científicas para difundila entre o público interesado.

Nos últimos anos (2024 e 2025), comezarase coa realización de edicións propias de textos que precisen novas lecturas. Neste sentido unha das obras que necesitaría unha revisión codicolóxica sería os *Miragres de Santiago*, un dos libros de referencia do noso acervo cultural e que non conta cunha edición de fácil acceso para un público non iniciado. A edición que se ten como obxectivo realizar orientaríase á elaboración dun texto rigoroso que, como transferencia de resultado, dese a coñecer un dos libros esenciais da historia literaria galega.

## PROXECTO CANTIGAS DE SANTA MARÍA



O Proxecto Cantigas de Santa María incorpora o estudo e edición da parte de poesía relixiosa afonsina ao corpus da poesía profana galego-portuguesa, implementando dese xeito tanto a base de datos MedDB, como complemento á base de datos bibliográfica *BirMed* cos aportes específicos relativos ás *Cantigas de Santa María* e a outra produción afonsina de interese. O proxecto, en última instancia, aspira a atender tamén os outros dous aspectos fundamentais desta magna obra literaria: o estudo musicolóxico e o iconográfico. Ata agora non existe unha aproximación exhaustiva a este monumento literario baixo esta triple perspectiva, que se ve favorecida pola recente dixitalización dos seus cancioneros, posta en liña no bienio anterior.

Polo de agora, está xa a disposición do público a transcripción paleográfica do Códice de Florencia (B.R. 20), que xa está en rede (<http://www.cirp.gal/w3/argamed/>), constituíndo o primeiro número do recurso *ARGAMED*. Con esta edición déronse os primeiros pasos para complementar, así mesmo, a base de datos paleográfica PalMed, que recolle as transcripcións en XML dos textos da lírica trobadoresca. A maiores, etiquetáronse os textos deste códice en XML e rematouse a súa versión en HTML para poder poñela a disposición dos usuarios nun formato máis áxil e sinxelo para a súa consulta tan pronto como se nos habilite un espazo web axeitado. Co etiquetado en XML dos textos das CSM (deste e dos demais códices que se etiquetarán no futuro) poderanse facer buscas similares ás que se poden facer actualmente en PalMed, ampliando dese xeito o repertorio dos textos da nosa poesía medieval e o correspondente campo para un estudo similar.

Para finais de 2025 preténdese:

- Ter a transcripción paleográfica do Códice dos Músicos (RBME B-I-2), que é o códice que contén toda a colección, con máis de catrocentas cantigas. O modelo a seguir é a transcripción paleográfica do Códice de Florencia (B.R. 20), sinalada arriba. Na medida do posible, abordarase tamén a transcripción do Códice Rico do Escorial (RBME T-I-1).
- Completar os criterios de etiquetado dos textos en XML para asegurar o crecemento e/ou futuras asociacións do proxecto.
- Completar os modelos de estruturas para o marcado dixital do texto en XML sobre os textos na súa transcripción paleográfica.
- Preparación dos textos do Códice Rico do Escorial (RBME T-I-1) en XML tal como se fixo coa versión paleográfica do Códice de Florencia (B.R. 20).
- Poñer a disposición do público unha versión en formato HTML da transcripción paleográfica dixital do Códice de Florencia (B.R. 20) e, na medida do posible, do Códice Rico do Escorial (RBME T-I-1). Este apartado depende de que se nos asigne un espazo web na páxina do CIRPH para poder presentar os textos.



- Avanzar na etiquetaxe dos textos en XML coa intención de que a actual ferramenta que sacou á luz PalMed integre tamén (ou se poida combinar dalgún xeito que se deberá explorar) os textos da poesía relixiosa que se estaba compoñendo ó mesmo tempo e no mesmo centro de produción de boa parte da poesía profana. Con esta extensión ofrecerase unha visión máis completa da poesía medieval galego-portuguesa. As buscas combinadas poderán complementarse coa escrita das CSM.
- Propoñer as edicións críticas do maior número de textos posible das CSM, acompañadas da respectiva edición modernizada do texto e da súa tradución a algunha outra lingua.
- Presentar en XML algunhas das edicións anteriormente mencionadas.
- Na medida en que os recursos informáticos do Centro o permitan, avanzar no deseño da ferramenta informática que combine a información filolóxica e outra de carácter máis xeral que temos recollida ó longo dos últimos anos. A posta a disposición do público só será posible cando se completen as tarefas anteriores e se acabe de deseñar a base de datos que inclúa a información posta a disposición dos servizos informáticos do Centro.

## **2.2. LITERATURA CONTEMPORÁNEA**

### **ACCIÓN 2.2.1**

#### **PROXECTO DICIONARIOS DE LITERATURA**

O proxecto Dicionario de Literatura consiste na elaboración dun gran dicionario xeral de termos literarios, que inclúe voces do campo da Retórica, Métrica, Semiótica, Poética, Narratoloxía, Literatura comparada, Estudos culturais, Estudos de Performance, Ciberliteratura, Hermenéutica, e outras disciplinas de coñecemento que interseccionen coa literatura, prestándolle unha atención detallada aos novos paradigmas teóricos, en especial aqueles que teñen que ver coa nova realidade tecnolóxica e dixital, e cos cambios a nivel sociocultural, relacionados coa produción e recepción de literatura e outras artes.

Ao tempo, aínda que neste Dicionario xa son integrados exemplos e referencias concretas da literatura galega, en relación coa literatura universal, mantense o proxecto de Dicionario vinculado a este campo literario concreto. Fronte á intención inicial de elaborar un gran dicionario xeral de literatura galega, optouse por parcelar máis o campo e comezar a pensar en dicionarios para cada unha das etapas asentadas pola periodización establecida pola historiografía contemporánea da literatura galega: Rexurdimento, Irmandades da Fala e Nós, Ditadura, Restauración democrática, etc. (1).



Despois de rematarse a redacción, está a traballarse na revisión do Dicionario sistémico-literario do Rexurdimento, e ao tempo, a elaborarse unha base de datos que permita a súa consulta a través de Internet.

Unha vez rematada a edición do terceiro tomo do Dicionario de Literatura (I-M) no 2021, e implementadas certas melloras na páxina web, particularmente no tocante ao deseño e visibilización da páxina e ao reconto estatístico das visitas e actualizacións, o seguinte período dedicárase á redacción de voces para o cuarto volume (N-P), á implementación de melloras, en paralelo, na base de datos, de acordo coa dispoñibilidade de traballo dos bolseiros contratados.

### **PROXECTO BIBLIOGRAFÍA DA LITERATURA GALEGA**

Este proxecto pretende reunir do xeito máis exhaustivo posible toda a información relacionada coa literatura galega contemporánea, dende 1800 até a actualidade, para a súa posterior publicación. A orientación actual está na dirección de confluír co proxecto *Informes de literatura*. Preténdese continuar cos labores de selección, compilación e revisión da base de datos. De momento, só é accesible para os distintos proxectos do Centro, pero estase a traballar para ampliar a súa dispoñibilidade.

### **PROXECTO INFORME ANUAL DE LITERATURA GALEGA**

O proxecto recolle nun informe anual, toda a produción literaria en galego (narrativa, poesía, teatro, ensaio, crítica; literatura infantil e xuvenil; literatura de transmisión oral; Literatura Medieval; revistas; premios) que se rexistre en libros, revistas ou prensa (tanto en papel coma electrónica). Este informe elabórase dende o ano 1995 e pretende apoiar a investigación, difundir esta información e observar a construción do sistema literario galego dende todos os ámbitos da comunicación literaria.

No traballo do informe recompílanse todas as informacións e reflexións que, atinxindo directa ou indirectamente ao ámbito da Literatura galega, apareceron por unha parte en soporte impreso durante o ano e, dentro deste tipo de soporte, nas súas máis diversas manifestacións: monografías, libros colectivos, revistas, xornais etc.

Ofrécese un comentario descritivo, non valorativo nin crítico, de cada unha das entradas rexistradas, é dicir, tanto dos libros relacionados coa literatura que viron a luz ao longo do ano como das recensións que sobre eles apareceron nas revistas e nos xornais durante ese mesmo período de tempo.

Os informes de Literatura que se referirán aso anos 2022, 2023, 2024 e 2025, editaranse exclusivamente como publicación electrónica en liña.



## **PROXECTO DE INVESTIGACIÓN EN LITERATURA INFANTIL E XUENIL**

Este proxecto, creado a finais do ano 2012, acolle os traballos que anualmente realizan sobre Literatura Infantil e Xuvenil galega e doutros ámbitos lingüísticos a Rede Temática "As Literaturas Infantís e Xuvenís do Marco Ibérico e Iberoamericano" (LIXMI); a Asociación de Investigadores en Literatura Infantil e X/Juvenil (ELOS), ELOS-GALICIA. Asociación de investigación en Literatura Infantil e Xuvenil e o Grupo de Investigación da Universidade de Santiago de Compostela LITER2I (GI-1819) que conta entre as súas liñas de investigación "Literatura Infantil e Xuvenil. Investigacións literarias, artísticas, interculturais e educativas".

Deste proxecto, naceu no CRPIH a Rede Ibérica de Literatura Infantil e Xuvenil, da que forman parte especialistas de toda a xeografía lingüística da península e que celebran un congreso todos os anos no CRPIH.

Para o período comprendido entre 2022 e o 2025, levaranse a cabo tarefas de difusión relacionadas directamente co ámbito da literatura infantil e xuvenil do marco ibérico e iberoamericano e organizaranse seminarios internacionais sobre o a mesma.

Ademais, participárase de maneira activa en actividades científicas e publicacións resultantes dos traballos de investigación levados a cabo.

### **3. ÁREA DE ANTROPOLOXÍA**

#### **ACCIÓN 3.1**

#### **PROXECTO DIMENSIÓNS DA IDENTIDADE COLECTIVA DE GALICIA**

Con motivo da conmemoración do corenta aniversario das primeiras eleccións autonómicas de Galicia, nos próximos meses presentarase o Atlas electoral de Galicia, no que se recollen os datos e mapas electorais de todos os procesos autonómicos, xunto cos índices electorais de cada proceso, e faise unha análise de como, dende unha visión neoinstitucionalista e construtivista, as institucións e a competición electorais apuntan a construción da identidade galega.

Os próximos anos, o estudo das dimensións da identidade centrarase nos factores emocionais que están detrás da construción das identidades, tanto dende a perspectiva dimensional das emocións coma dende a perspectiva discreta.

Este intento de tratar agora as dimensións emocionais da identidade, lévanos a formular catro obxectivos xerais para os próximos catro anos:



1. Mostrar a presenza de factores emocionais asociados á construción da identidade colectiva de Galicia.
2. Dende unha perspectiva emocional, analizar qué tipo de emocións están máis presentes nos diversos espazos e modelos de construción da identidade.
3. Dende unha perspectiva discreta, detectar cales son as emocións máis presentes na identidade galega.
4. Delimitar a existencia dun réxime emocional colectivo e das diferentes arquitecturas emocionais asociadas a ese réxime.
5. Amosas a existencia dunha semántica emocional asociada á construción deste réxime emocional e de semánticas particulares que inciden na construción das diferentes arquitecturas emocionais.



**ANEXO II**

O custo total do convenio, coas achegas correspondentes de ambas as partes asinantes, distribuirase anualmente do seguinte xeito:

Accións	Custo total anual	Horas estimadas anuais	Achega Consellería anual	Contribución USC anual
1.1.1. Traballos de continuación dos proxectos Recursos para o desenvolvemento das tecnoloxías da fala, TERMIGAL e ALIR	41.300,00 €	1380	38.500,00 €	2.800,00 €
1.1.2. Traballos de continuación dos proxectos BILEGA, CORGA e XIADA	77.300,00 €	2570	72.600,00 €	4.700,00 €
1.1.3. Traballos de continuación do proxecto CODOLGA	15.960,00 €	532	11.960,00 €	4.000,00 €
2.1.1. Traballos de continuación no macroproxecto ARGAMED (Lírica galego-portuguesa; Prosa literaria; Cantigas de Santa María; BirMED)	38.300,00 €	1280	33.400,00 €	4.900,00 €
2.2.1. Traballos de continuación dos proxectos Dicionarios de literatura; Bibliografía da literatura galega; Informe anual da literatura galega; Literatura infantil e xuvenil	7.440,00 €	248	5.200,00 €	2.240,00 €
3.1. Traballos de continuación do proxecto Identidade colectiva de Galicia	4.500,00 €	150	0,00 €	4.500,00 €
<b>TOTAL ANUAL</b>	<b>184.800,00 €</b>	<b>6160</b>	<b>161.660,00 €</b>	<b>23.140,00 €</b>
<b>TOTAL CONVENIO (CUATRIANUAL)</b>	<b>739.200,00 €</b>	<b>24640</b>	<b>646.640,00 €</b>	<b>92.560,00 €</b>

