

CONVENIO DE COLABORACIÓN ENTRE A CONSELLERÍA DE CULTURA, EDUCACIÓN E UNIVERSIDADE, A AXENCIA PARA A MODERNIZACIÓN TECNOLÓXICA DE GALICIA E A UNIVERSIDADE DE SANTIAGO DE COMPOSTELA PARA DESENVOLVER O PROXECTO NÓS: O GALEGO NA SOCIEDADE E ECONOMÍA DA INTELIXENCIA ARTIFICIAL

En Santiago de Compostela, 29 de abril de 2022

REÚNENSE

Don Román 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 11/1988, do 20 de outubro e pola Lei 12/2007, do 27 de xullo, de acordo co seu nomeamento polo Decreto 112/2020, do 6 de setembro e no uso das atribucións que lle confire o Decreto 198/2020, do 20 de novembro, publicado no DOG número 242, do 1 de decembro, no que se establece a estrutura orgánica da Consellería de Cultura, Educación e Universidade e do acordo do Consello da Xunta de Galicia do 27 de maio de 1991, sobre convenios de cooperación con outros entes públicos e de colaboración con particulares, publicado pola Resolución da extinta Consellería de Economía e Facenda do 8 de abril de 1991 (DOG núm 82, do 30 de abril).

Dona Mar Pereira Álvarez, directora da Axencia para a Modernización Tecnolóxica de Galicia, (en diante Amtega) cargo para o que foi nomeada polo Consello da Xunta na súa reunión do 15 de marzo de 2012, en función das atribucións que lle confire o Decreto 252/2011, do 15 de decembro, polo que se crea a Axencia para a Modernización Tecnolóxica de Galicia e se aproban os seus estatutos, modificado polo Decreto 9/2021, do 21 de xaneiro, polo que se suprime o Centro Informático para a Xestión Tributaria, Económico-Financeira e Contable (Cixtec) e se modifica o Decreto

252/2011, do 15 de decembro, polo que se crea a Amtega e se aproban os seus estatutos.

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 35/2022, do 17 de marzo, (DOG núm. 56 do 22 de marzo de 2022), actuando en nome e representación da devandita entidade.

Interveñen en función dos seus respectivos cargos e no exercicio das facultades que teñen conferidas e,

EXPOÑEN

1. Que o artigo 27 do Decreto 198/2020, atribúelle á Consellería de Cultura, Educación e Universidade a través da Secretaría Xeral de Política Lingüística (en adiante SXPL), entre outras, as funcións de promover, coordinar e desenvolver a política de investigación que favoreza a normalización lingüística nas súas diferentes manifestacións, especialmente no eido do léxico e a terminoloxía, así como, promover o diálogo, a cooperación e a colaboración con outras institucións competentes en materia de política lingüística e desenvolver medidas para o fomento do uso do idioma nos diferentes eidos sociais e económicos de Galicia, co fin de incrementar a dispoñibilidade de produtos e servizos en galego.

2. Que a Axencia para a Modernización Tecnolóxica de Galicia (en adiante Amtega) encadrase nas entidades instrumentais do sector público autonómico reguladas no título III da Lei 16/2010, do 17 de decembro, de organización e funcionamento da Administración xeral e do sector público autonómico de Galicia. Está adscrita á Presidencia da Xunta de Galicia e conta con personalidade xurídica diferenciada respecto da Administración xeral da Comunidade Autónoma de Galicia, patrimonio e tesouraría propios e autonomía de xestión nos termos establecidos pola normativa vixente e ten como obxectivos básicos a definición, o desenvolvemento e a execución dos instrumentos da política da Xunta no ámbito das tecnoloxías da información e comunicacións, a innovación e o desenvolvemento

tecnolóxico. Os seus estatutos atribúenlle, entre outras, as seguintes competencias e funcións: a dirección e xestión de todas as actuacións da Xunta de Galicia en materia de tecnoloxías da información e as comunicacións e o deseño e liderado de iniciativas que contribúan a impulsar o desenvolvemento da Sociedade da Información en Galicia, garantindo os dereitos dos cidadáns e eliminando as barreiras que se opoñan á expansión e ao uso das novas tecnoloxías.

3. Que a Universidade de Santiago de Compostela (en adiante USC) é unha entidade que leva a cabo actividades de investigación, docencia e desenvolvemento científico e tecnolóxico, que ten entre os seus fins o de procurar o seu enraizamento na sociedade galega que a sustenta e, en xeral, contribuír ao seu progreso social, económico, cultural e tecnolóxico, e é unha peza clave no proceso de normalización lingüística tanto polos diferentes campos culturais que abrangue como pola súa condición de referente social de prestixio o que a configura como un dos principais axentes xeradores de coñecemento de Galicia. A referida institución universitaria conta entre os seus centros de investigación co Instituto da Lingua Galega (ILG), referente de recoñecido prestixio na súa contribución ao desenvolvemento da investigación científica sobre a lingua galega. Dende o ano 1975, o centro de investigación está traballando na recollida de material dialectal no dominio lingüístico galego e, froito desta recollida, é un ficheiro lexicográfico que contén o vocabulario galego extraído de obras editadas ou inéditas que apareceron en forma distinta á de dicionario. Son tamén resultado do seu labor os materiais do Atlas Lingüístico Galego, do que se levan publicado varios volumes. O ficheiro lexicográfico como os cadernos do ALGa, están dispoñibles para os investigadores na sede do ILG, e son hoxe unha fonte de información imprescindible para calquera traballo sobre gramática e léxico, o centro dispón do Arquivo do Galego Oral, así como o Corpus Oral Informatizado da Lingua Galega (CORILGA), que permitirá investigar sobre a variación e o cambio no galego falado actualmente, e no que se incorporan ferramentas de tecnoloxías da fala.

A USC, a través do Centro Singular de Investigación en Tecnoloxías Intelixentes, (CiTIUS) dispón, así mesmo, de recursos materiais e persoais no ámbito tecnolóxico e

de investigación que complementan as necesidades para o desenvolvemento e execución do Proxecto Nós. Dito centro conta cunha dilatada experiencia en investigación, desenvolvemento e innovación no ámbito da intelixencia artificial e das tecnoloxías da linguaxe natural, nomeadamente nas linguas ibéricas, como o galego, portugués e castelán, e colabora con empresas do sector para dinamizar e poñer en valor emprendementos vinculados a proxectos relacionados coa Intelixencia Artificial (IA) e as Tecnoloxías Lingüísticas (TL)

4. Que a Unión Europea está a traballar na consolidación dun modelo de regulación e en políticas que permitan o desenvolvemento e aproveitamento das tecnoloxías dixitais en favor da sociedade e a da economía, o que lle ten levado a propoñer, en 2020 e 2021, a Proposta de Regulamento do Parlamento Europeo e do Consello polo que se establecen normas harmonizadas en materia de intelixencia artificial (Lei de intelixencia artificial) e se modifican determinados actos legislativos da Unión (COM/2021/206). Isto comporta que as iniciativas innovadoras que desenvolva a Xunta de Galicia neste eido deban introducir aspectos de investigación de vangarda que permitan que tales desenvolvementos se planifiquen prevendo que sexan compatibles non só coa regulación actual, senón cos futuros desenvolvementos normativos.

5. No ámbito normativo da Comunidade Autónoma de Galicia, a Lei 6/2013, do 13 de xuño, do sistema universitario de Galicia establece no seu artigo 111 que a transferencia de coñecemento é un dos mecanismos polos que o valor creado na universidade a través da investigación retorna á sociedade. Así pois, pola súa contribución ao desenvolvemento e ao crecemento económico, a transferencia constitúe unha actividade fundamental para a universidade.

6. Nos últimos anos, o impulso da presenza da lingua galega no desenvolvemento dixital xa quedou recollido na estratexia tecnolóxica da Xunta de Galicia, tanto na Axenda Dixital 2014.gal, como na posterior Axenda Dixital de Galicia 2020. Neste ámbito realizáronse importantes esforzos para dotar o galego de ferramentas tecnolóxicas (tradutores automáticos, dicionarios en liña, plataformas en liña para a difusión do coñecemento da lingua galega etc.) e poñelas á disposición do conxunto da cidadanía a través de iniciativas como os recursos lingüísticos do portal web

<https://www.lingua.gal> da Secretaría Xeral de Política Lingüística da Xunta de Galicia e, outros, promovidos por diversas institucións e entidades públicas ou privadas.

Porén, a situación do galego no contexto dixital non é equiparable a outras linguas como o inglés ou o castelán, polo que precisa dun impulso decidido que a sitúe como lingua de uso convencional nos recursos tecnolóxicos de uso común pola sociedade e para acadar que o galego dea un paso decidido cara adiante no eido das tecnoloxías intelixentes, o “Proxecto Nós” e a súa estratexia de intelixencia artificial ao servizo da lingua galega, pretende ter visibilidade a través da acción conxunta das partes deste convenio e, nesa dimensión colaborativa, aliñarase tamén cos principios inspiradores da nova Estratexia de Especialización Intelixente de Galicia RIS3 2021-2027, para lles dar respostas ás demandas e ás necesidades da sociedade.

7. Nós é un proxecto inédito, que leva a denominación dunha xeración de intelectuais galegos que foron quen de impulsar a cultura galega a través da súa lingua e, igual que fixera a intelectualidade galega do século XIX, o proxecto nace, na actualidade, coa virtualidade de configurarse como a pedra alicerce do ecosistema dixital que permitirá xerar recursos dixitais e lingüísticos transversais que lle proporcionaran ao galego a posibilidade de se incorporar a plataformas tecnolóxicas, das cales, ata o de agora, estivera excluído: plataformas audiovisuais, videoxogos e noutros recursos tecnolóxicos e dispositivos dixitais presentes e futuros.

A finalidade que persegue o Proxecto Nós co seu desenvolvemento é acadar unha ampla variedade de recursos lingüísticos en galego de alta calidade e que serán publicados en formato dixital baixo licenzas libres e a través de ferramentas con códigos abertos para ser empregadas, libre e gratuitamente en calquera tipo de soporte dixital polas diferentes institucións públicas e, mesmo, polos axentes que conforman o tecido económico empresarial e, asemade, pola sociedade en xeral.

A infraestrutura de recursos tecnolóxicos e lingüísticos que se xeren co desenvolvemento do proxecto permitirá que o galego experimente un salto cualitativo dentro do universo dixital, fito histórico que favorecerá o proceso de transformación dixital da acción da Xunta de Galicia e doutras administracións

públicas e garantirá que a cidadanía en xeral poida participar no mundo dixital sen renunciar a súa lingua ao mesmo nivel que un falante dunha lingua global, como pode ser o inglés ou o castelán. Os recursos resultantes do proxecto clasificaranse en tres categorías: recursos lingüísticos, tecnoloxías lingüísticas e demostradores. Nomeadamente desenvolverase un sintetizador de voz, que vai permitir a terceiros incluír nos seus produtos voces artificiais de alta calidade en galego, podendo escoller xénero e idade; un recoñecedor de fala, que posibilite a interacción persoa-máquina; un xerador de lingua natural que converta datos en oracións en galego; sistemas de diálogo xenérico, así como un elenco de respostas e preguntas, tanto a través de texto como de voz; un tradutor automático a outras linguas e viceversa; e un corrector lingüístico automático, tanto ortográfico como léxico e gramatical.

8. O proxecto na súa dimensión global encádrase dentro das estratexias galega, española e europea orientadas ao desenvolvemento de tecnoloxías, ferramentas e recursos para a proxección das linguas nos ámbitos de aplicación da intelixencia artificial. Neste contexto, a Xunta de Galicia aprobou a Estratexia Galicia Dixital 2030, que identifica a innovación tecnolóxica como instrumento para a proxección dos sinais de identidade de Galicia e, en particular, da lingua galega para promover o coñecemento e o seu uso en todo tipo de soportes e formatos dixitais. Participará tamén do proceso participativo da Xunta de Galicia a través da Axencia Galega de Innovación, para a elaboración da Estratexia de Especialización Intelixente RIS3 para o período 2021-2027, que definirá os investimentos para o dito período no ámbito da investigación e a innovación e que reforzará especialmente, tal como esixe a Comisión Europea, unha maior implicación dos axentes sociais e da cidadanía en xeral.

9. No ano 2021 a Xunta de Galicia a través da Consellería de Cultura, Educación e Universidade e a Amtega e, a USC asinaron un convenio de colaboración para a xeración de recursos lingüísticos e tecnolóxicos en lingua galega. Proxecto Nós: O galego na sociedade e economía da intelixencia artificial, cuxa vixencia rematou o 15 de decembro de 2021. A finalidade deste convenio foi a elaboración conxunta do plan de traballo do Proxecto Nós durante as seguintes catro anualidades posteriores ao remate da súa vixencia. No citado plan recóllese o detalle dos diferentes proxectos,

recursos, demostradores e casos de uso a desenvolver no proxecto durante o período de execución que desenvolven as tres liñas estratéxicas seguintes:

- Liña 1: conversación (oral ou escrita) con persoas
- Liña 2: calidade da lingua
- Liña 3: xestión da información

10. Que é interese das partes asinantes, no marco de colaboración establecida, manter e intensificar esta colaboración, mediante a sinatura dun novo convenio de colaboración para desenvolver o citado proxecto co fin de poñer á disposición da sociedade en xeral recursos lingüísticos en galego de alta calidade, en liña e con carácter gratuito para todas as persoas usuarias.

11. Que a Xunta de Galicia contribuirá ao financiamento do convenio con fondos propios e con fondos europeos como parte da resposta da Unión Europea para reparar os danos económicos e sociais xerados pola pandemia da COVID-19 no marco do Eixe REACT-UE do programa operativo FEDER- Galicia 2014-2020.

12. Que as partes son conscientes da responsabilidade que teñen no proceso normalizador, así como á hora de cumprir cos obxectivos previstos na Lei 3/1983, de normalización lingüística e no Plan xeral de normalización da lingua galega que aprobou, por unanimidade, o Pleno da Cámara do Parlamento de Galicia na súa sesión do día 21 de setembro de 2004, xa que logo, a comunidade autónoma necesita dispor de traballos de investigación e dunha estratexia común para o desenvolvemento dun proxecto, como Nós, que permitirá xerar os recursos necesarios para colocar o galego na vangarda das tecnoloxías intelixentes, nomeadamente naquelas áreas relacionadas coas tecnoloxías lingüísticas e, de modo xeral, na sociedade e economía da intelixencia artificial.

13. Que a sinatura deste convenio ten por finalidade a prestación dun servizo de carácter público e social que a través da transferencia do coñecemento incide na

sociedade en xeral, e que se executará mediante a acción coordinada e colaborativa pública a través de diferentes departamentos da Xunta de Galicia e a USC, por medio dos organismos especializados dela dependentes, como son, O ILG e o CiTIUS, por ser estes, referencias indiscutibles no eido da investigación lingüística e das tecnoloxías intelixentes. Son as únicas entidades publicas que no ámbito da Comunidade Autónoma Galega contan cos coñecementos técnicos, traballos de investigación avanzada en materia de intelixencia artificial e tecnoloxía lingüística, así como dos medios materiais e humanos máis especializados para realizar as tarefas de investigación necesarias nos paradigmas avanzados da intelixencia artificial (análise de datos sobre procesos, analítica avanzada de procesos, tecnoloxía da linguaxe natural e intelixencia artificial responsable) para a realización do obxecto do convenio. En consecuencia, pola singularidade das accións que se van desenvolver ao abeiro deste convenio e a súa especialidade non é posible promover a concorrência en convocatoria pública.

Por todo isto, as partes asinantes acordan suscribir este convenio de colaboración coa finalidade de darlle un pulo decidido a lingua galega no ecosistema dixital e na vangarda das tecnoloxías intelixentes.

CLÁUSULAS

Primeira.- Obxecto do convenio

1. O obxecto deste convenio é establecer as condicións polas que se rexerá a colaboración entre a Consellería de Cultura, Educación e Universidade, a través da SXPL e do Centro Ramón Piñeiro para a Investigación en Humanidades, (en adiante CRPIH) dela dependente, a Amtega e a USC a través do ILG e CiTIUS, para o desenvolvemento gradual e por fases do Proxecto Nós: o galego na sociedade e economía da intelixencia artificial, co seguinte plan de traballo:

1.1 Desenvolveranse as tres liñas estratéxicas de acción xa identificadas no apartado 9 da parte expositiva deste convenio. Dos oito subproxectos dos que consta o Proxecto Nós completo (a desenvolver en catro anualidades) atenderase

este ano aos seis subproxectos de maior impacto en relación coas tres liñas estratéxicas xa indicadas na parte expositiva:

- Síntese de voz
- Recoñecemento da fala
- Xeración automática de textos
- Sistemas de diálogo
- Tradución automática
- Corrección e avaliación lingüística

1.2 Para cada un dos subproxectos definírase o conxunto de recursos, prototipos de proba e un demostrador imprescindibles para súa realización, e obterase un demostrador prioritario co obxectivo de dar visibilidade ao traballo realizado na elaboración dos corpus lingüísticos e modelos de lingua, tal como se indica no anexo.

1.3 Así mesmo, no plan de traballo para o ano 2022 identificaranse seis paquetes de traballo do proxecto, sobre os que incidirá as tarefas dos investigadores/as:

- WP1: Xestión global, legal e ética. Concentra as tarefas de xestión, riscos, persoal e cuestións legais e éticas.
- WP2: Monotorización do estado da arte. Consiste nun estudo exhaustivo do estado da arte das diferentes tecnoloxías, o que supón o deseño de prototipos dos recursos e dun plano de experimentación.
- WP3: Obtención de recursos lingüísticos. Alén doutras tarefas específicas para cada subproxecto, este paquete tamén recolle a adquisición e transformación dos corpus de referencia existentes, a obtención de corpus web de gran tamaño de carácter xeral e a avaliación da viabilidade de obter outros tipos de *datasets* (tarefa 0).
- WP4: Construción e avaliación de modelos. Alén doutras tarefas específicas para cada subproxecto, este paquete tamén recolle a construción de modelos

transversais e a elaboración de ferramentas lingüísticas básicas a partir destes (tarefa 0).

- WP5: Desenvolvemento de demostradores e casos de uso. Inclúe as tarefas para experimentar os demostradores e os casos de uso.

- WP6: Diseminación e difusión. Inclúe tanto a diseminación do software, modelos e recursos lingüísticos (repositorios) como outra información do proxecto (redes sociais, web).

2. Ademais da definición do obxecto do convenio nos parágrafos anteriores, as liñas de acción que se van desenvolver durante o ano 2022 xúntanse como anexo deste convenio e formarán parte integrante del.

Segunda.- Partes do convenio

1 A Xunta de Galicia coa implicación dos seguintes departamentos:

1.1 A Consellería de Cultura, Educación e Universidades, a través da SXPL e do CRPIH, dela dependente.

1.2 A Amtega

2 A USC a través do ILG e o CiTIUS, por ser centros de referencia en investigación lingüística e de tecnoloxías intelixentes.

Terceira.- Compromisos das partes

As partes asinantes deste convenio comprométense a desenvolver de forma colaborativa e conxunta no ámbito das súas competencias as accións necesarias para o eficaz cumprimento do obxecto deste convenio e nomeadamente as seguintes:

1. A Xunta de Galicia, a través da SXPL e o CRPIH dependentes da Consellería de Cultura, Educación e Universidades e da Amtega asume os seguintes compromisos:

- A participar activamente nas accións estratéxicas para o desenvolvemento do plan de traballo do Proxecto Nós durante o ano 2022, en especial nos aspectos vinculados ao seu ámbito competencial.

- A contribuir ao financiamento do obxecto deste convenio con fondos propios polo que se refire á SXPL e con fondos europeos polo que se refire á Amtega nos termos previstos na cláusula quinta deste convenio. O financiamento parcial con fondos europeos responde a estratexia da Unión Europea para reparar os danos económicos e sociais xerados pola pandemia da COVID -19 no marco do Eixo REACT-UE do programa operativo FEDER- Galicia 2014-2020.

- A cumprir as obrigas en materia de información e comunicación segundo o establecido no apartado 2.2 do anexo XII do RDC e de acordo coas especificacións técnicas do regulamento 821/2014.

- Ao mantemento dos investimentos dedicándoos á finalidade para a que foron realizados.

- A Amtega terá a obriga de aplicar medidas antifraude eficaces e proporcionadas no seu ámbito de xestión e a cumprir a normativa comunitaria, estatal e autonómica de aplicación, en particular, a normativa de subvencións (art. 125.4.c do Regulamento(UE) Nº 1303/2013). A detección de feitos que puideran ser constitutivos de fraude ou irregularidade serán comunicados ao Servicio Nacional de Coordinación Antifraude (SNCA) nos termos establecidos na comunicación 1/2017 del SNCA, do 6 de abril, sobre a forma na que poden proceder as persoas que teñan coñecemento de feitos que poidan ser constitutivos de fraude ou irregularidade en relación con proxectos ou operacións financiados total ou parcialmente con cargo a fondos da Unión Europea.

2) A USC, a través do ILG e o CiTIUS comprométese:

- A colaborar activamente en todas as actuacións previstas neste convenio e no desenvolvemento das tarefas investigadoras e de realización dos resultados previstos no Plan de traballo para o ano 2022 segundo a metodoloxía e a planificación de datas para a obtención de resultados no desenvolvemento do proxecto ao longo desta anualidade.

- A asignar os medios humanos e materiais que sexan necesarios e, de ser o caso, a contratación das colaboracións externas que se precisen para a execución das actuacións previstas neste convenio.
- A achegar coa xustificación os informes e documentos que acrediten fidedignamente o desenvolvemento do obxecto do presente convenio, así como os informes e demais documentos requiridos pola Xunta de Galicia.
- A contribuir ao financiamento do obxecto deste convenio con fondos propios nos termos previstos na cláusula quinta deste convenio.
- Ao mantemento dos investimentos dedicándoos á finalidade para a que foron realizados.
- A facer figurar expresamente o financiamento da Xunta de Galicia e dos fondos europeos na totalidade das actuacións asociadas a este Convenio.

Cuarta.- Responsabilidade do proxecto

1) A Consellería de Cultura, Educación e Universidade, a través da SXPL e o CRPIH, a Amtega, e a USC a través do ILG e o CiTIUS, serán as encargadas de realizar as actuacións necesarias para desenvolver as actividades obxecto do convenio, nun proceso de colaboración continuo. As entidades asinantes, como coorganizadoras, terán unha participación activa e esencial na realización do obxecto do convenio e compartirán, nos termos que se indiquen neste convenio, a titularidade dos resultados dos traballos realizados.

2) A participación do equipo de traballo ou equipos de traballo dos investigadores da USC nas actividades recollidas neste convenio efectuarase conforme a regulamentación propia da USC recollida nos seus estatutos e no Regulamento para a realización de actividades de investigación aprobado polo Consello de Goberno desta universidade o 20 de xuño de 2005 ao abeiro do establecido no artigo 83 da Lei orgánica 6/2001, do 21 de decembro, de universidades, modificada pola Lei orgánica 4/2007, do 12 de abril.

3) Calquera das institucións e entidades que participan no proxecto, cada unha no seu ámbito competencial de actuación, dispón de persoal asignado aos servizos xurídicos, persoal docente e/ou investigador, lingüistas e a infraestrutura técnica e a tecnoloxía máis adecuada para garantir a consecución dos obxectivos de interese público perseguidos coa sinatura deste convenio. Se fose necesario contratar medios humanos, técnicos ou materiais complementarios para o axeitado desenvolvemento do obxecto do convenio, a devandita contratación realizarase de acordo coa normativa específica que resulte de aplicación ao abeiro da Lei 9/2017, de 8 de novembro, de contratos do sector público, pola que se traspoñen ao ordenamento xurídico español as directivas do Parlamento Europeo e do Consello 2014/23/UE e 2014/24/UE, de 26 de febreiro de 2014.

4) A sinatura deste convenio non implica relación laboral, contractual ou de calquera outro tipo entre os profesionais que vaian desenvolver as actividades previstas nel e mais a Xunta de Galicia, de tal xeito que a esta non se lle poderá esixir responsabilidade ningunha –nin directa, nin indirecta, nin subsidiaria– polos actos ou feitos acaecidos no desenvolvemento do convenio.

Quinta.- Orzamento e financiamento do convenio

1. O orzamento previsto para a realización do obxecto do presente convenio durante o ano 2022 ascende a un total de **626.218,24 €**

1.1. **A Xunta de Galicia**, a través da da Consellería de Cultura, Educación e Universidade e da Amtega colaborará na realización das actividades obxecto deste convenio cunha cantidade máxima de **cincocentos mil euros (500.000,00 €)**, e unha parte deste importe será sufragado con fondos europeos:

a) A Consellería de Cultura, Educación e Universidade a través da SXPL colaborará con fondos propios cun importe máximo de **cento sesenta mil euros (160.000,00 €)** con cargo a aplicación orzamentaria

10.50.151A.640.1. A referida achega equivale a unha porcentaxe sobre o importe total do proxecto do **25,55%**.

b) A Amtega colaborará con fondos europeos cun importe máximo de **trescentos corenta mil euros (340.000,00 €)** con cargo a aplicación orzamentaria **04.A1.571A.626.0** e financiado ao 100% no marco do Eixe REACT-EU do Programa Operativo FEDER Galicia 2014-2020 como resposta da Unión Europea á pandemia do COVID-19. A referida achega equivale a unha porcentaxe sobre o importe total do proxecto do **54,29%**.

Os devanditos importes destinaranse ao cofinanciamento dunha parte dos custos necesarios para o desenvolvemento do obxecto do convenio segundo o plan de traballo do ano 2022 nun proceso de colaboración continuo baseado nos principios inspiradores da intelixencia artificial e das tecnoloxías da linguaxe, de acordo coa seguinte táboa:

Paquete de traballo	Persoal/recurso	Duración	Orzamento
WP1. Xestión global, legal e ética	1 Xestor/a de proxecto Asesoría legal e ética	3 meses	13.908,96 €
		8 meses	15.000,00 €
	Estudo impacto socio-económico	8 meses	13.500,00 €
WP2. Estado da arte	3 Investigadores/as posdoutorais	8 meses ¹	98.180,64 €
WP3 Recursos lingüísticos	Investigadores posdoutorais (WP2)	-	-
WP4 Construción de modelos			
0. Elaboración de corpus, modelos transversais e ferramentas básicas	Recursos de computación (servidores, nube)	-	-
		-	40.000,00 €
1. Síntese de voz	Contratación de recursos humanos	8 meses	29.832,86 €
	Contratación de recursos humanos	7 meses	20.052,06 €
	Colaboracións externas (gravacións de voz)		19.000,00 €

Paquete de traballo	Persoal/recurso	Duración	Orzamento
2. Recoñecemento da fala	Contratación de recursos humanos	8 meses	20.052,06 €
	Colaboracións externas	8 meses	46.337,32 €
3. Xeración automática de textos	Contratación de recursos humanos	7 meses	20.052,06 €
4. Sistemas de diálogo	Contratación de recursos humanos	7 meses	40.104,12 €
5. Tradución automática	Contratación de recursos humanos	8 meses	29.832,86 €
6. Corrección e avaliación lingüística	Contratación de recursos humanos	7 meses	20.052,06 €
WP5. Demostradores e casos de uso	Investigadores posdoutorais (WP2), contratados/as (WP3,WP4)	-	-
WP6. Diseminación e difusión	Software, web, redes sociais Viaxes e asistencia a eventos		10.000,00 €
			14.095,00 €
TOTAL CUSTOS DIRECTOS			450.000,00 €
CUSTOS INDIRECTOS (10%)			50.000,00 €
*TOTAL + CUSTOS INDIRECTOS			500.000,00 €

*A referida achega é unha porcentaxe do importe total do proxecto e financiará exclusivamente unha parte dos custos do obxecto do convenio sen que a USC obteña beneficio ou lucro ningún.

1.2. A USC, a través do ILG e o CiTIUS, contribuirá, con cargo aos seus orzamentos propios, a financiar a diferenza entre a achega económica da Xunta de Galicia e o orzamento total do convenio. Para tales efectos, a USC cofinanciará unha parte dos custos nos que incurran as partes asinantes pola realización das obrigas derivadas deste convenio e que non financiará a Xunta de Galicia (gasto en persoal, material, instalacións, equipamentos.....). O importe estimado desta asignación é de **cento vinte e seis mil douscentos dezaoto euros e vinte catro céntimos. (126.218, 24 €)** que equivale á porcentaxe do **(20,16%)** sobre o importe total do proxecto, sen prexuízo do seu incremento en atención ás horas de traballo efectivamente dedicadas ao desenvolvemento do obxecto deste convenio polo persoal propio docente e investigador adscrito a este proxecto.

1.3 O custo referido ao **valor/hora** do equipo ou dos equipos de traballo que participen na dirección técnica e de execución do plan de traballo do proxecto Nós durante a vixencia do convenio será a seguinte :

Achega USC dedicación do persoal propio	Perfil	h/ semana	Custo/h	Total horas	Custo total
	CiTIVS				
	CU	4	59.55	116	6907.80
	TU	12	41.69	348	14508.12
	CU	12	51.43	348	17897.64
	RyC	12	22.17	348	7715.16
	RyC	2	27.33	58	1585.14
	JCI	2	22.61	58	1311.38
	TU	2	39.18	58	2272.44
	CDr	2	44.37	58	2573.46
	CDr	2	48.34	58	2803.72
	TU	2	48.92	58	2837.36
	ILG				
	TU	4	36.98	116	4289.68
	TU	4	43.24	116	5015.84
	CDr	12	46.5	348	16182.00
	CU	12	54.21	348	18865.08
	CU	6	54.33	174	9453.42
SUMA		90		2610	114.218,24

Custos do material bibliográfico e documental, recursos informáticos, técnicos e tecnolóxicos, e outros necesarios para o desenvolvemento do proxecto

12.000

TOTAL USC RECURSOS PROPIOS

126.218,24

Neste importe inclúense, ademais, de parte dos custos de dirección técnica e da execución do proxecto, tamén, os custos do material bibliográfico e documental necesario, os recursos e medios informáticos, técnicos e tecnolóxicos, necesarios para o bo desenvolvemento do proxecto nas sedes dos centros investigadores.

1.4 O importe do convenio destinarase na súa totalidade a cubrir o custo do desenvolvemento do proxecto segundo o programa de traballo do ano 2022, sen que a USC obteña lucro ningún.

Sexta.- Forma de pagamento e xustificación do convenio

1. A documentación requirida na fase xustificativa presentarase electronicamente accedendo ao expediente na Carpeta do cidadán da persoa solicitante/representante, <https://sede.xunta.gal/carpeta-do-cidadan>. Código de procedemento PR004A.

2. A **USC**, para recibir as achegas económicas das unidades ou entes cofinanciadores da Xunta de Galicia por mor dos custos necesarios para o desenvolvemento do proxecto segundo o Plan de traballo para o ano 2022, presentará electronicamente, coa data límite do **10 de decembro** a seguinte documentación:

- Un certificado da persoa responsable da Xerencia da USC ou do órgano de control que corresponda, que acredite a toma de razón en contabilidade de que se executou o obxecto do convenio e que as achegas das partes cofinanciadoras foron destinadas integramente ao desenvolvemento do obxecto do convenio e ao pagamento dos custos por hora do persoal investigador propio e, de ser o caso, os custos do persoal externo contratado, que irá acompañado dunha relación valorada das horas dedicadas ao proxecto polo equipo ou equipos de traballo da USC e, tamén, dos colaboradores externos.

- Unha memoria asinada pola persoa investigadora principal do equipo de traballo (ou dos equipos de traballo, de configurarse máis dun) na que se fará constar de forma detallada as accións realizadas, os resultados obtidos e as incidencias máis significativas.

3. No caso de que a acreditación dos custos do convenio e, máis concretamente, a súa equivalencia en horas de dedicación ao proxecto, sexa inferior ao importe do

convenio, a achega das unidades ou entes cofinanciadores diminuírase na mesma proporción.

4. Para os efectos de axilizar a tramitación dos pagamentos do convenio polas partes asinantes, a Consellería de Cultura, Educación e Universidade a través da SXPL, asume as tarefas de comprobación da xustificación que presente a USC para acreditar as horas de dedicación ao proxecto polos seus especialistas. Sen prexuízo diso, logo da comprobación que corresponda, a SXPL dará conta do resultado á Amtega. Cada ente cofinanciador tramitará o expediente administrativo para o pagamento da súa achega e a Amtega cumprirá as obrigas que lle corresponda en materia de información, comunicación e de medidas antifraude como consecuencia do financiamento parcial do convenio con fondos europeos.

5. O pagamento realizarase na conta bancaria que sinale a USC mediante certificado acreditativo da súa titularidade.

6. Antes de proceder á sinatura deste convenio e para o seu pagamento, a USC deberá presentar:

- Un certificado de non ter pendente de pagamento ningunha débeda coa Administración pública da Comunidade Autónoma de Galicia, expedido pola Consellería de Facenda.
- Un certificado de estar ao día das obrigas tributarias, expedido pola correspondente delegación ou axencia do Ministerio de Economía e Facenda.
- Un certificado de estar ao corrente das obrigas sociais, expedido pola Tesourería da Seguridade Social.

A sinatura deste convenio comportará a autorización ao órgano xestor (SXPL) para solicitar as certificacións que deban emitir a Axencia Estatal da Administración Tributaria, a Tesourería da Seguridade Social e a Consellería de Facenda e Administración Pública da Xunta de Galicia.

7. A USC facilitará toda a información que lle requira a Xunta de Galicia a través da SXPL ou da Amtega, así como a Intervención Xeral da Comunidade Autónoma, de ser o caso, a Intervención Xeral da Administración do Estado, o Consello de Contas e/ou o Tribunal de Contas no exercicio das súas funcións de fiscalización e control. Igualmente, estará suxeita ás comprobacións e verificacións a realizar polo OI, AG ou AC e, de ser o caso, polos órganos de control da Comisión Europea e do Tribunal de Contas Europeo ou outras instancias de control.

8. A SXPL facilitaralle unha copia de toda a documentación presentada pola USC á Amtega xunto coa certificación da realización das actuacións pola USC, no momento que proceda tramitar os pagamentos. O cumprimento das obrigas derivadas do financiamento europeo lle corresponderá á Amtega.

Sétima.- Resultados conxuntos

1. Os resultados obtidos como consecuencia da colaboración obxecto deste convenio corresponderán de xeito conxunto ás partes asinantes e, cada unha delas, poderán facer uso gratuíto destes co consentimento unánime das outras partes.

2. Os dereitos incorporais de propiedade intelectual de carácter patrimonial respecto dos recursos dixitais ou tecnolóxicos que, de ser o caso, se xeren nas diferentes fases de desenvolvemento do proxecto, de titularidade común das partes asinantes, pola sinatura do presente convenio cederanse de xeito gratuíto á Xunta de Galicia para o seu destino ao uso público con formatos dixitais baixo licenzas libres e con códigos abertos para ser empregados, libre e gratuitamente, polas persoas destinatarias.

3. Exceptúase do disposto no parágrafo anterior o resultado da actividade propia do labor investigadora dos traballos de investigación que se utilicen como medios para o desenvolvemento do obxecto do convenio e cuxa propiedade intelectual lle corresponderá á USC.

Oitava.- Difusión da colaboración das partes

Nas actuacións informativas e divulgativas, así como nas publicacións que se fagan relacionadas coas actividades deste convenio, farase constar nun lugar destacado a colaboración das partes asinantes e utilizaranse, os anagramas ou logos que correspondan, segundo o previsto pola normativa de identidade corporativa da Xunta de Galicia da USC e atenderase a normativa que regula a imaxe corporativa dos proxectos financiados ao abeiro do Eixo REACT UE do Programa Operativo FEDER Galicia 2014-2020.

Sexa cal sexa o medio de difusión, cada unha das partes comprométese a facer mención das outras partes e deste convenio.

Novena.- Comisión de seguimento

Crearase unha comisión de seguimento das accións obxecto do presente convenio, facultada para solventar as dúbidas que poidan xurdir na aplicación e interpretación das súas cláusulas e, así mesmo, proporlles ás partes as resolucións de cantas dúbidas xurdan na súa execución. Estará presidida pola persoa representante da SXPL ou pola persoa na que esta delegue, e integrarase ademais por:

- Unha persoa representante da USC, nomeada polo seu titular.
- Unha persoa representante da Amtega, nomeada polo seu titular.

A comisión de seguimento manterá reunións periódicas de coordinación, seguimento e control das actividades realizadas e xestionará calquera modificación que xurda no desenvolvemento do proxecto nun proceso de colaboración mutua entre as partes asinantes. Ás reunións da comisión asistirán os seus membros ou as persoas en quen deleguen.

A comisión poderá adoptar as decisións e ditar as instrucións necesarias para asegurar o cumprimento do obxecto do presente convenio. Os acordos da comisión serán inmediatamente executivos.

Décima.- Vixencia

Este convenio entrará en vigor o mesmo día da súa sinatura dixital pola persoa asinante que o faga en último lugar e terá vixencia ata o **20 de decembro** do ano 2022 agás que sexa denunciado por escrito por algunha das partes asinantes polo incumprimento dos compromisos que nel se recollen, cun mes de antelación polo menos.

Antes da finalización do prazo de vixencia, os asinantes poderán acordar unánimemente a súa prórroga por un máximo dun ano máis. De prorrogarse o convenio, na correspondente addenda estableceríanse as obrigas concretas das partes e o Plan de traballo para o ano 2023.

Décimo primeira.- Utilización da lingua galega

Todas as accións e actividades que se realicen ao abeiro deste convenio deberán levarse a cabo de acordo co indicado 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 o respecto rigoroso á toponimia oficial nos termos previstos no artigo 10 da citada lei.

O incumprimento desta estipulación será considerada causa específica de resolución deste convenio.

Décimo segunda.- Rexistro de Convenios. Transparencia e bo goberno

A sinatura deste 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, co artigo 8 da Lei 19/2013, do 9 de decembro (BOE 10-12-2013), de transparencia, acceso á información pública e bo goberno e co Decreto 126/2006, do 20 de xullo, polo que se regula o Rexistro de Convenios da Xunta de Galicia

Décimo terceira.- Causas de resolución

Este convenio resolverase:

- 1 Por incumprimento total ou parcial dalgunha das estipulacións que o regulan.
- 2 Polo transcurso do tempo estipulado sen que se culminen as actividades acordadas.

De se resolver o convenio, as partes asinantes poderán, se é o caso, esixir a devolución das cantidades achegadas.

Décimo cuarta- Protección de datos de carácter persoal

As partes asinantes comprométese a cumprir, en todo momento, as disposicións contidas na normativa básica estatal e autonómica que xustifica a lexitimidade do tratamento dos datos persoais e dos dereitos dixitais.

Na aplicación deste convenio respectaranse cantas esixencias establece o Regulamento UE 679/2016 xeral de protección de datos e a Lei orgánica 3/2018, do 5 de decembro (BOE 06/12/2018) de protección de datos persoais e garantía dos dereitos dixitais, adoptando cantas medidas resulten necesarias para cumprir as súas previsións e, en particular, para garantir a seguridade e integridade dos datos persoais e dos dereitos dixitais e, a súa protección fronte a alteracións, perdas, tratamentos ou accesos non autorizados.

Serán de aplicación todos os dereitos e obrigas derivados da normativa aplicable en materia de protección de datos e garantía dos dereitos dixitais.

Décimo quinta.- Administración dixital

As partes asinantes deste convenio 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 a USC.

Décimo sexta.- Natureza xurídica e xurisdición

Este convenio ten carácter administrativo e os seus efectos rexeranse polo establecido nas súas estipulacións, e quedará suxeito ao disposto nos artigos 47 e seguintes da Lei 40/2015, do 1 de outubro, do réxime xurídico do sector público.

Así mesmo, será de aplicación á Xunta de Galicia a Lei 14/2013, do 26 de decembro, de racionalización do sector público autonómico, polo que se refire ao contido da colaboración mediante o desempeño conxunto de tarefas de interese público. O artigo 14 prevé que a Administración xeral da comunidade autónoma e as entidades do sector público, ou estas entre elas, poderán formalizar acordos de colaboración ou cooperación para o servizo aos intereses xerais da súa competencia.

Doutra banda, aínda que é un dos negocios xurídicos excluídos do ámbito de aplicación da Lei 9/2017, de 8 de novembro, de contratos do sector público, pola que se traspoñen ao ordenamento xurídico español as directivas do Parlamento Europeo e do Consello 2014/23/UE e 2014/24/UE, de 26 de febreiro de 2014, no non previsto aplicaráselle os principios recollidos na citada lei, 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, pola que se aproba o texto refundido da Lei de réxime financeiro e orzamentario de Galicia e demais normativa vixente na materia.



XUNTA
DE GALICIA



UNIÓN EUROPEA



UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA



Xacobeo 21-22

En atención á tipoloxía de fondos que financian as actuacións previstas no convenio, será de aplicación o Regulamento 1303/2013, e a normativa comunitaria de desenvolvemento, así como a Orde HFP/1979/2016, de 29 de decembro, modificada pola Orde HAC/114/2021, de 5 de febreiro polo que se aproban as normas sobre os gastos subvencionables.

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

En proba de conformidade co seu contido as partes asinan telematicamente este convenio.

O conselleiro de Cultura, Educación e
Universidade

Román Rodríguez González

A directora da Axencia para a
Modernización Tecnolóxica de Galicia
(Amtega)

Mar Pereira Álvarez

O reitor magnífico da Universidade de
Santiago de Compostela

Antonio López Díaz

Anexo:**Plan de traballo para o ano 2022**

Realizaranse os seguintes seis subproxectos:

- Síntese de voz
- Recoñecemento da fala
- Xeración automática de textos
- Sistemas de diálogo
- Tradución automática
- Corrección e avaliación lingüística

As tarefas correspondentes desenvolveranse cun alcance menor ao previsto para o proxecto completo, en canto á extensión e profundidade dos recursos lingüísticos e dos modelos computacionais. No referente aos demostradores, nas reunións técnicas mantidas coa AMTEGA acordouse desenvolver un demostrador prioritario, seleccionado de entre o catálogo de demostradores identificados previamente, e cuxo obxectivo é dar visibilidade ao traballo desenvolvido na elaboración de corpus lingüísticos e modelos de lingua:

Subproxectos	Demostradores propostos
Síntese de voz	Sistema de conversión texto-voz (TTS) de alta calidade e naturalidade

Ademais do demostrador, implementaranse os prototipos de probas que se indican de seguido, algúns dos cales se presentarán en congresos especializados:

- Prototipo de probas do **recoñecedor automático da fala**
- Prototipo de probas do **tradutor neuronal español-galego / inglés-galego**

- Prototipo de probas do **xerador automático de descrições textuais de datos en tempo real.**

Por último, crearanse tamén recursos textuais e de voz que se porán a disposición de toda as entidades interesadas (institucións, empresas, ...) en repositorios abertos con licenza libre. Concretamente, crearanse os seguintes corpus e repositorios:

- **Creación dun banco de voces** da lingua galega, por medio dunha plataforma áxil e flexible que permitirá rexistrar voces que len texto controlado. A participación neste recurso será promocionada mediante unha campaña de difusión nos medios de comunicación.
- **Corpus de voz**, de utilidade para síntese
- **Corpus de audio e texto aliñados**, de utilidade para o recoñecemento da fala
- **Corpus textual** do galego a escala web (250G), de utilidade para todos os subproxectos textuais
- **Corpus paralelos galego-castelán e galego-inglés**, de utilidade para o subproxecto de tradución automática
- **Corpus de dominio específico para a xeración automática neuronal de descrições textuais de datos**

Os corpus desenvolvidos non só serán de utilidade para acadar os obxectivos do Proxecto Nós, senón que tamén, ao dispoñibilizárense en aberto, poderán servir para xerar modelos de lingua integrables en sistemas e aplicacións de empresas interesadas en utilizar o galego nas súas tecnoloxías. Tanto os demostradores como os corpus lingüísticos non serán produtos finais, pois continuarán a mellorarse incrementalmente en anos sucesivos con novos desenvolvementos e axustes. Dado que un dos obxectivos estratéxicos é crear e fortalecer unha industria galega de tecnoloxías lingüísticas, desde o arranque do proxecto, o

diálogo coas empresas de sectores afíns será prioritario, e concretarémolo coa organización de reunións, seminarios e xornadas, así como coa participación destas en estudos e informes.

CRONOGRAMA DE FITOS

Subproxecto ou WP	Fito	Data
Obtención de recursos lingüísticos	Novo corpus web para o galego construído coas ferramentas CCNet, dispoñíbel para a comunidade.	Xuño 2022
Construción e avaliación de modelos	Presentación de modelos semánticos neuronais para o galego	Xuño 2022
Obtención de recursos lingüísticos	Corpus de avaliación sintáctica de modelos neuronais para o galego.	Xuño 2022
Tradución automática	Presentación do prototipo de probas do tradutor neuronal español-galego / inglés-galego	Xullo 2022
Recoñecemento de fala	Presentación do prototipo de probas do recoñecedor automático da fala	Xullo 2022
Xeración automática de textos	Corpus nun dominio específico para a xeración automática neuronal de descricións textuais de datos	Xullo 2022
Obtención de recursos lingüísticos	Presentación dun corpus de composicionalidade semántica para o galego , e avaliación de modelos neuronais para o seu tratamento computacional, en SemEval 2022 (NAACL).	Xullo 2022
Recoñecemento da fala	Creación dun banco de voces da lingua galega, por medio dunha plataforma áxil e flexible que permitirá rexistrar voces que len texto controlado. A participación neste recurso será promocionada mediante unha campaña de difusión nos medios de comunicación.	Novembro 2022
Obtención de recursos lingüísticos	Corpus textual do galego a escala web (250G), de utilidade para todos os subproxectos textuais	Novembro 2022
Obtención de recursos lingüísticos	Corpus paralelos galego-castelán e galego-inglés , de utilidade para o	Decembro 2022

	subproxecto de tradución automática	
Obtención de recursos lingüísticos para síntese de voz	Corpus de voz para a síntese de voz	Decembro 2022
Obtención de recursos lingüísticos para recoñecemento da fala	Compilación de corpus para o recoñecemento da fala (de utilidade tamén para outros subproxectos textuais)	Decembro 2022
Síntese de voz	Demostrador: Sistema de conversión texto-voz (TTS) de alta calidade e naturalidade	Decembro 2022
Xeración automática de textos	Presentación do prototipo de probas do Xerador automático de descrições textuais de datos en tempo real	Decembro 2022