

O galego nas buscas na rede

O desafío
da relevancia



Presidente de PuntoGal

Manuel González González

Vicepresidente: Camilo Regueiro

Secretaria: Edita de Lorenzo

Tesoureiro: Marcus Fernández

Vogais: Sergio Aramburu, Manuel Bermejo, Javier Fraga, Concha Losada, Moncho Paz

Director Xeral

Darío Janeiro Pereira

Textos e produción

Luis Antelo, Noa Expósito, Bárbara G. Vilariño, Cristina Ríos Prado

Aportacións de

Alejandro Arce, Bárbara G. Vilariño, Sergio Ramírez, Xabier Rolán, Anxo Sánchez, Saúl Vérez

Este informe foi publicado coa colaboración da Secretaría Xeral da Lingua.

1ª edición: A Coruña, 2024

© PuntoGal

Rúa Tabernas, 11

15001 A Coruña

Tel. 670226268

info@dominio.gal

www.dominio.gal

© As autoras e autores

Imaxe da capa: Xan Sabarís

Maquetación: Sabarís Media

Reservados todos os dereitos. A infracción dos devanditos dereitos pode constituír un delito contra a propiedade intelectual.s

Depósito legal: C 1735-2024

As buscas na rede: a posición do galego e a súa mellora	5
O lugar no mundo.....	6
Google e o galego: si pero non.....	7
Que é PuntoGal.....	9
O dominio da Administración galega	10
Presenza de PuntoGal en foros de interese.....	10
O galego na rede	12
O galego	13
Uso do galego nas TIC	16
Iniciativas empresariais.....	22
Proxecto Nós.....	22
Nova Economía da Lingua.....	23
Balance da situación.....	24
Escaso desenvolvemento tecnolóxico en galego.....	24
Invisibilidade do contido nos motores de busca.....	24
Lingua “non grata” na publicidade.....	25
Os buscadores.....	26
O mercado: historia, evolución e uso actual.....	27
Tipos de buscadores	27
A relevancia económica dos buscadores	28
Google: o buscador dominante.....	29
Os outros negocios e ferramentas.....	30
Trabas xudiciais	32
Os buscadores en dispositivos móbiles.....	33
Buscas e intelixencia artificial: novas perspectivas	35
Google e os resultados con intelixencia artificial.....	35
OpenAI e a súa incidencia nos resultados orgánicos	37
E que pasa con Galicia?	37
A detección do problema nas linguas minorizadas.....	39
O exemplo do catalán.....	41
O caso vasco	45
O posicionamento.....	46
A importancia socioeconómica das buscas	47
O SEO	49
O SEM.....	51
Recomendacións para webmasters de Google.....	53
Descubrir como ve Google a web.....	53
Comprobar as ligazóns.....	54
Comprobar como se usa JavaScript	54
Informar a Google cada vez que o contido cambie.....	54
Non esquecer poñer texto nas páxinas	54

Controlar o contido que ve Google.....	55
Habilitar os resultados enriquecidos do sitio.....	56
Filtrar resultados para un idioma en particular	56
Ter en conta as persoas que realizan buscas en varios idiomas.....	56
Manterse informada das novidades nas buscas	56
Situación nos outros buscadores.....	57
Bing (Microsoft)	58
Yahoo! Search	58
DuckDuckGo	59
O posicionamento do galego.....	60
Situación actual.....	62
Recomendacións das expertas reunidas por PuntoGal.....	63
Situación actual.....	65
Recomendacións das expertas reunidas por PuntoGal.....	68
Como optimizar os navegadores e buscadores en galego.....	71
Proxecto de monitorización	74
O proxecto catalán	76
A proposta de monitorización de PuntoGal	78
Requisitos cos que deberá contar o proxecto	81
Prototipo: solución aportada.....	81
Prototipo: hardware.....	82
Prototipo: software	84
Prototipo: termos de busca.....	86
Fase 2: seguintes pasos.....	87
Resultado proba piloto.....	88
Resultado proba piloto	88
Como funciona?.....	89
Comportamento dun sensor	90
Servidor central.....	91
Accións necesarias.....	92
Primeiros resultados do piloto.....	93
Próximos pasos e conclusións.....	94
ANEXO: Ferramentas TIC en galego	96
ANEXO: guías PuntoGal	100
Bibliografía e referencias.....	102

As buscas na rede: a posición do galego e a súa mellora

Cando buscamos calquera información en Internet, as usuarias tenden a interactuar con aquela que o buscador ofrece nos primeiros lugares: canto máis arriba apareza un resultado dunha busca, maior é a probabilidade de que faga uso del. De aí que sexa tan importante o aparecer ben posicionados nos buscadores de Internet.

O dominio .gal ten como principal finalidade darlle visibilidade a Galicia na rede e, de maneira moi especial, á súa lingua e á súa cultura, xa que se trata dun dominio lingüístico-cultural. Por iso, a asociación PuntoGal, encargada da xestión deste dominio, ten a obriga moral de esforzarse para que as súas usuarias teñan a máxima visibilidade nos buscadores de Internet e, de maneira moi especial, en Google polo seu nivel de implantación.

Tanto en Cataluña coma en Galicia tense a percepción de que as webs en catalán e en galego, respectivamente, aparecen en xeral peor posicionadas que as webs en castelán. Esta percepción fixo saltar a alarma en sectores preocupados pola normalización lingüística do catalán e do galego e en organismos gobernamentais responsables do fomento do uso destas linguas. Tanto en Cataluña coma en Galicia puxéronse en marcha iniciativas para comprobar se esta percepción se corresponde coa realidade, e, en caso de ser real, analizar por que se produce e deseñar estratexias para reverter a situación.

En Galicia, este traballo sobre o posicionamento do galego nos buscadores esta realizándoo PuntoGal, co apoio da Secretaría Xeral da Lingua da Xunta de Galicia. O estudo debe deseñar procedementos de seguimento do posicionamento do galego nos vindeiros anos, e debe, sobre todo, propoñer estratexias para melloralo. E neste aspecto é fundamental o Search Engine Optimization (SEO), é dicir, a aplicación de técnicas que melloren a posición que ocupa unha web nos resultados dos distintos buscadores.

Estamos seguros de que as recomendacións contidas neste estudo, tanto as referidas ao interior da propia páxina (on page) como a elementos externos que van máis aló dun sitio web (off page), terán unha influencia decisiva para o posicionamento do galego nas buscas de marcas, nas transaccionais e, sobre todo, nas informacionais, que constitúen aproximadamente o 80% das buscas dos usuarios.

Agradecemos á Secretaría Xeral da lingua da Xunta de Galicia a súa preocupación por este tema, porque a visibilidade en Internet é, no mundo actual, un elemento fundamental para a normalización da lingua galega.

Manuel González
Presidente de PuntoGal

O lugar no mundo

A mala noticia é que un número cada vez máis pequeno de corporacións multinacionais determinan o funcionamento da rede. Empresas sen rostro e con oficinas situadas a miles de quilómetros. Empresas que pensan en obxectivos económicos, algoritmos, datos de usuarias ou publicidade, pero non especialmente preocupadas polos dereitos dixitais, a desinformación, a privacidade ou as linguas minorizadas.

Meta, Alphabet, ByteDance, X Corp., Microsoft ou Apple constrúen unha realidade que determina as nosas vidas. Porque vivimos na sociedade en rede versión século XXI. Compramos, conducimos, informámonos, relacionámonos, estudamos ou realizamos trámites coas administracións mediante unha constelación de pantallas conectadas. Moitas veces non sabemos a onde, pero conectadas.

Neste complexo escenario, as linguas e as culturas buscan o seu sitio. De forma orgánica, impulsándose coa forza das súas falantes. Pero tamén loitando por el, baixo a ameaza da desaparición ou, no mellor dos casos, da irrelevancia. Este debe ser un obxectivo prioritario das administracións públicas, que precisa de recursos e de determinación. E, igualmente, dun esforzo conxunto como sociedade que abrangue tanto á cidadanía como ás empresas e institucións. Cada unha no seu ámbito, nas súas posibilidades, pero de forma activa e consciente.

Aínda así, a folla de ruta é difícil. A monitorización da presenza do galego en buscadores, aplicacións, sistemas operativos, redes sociais, sistemas de publicidade; a presión para conseguir que a lingua estea presente neles; a incorporación de novos contidos no noso idioma en Internet; ou a aparición continúa de novas tecnoloxías que cambian o campo de xogo -ou a sociedade enteira, como a intelixencia artificial- precisan de moitos ollos, moitas mans, moitos recursos, moito interese.

Este documento é, por iso, un gran de area. Ou mellor, un xermolo. Como parte fundamental da nosa vida dixital, usamos Google e outros buscadores a diario, dúcias de veces, para todo. Pero o famoso algoritmo determina en que posición aparecen os resultados. E tamén, o lugar do galego. Froito desa preocupación, PuntoGal inicia un proxecto a longo prazo de monitorización da lingua en buscadores. Para saber que está a acontecer nun aspecto tan relevante da nosa vida dixital, divulgar recomendacións, exercer presión con datos reais, sumar esforzos con outras linguas nunha situación semellante.

As páxinas que compoñen o informe son un primeiro paso, unha avaliación inicial do estado da cuestión e unha folla de ruta para os seguintes anos. Un traballo necesario para poñer en marcha o que virá despois. Unha iniciativa de PuntoGal que conta co apoio da Secretaría Xeral da Lingua e a colaboración doutras persoas, entidades e empresas. E do equipo e a directiva do dominio. Entre as entidades amigas, quero salientar á Fundació PuntCat, os primeiros en facer saltar a alarma sobre esta cuestión. Grazas a todas elas. E tamén agardamos cos brazos abertos e ánimo de colaboración a quen queira sumarse. Contacta con nós se estás interesada.

PuntoGal celebra a súa primeira década este 2024. Dez anos de funcionamento do rexistro que xestiona unha entidade creada en 2006. Unha entidade que axuda de forma decidida a manter a identidade do país na rede. O noso ADN é o compromiso coa lingua, a cultura e a tecnoloxía galegas. E este proxecto que iniciamos agora é puro ADN PuntoGal.

Darío Janeiro

Director xeral de PuntoGal

Google e o galego: si pero non

A relación de Google e a nosa lingua foi sempre complicada, xa que inicialmente o galego era insignificante para a compañía detrás do motor de busca, de xeito que non ofrecía os seus servizos na nosa fala, algo que por fortuna foi mudando co tempo, pero mantendo unha clara discriminación fronte a outras linguas ibéricas, até o punto de chegar a considerarse a Google como un dos grandes inimigos da nosa fala, ao situala nunha posición de inferioridade na Internet que podemos atribuír basicamente a un factor: falla de entendemento dos contextos culturais plurilingües.

Para entender mellor a situación podemos facer memoria, e remontarnos ao ano 2008, cando se fixo evidente un cambio no algoritmo para os resultados de busca de Google que soterraba os contidos en galego dun xeito alarmante. Resulta que para que os resultados de busca en galego figurasen nunha posición aceptable dentro dos resultados de busca en Galicia era preciso que os usuarios tivesen personalizado a lingua do buscador para forzar a súa versión en galego (algo que a meirande parte dos usuarios non sabe facer), do mesmo xeito que os usuarios tampouco forzan o galego no seu navegador web ou no seu sistema operativo, pois traballan nun contorno que xa entenden (o español) polo que non perciben a necesidade de configurar as súas ferramentas en galego.

Até antes do cambio do algoritmo de busca, cando realizabamos unha consulta de Google con texto en galego, os resultados en galego saían nas primeiras posicións, pero tras os cambios de funcionamento do buscador, para que os contidos en galego saísen nunha boa posición ante unha consulta en galego era preciso que o buscador estivese configurado en galego, porque en caso contrario atoparíamos antes unha morea de resultados en español promovidos a posicións que non correspondían ao seu peso real dende o punto de vista dun usuario galego.

Esta situación provocou que moitas páxinas galegas, con soporte de varias linguas, mudasen a configuración do seu espazo na web para darlle prioridade ao castelán, conseguindo así un mellor posicionamento, algo que resulta crucial especialmente no caso dos negocios.

Google estaba a forzar que a sociedade galega tivese que agochar a súa lingua na web se quería ter visibilidade, inducendo a colocar a nosa propia fala nunha posición secundaria para non recibir o castigo dun axente estranxeiro cun poder de influencia innegable.

Se as galegas tivéssemos ordenadores con Windows en galego, dispositivos móbiles en galego, navegadores web en galego e Google en galego, a situación podería mudar, pero actualmente ningunha desas cousas son especialmente doadas de conseguir, sendo especialmente triste o caso do buscador na nosa lingua, xa que Google ten contratado o dominio google.gal, pero non fai uso do mesmo, mentres que si emprega google.cat e google.es, deixando clara a posición de discriminación á sociedade galega.

No caso dos dispositivos móbiles a responsabilidade de Google é aínda maior, xa que son os promotores de Android, sistema operativo móbil que tardou tantos anos en ser consciente do plurilingüismo, que cando por fin engadiron soporte para o mesmo xa era demasiado tarde, facendo imposible vivir en galego a tecnoloxía do noso día a día. Se hoxe configuramos un móbil con Android colocando o galego como lingua principal e o español como lingua secundaria, en moitas ocasións atoparemos como o noso trebello chega a mesturar linguas na mesma pantalla, pero non dúas linguas que o usuario considera nativas, senón que cando un texto non está localizado ao galego... aparece en inglés! Mestura que pode acontecer incluso nunha mesma frase, dando lugar a unha sorte de *castrapinglis* difícil de xustificar. Deste xeito disuádese á meirande parte dos usuarios de configurar os seus dispositivos en galego, pois achega unha

experiencia de pouca calidade (especialmente se tempos en conta que ningunha plataforma social relevante ten aplicación móbil en galego).

Podemos ter así a percepción de que Google está a favor do galego se atendemos a que conta con versións na nosa lingua de moitas das súas páxinas web, pero logo vemos que nos elementos críticos (o nome de dominio do seu buscador, as aplicacións móbiles...) ignoran completamente a nosa lingua, deixando a balanza completamente desequilibrada na nosa contra.

Discriminación económica

Google non só discrimina ao galego a nivel de posicionamento nos resultados de busca, senón que tamén ten unha política empresarial contraria ás linguas minorizadas, que non acepta nos seus servizos máis importantes de cara aos medios de comunicación: Google News e Google Ads.

Isto quere dicir que un medio de comunicación que estea escrito unicamente en galego non pode figurar na plataforma de novas de Google, pero tampouco pode nin mercar nin vender anuncios na plataforma publicitaria da compañía. O caso publicitario é especialmente lesivo para os negocios galegos, xa que nin poden explotar os seus espazos publicitarios a través de Google, nin poden anunciarse nos espazos do catálogo da compañía de Mountain View, deixando aos negocios galegos nunha situación de grande incapacidade.

E a longo prazo?

Aínda que o posicionamento en buscadores é algo que aínda está lonxe de solucionarse, a situación actual para o galego na web é crítica, pero estamos nun escenario de cambio que é unha auténtica incógnita, de xeito que a situación pode virar para mellor ou para peor. Falamos da aplicación da intelixencia artificial ás buscas na web.

Os modelos de linguaxe actuais, unha vez adestrados, aínda que non se personalizaron para as linguas minorizadas, aprendéronas igualmente, polo que actualmente sistemas como ChatGPT ou Gemini poden interactuar na nosa lingua, e unha vez que estas tecnoloxías pasen a aplicarse aos resultados de busca (como fan Perplexity e SearchGPT) poderemos atopar unha situación totalmente diferente á actual (matando todo o traballo en SEO feito até o momento), na que o sistema poida eliminar a discriminación dos espazos web na nova lingua, ou xusto o contrario (pois a IA pode ignorar o factor lingüístico se fai as inferencias mediante traducións a unha lingua franca), dependendo moito tamén de se conseguimos que estas tecnoloxías se adapten ben ás nosas necesidades. Se os servizos futuros ofrecen máis prestacións en función da lingua na que nos comuniquemos con eles, poderemos caer nun pozo do que sería precisa moita axuda para saír.

Marcus Fernández

Xunta directiva PuntoGal

Que é PuntoGal

PuntoGal, creada en 2006 por quince socias fundadoras representativas da sociedade, conta hoxe con máis de 115 entidades de diferentes ámbitos da sociedade galega, cunha ampla e moi diversa representatividade. Durante a elaboración da candidatura sumáronse miles de sinaturas de apoio de cidadáns e o respaldo de centos de entidades galegas, que dende entón forman parte da iniciativa. As tres universidades galegas, o Consello da Cultura Galega, a Real Academia Galega, o Museo do Pobo Galego, colexios profesionais como o de Telecomunicacións ou o de Xornalistas, a Mesa pola Normalización Lingüística, a CRTVG, empresas significativas dentro do ámbito dos dominios como Dinahosting, ou centros galegos no exterior son algunhas das súas socias.

Despois dun longo proceso que requiriu a elaboración dunha complexa candidatura con informes xustificativos, económicos e técnicos, ICANN (o organismo que regula Internet e realiza a asignación de nomes e direccións, así como o seu funcionamento) aprobou o dominio e delegouno. O proceso precisou da implicación de varios gobernos da Xunta de Galicia de distinta cor política, das aportacións intelectuais de entidades como a RAG ou as universidades, e do traballo solidario e anónimo de moitas outras persoas e institucións.

O obxectivo conseguiuase en xuño de 2013, data formal da aprobación do dominio para a lingua e a cultura galega. O dominio foi delegado finalmente en abril de 2014, de xeito que PuntoGal se converteu nun operador de dominio de feito. Nese mesmo ano presentouse, o día previo ás Letras Galegas, a primeira web baixo un dominio .gal (www.dominio.gal) nun acto que se celebrou na Cidade da Cultura en Santiago de Compostela. O 25 de xullo, déronse a coñecer os 93 primeiros dominios que se activaron nunha fase de prelanzamento. A continuación, houbo una sucesión de fases de rexistro ata o día 2 de decembro dese ano, momento no que comezou a fase Aberta, que se presentou na Escola de Telecomunicacións de Vigo en 2014.

A partir dese momento, os dominios xa se puideron rexistrar sen ningunha restrición de elixibilidade. Os únicos criterios que deben ter en conta é que sexan páxinas con contido en galego, que estean relacionadas coa cultura galega ou que inclúan servizos en liña dirixidos á nosa comunidade.

O dominio da Administración galega

A entidade conta co apoio da maioría dos sectores da sociedade galega, o que quedou reflectido na declaración institucional que, en xuño de 2014, realizou o Parlamento de Galicia, por unanimidade. No texto salientábase "o grande logro que supón para o noso pobo a consecución do dominio .gal", e que "ter un dominio propio é un paso importante para a comunidade, que agora ten unha marca propia que identifica a nosa lingua e cultura na Internet".

Ademais, en maio de 2015, a Xunta de Galicia aprobou o decreto 74/2015 que regula a organización da presenza na Internet da Administración xeral e do sector público autonómico de Galicia baixo o dominio de nivel superior .gal, coa finalidade de proxectar unha "imaxe homoxénea na Internet da Administración xeral e do sector público autonómico de Galicia por medio do impulso e difusión da riqueza cultural, patrimonial, lingüística e social de Galicia baixo o dominio .gal".

Pouco a pouco, PuntoGal vai figurando en distintas leis galegas como dominio de referencia. Por exemplo, na Lei de Administración Dixital de Galicia 4/2019, do 17 de xullo, no artigo 58 menciona a .gal como o dominio de nivel superior baixo o que estarán os portais web dos servizos ao público. Así mesmo, as contas de correo electrónico persoais, departamentais ou corporativas empregarán o dominio de correo de nivel superior .gal.

PuntoGal mantén dende a súa creación unha política de participación activa en entidades e institucións representativas do sector, tanto para expoñer o punto de vista do sector galego como para coñecer as tendencias no mercado e na regulamentación tecnolóxica relacionada co ámbito da rede e dos dominios en particular.

Presenza de PuntoGal en foros de interese

Galicia é das poucas comunidades a nivel mundial que conta cunha entidade que xestiona un dominio propio, distintivo para as páxinas web con vinculacións co país. Como tal, PuntoGal está recoñecida por ICANN como delegada dun dominio que conta, a día de hoxe, con máis de sete mil rexistros.

A existencia de PuntoGal permítelle a Galicia estar presente neste foro e aportar o seu punto de vista específico de dominio sen ánimo de lucro para unha comunidade cultural e lingüística. Ao mesmo tempo, permite o contacto permanente e as actividades sobre os temas que preocupan á comunidade da rede, como os problemas da ciberseguridade e a súa prevención, a aceptación universal de novas linguas e alfabetos, a inclusividade da rede, desde a protección contra os discursos de odio ata as diferenzas de xénero na forza laboral nas industrias TIC, etc.

A participación de .gal neste tipo de organizacións esténdese ademais a outros ámbitos. Así, forma parte dende hai varios anos do Geotld Group, a asociación que reúne aos dominios xeográficos e culturais. A actividade do grupo céntrase en construír políticas de defensa das comunidades dentro da industria global da rede e en crear ferramentas comúns que axuden ao desenvolvemento dos seus membros.

Neste sentido son salientables os contactos co tecido das industrias TIC e audiovisuais galegas e a integración do dominio como fonte de información e como portavoz en foros internacionais. PuntoGal está integrados en asociacións empresariais como o Cluster TIC, o da Comunicación e mantén relacións contantes cos do ecommerce, o dos videoxogos ou asociacións profesionais coma as dos enxeñeiros de telecomunicación ou a dos informáticos.

PuntoGal forma parte, polo tanto, das entidades que traballan na gobernanza da rede, o que lle confire un estatus singular e privilexiado para coñecer de primeira man a evolución do sector, as oportunidades e ameazas, os procedementos técnicos, as especialistas na materia e, non menos importante, o acceso a datos e estatísticas dos dominios. PuntoGal non só xestiona un ecosistema de provedores e usuarias arredor dos seus miles de dominios, senón que ten os coñecementos e a experiencia para obter e analizar información de distintas fontes aplicadas á saúde da comunidade onde realiza a súa actividade.

Esta posición privilexiada debe aproveitarse en favor de Galicia. Os estatutos da entidade deixan claro, e así se fixo dende a súa fundación, que o fin último é fomentar a presenza de Galicia e o seu idioma na rede, e dedicar todos os seus esforzos a mellorar dita presenza. Por este motivo, PuntoGal é a entidade sen ánimo de lucro que xestiona o dominio .gal, e considera que, ademais de aumentar a visibilidade do país, é preciso analizar o mapa actual coas ferramentas ás que pode ter acceso, como outro servizo a dita comunidade.-

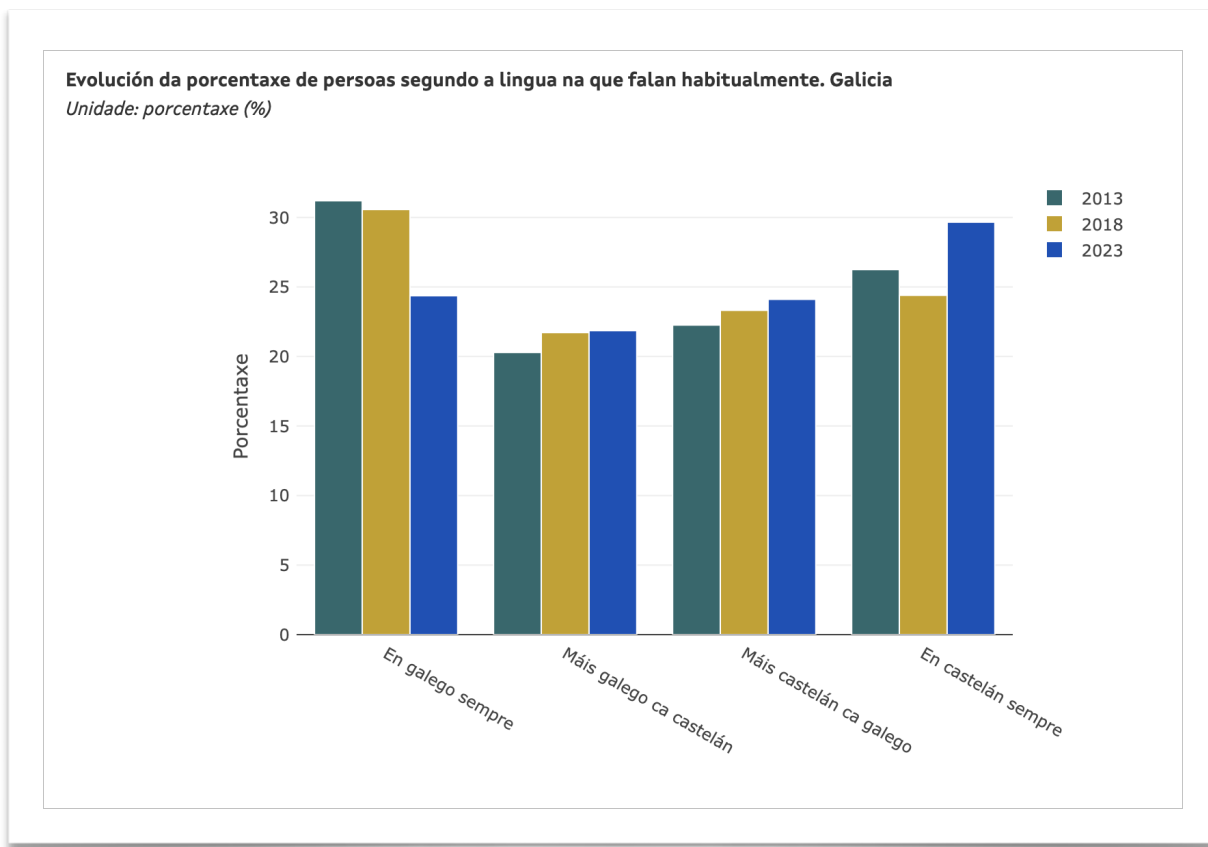
PuntoGal só concibe o seu traballo de xeito circular: os ingresos que recibe como administrador do rexistro do dominio .gal devólvense á sociedade galega en programas de colaboración, patrocinio ou investigación en campos como a cultura, o avance tecnolóxico e as actuacións comunitarias. Desde esta posición, PuntoGal recoñece a importancia que teñen os buscadores para garantir a relevancia e a sostibilidade económica de contidos e servizos na rede. Os seus algoritmos non sempre garanten as opcións minoritarias ou minorizadas, neste caso en razón da lingua na que se expresan. Por iso, co apoio da Secretaría Xeral da Lingua, abórdase este primeiro traballo no que se fai un diagnóstico da situación e se establecen as bases para a creación dun sistema que permita facer un seguimento no tempo para comprobar a evolución do tratamento do galego por parte das ferramentas máis usadas polo público.

O galego na rede



O galego

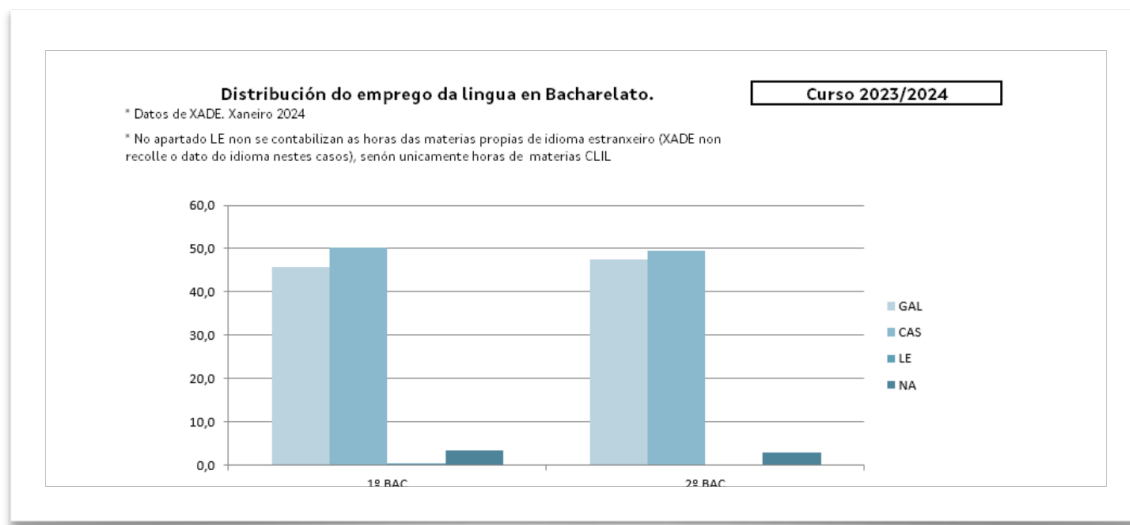
O galego é unha lingua romance con arredor de 2.400.000 de falantes, segundo datos do Observatorio da Lingua Galega (2023). Conforme cifras do Instituto Galego de Estatística, a poboación que declara falar habitualmente "en castelán sempre" chega ao 29,2%, cinco puntos máis que nos últimos datos publicados. A poboación que indica que "sempre" fala en galego cae por baixo do 24%, un 23,99% que, segundo a serie estatística consolidada polo IGE, supón unha caída superior aos 6 puntos. Coa data de outubro de 2024, as falantes habituais de castelán superan ás de galego por primeira vez na serie histórica.



Fonte: IGE. Enquisa estrutural a fogares. Coñecemento e uso do galego.

O galego é unha das linguas recoñecidas pola Carta Europea de Linguas Rexionais e Minoritarias do Consello de Europa, aínda que existen certos aspectos desa carta que nunca se chegaron a cumprir, especialmente en áreas como a educación, a administración de xustiza, a comunicación e o consumo. Neste contexto, Monteagudo (2024) sinala medidas como "garantir que o ensino en galego estea dispoñible en todos os niveis apropiados(...) establecer unha liña de ensino en lingua galega desde o primeiro ensino ata a Universidade; facer un seguimento pormenorizado do uso do galego nos procedementos penais, civís e administrativos; desenvolver un programa informático que contemple a posibilidade de utilizar formularios, modelos ou documentos en galego para asistir eficazmente no proceso

xudicial, así como proporcionar á cidadanía canles de comunicación coa administración de xustiza en galego. Como se bota de ver, o foco ponse, coma en informes anteriores, dunha parte na educación, e doutra parte no ámbito xudicial".



Distribución do uso de lingua no bacharelato, segundo o sexto reporte de avaliación da Carta Europea de Linguas Rexionais e Minoritarias.

O galego fálase no territorio da actual Comunidade Autónoma de Galicia e tamén nas áreas máis occidentais de Asturias, León e Zamora, así como na diáspora galega que emigrou a outras partes de España, América Latina, Estados Unidos e algúns países europeos, principalmente Suíza. Na actualidade, a lingua habitual de case o 90% dos aproximadamente 6.000 habitantes do Val do Ellas, no noroeste de Estremadura, segundo Haßler, G. (2006).

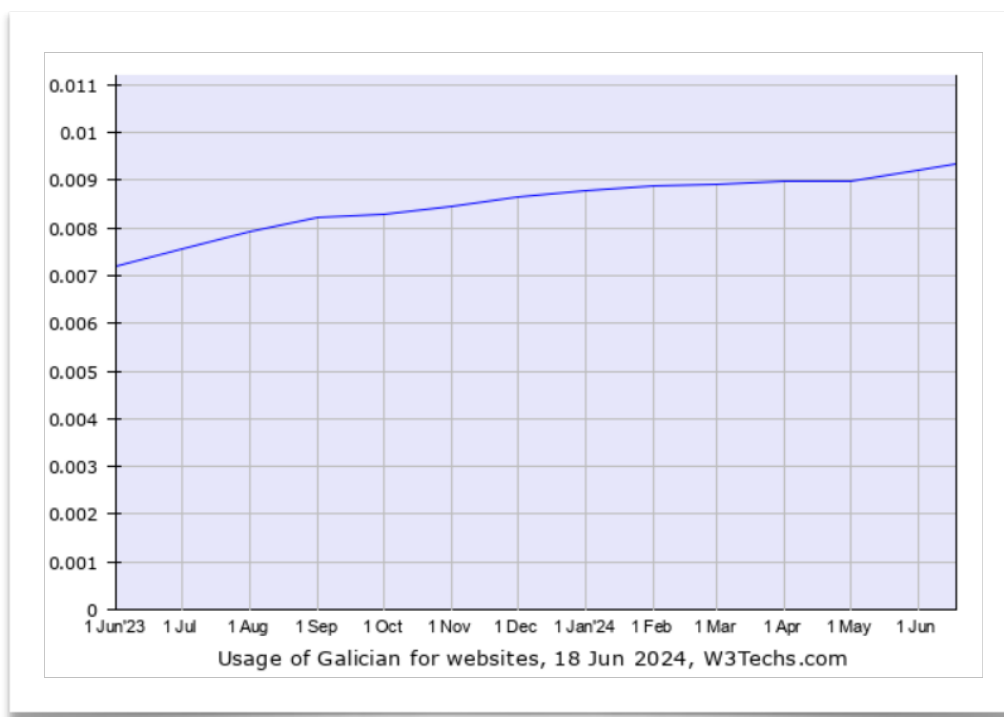
A situación legal do galego está amparada polas seguintes leis:

- Estatuto de Autonomía de 1981.
- Lei de Normalización Lingüística de 1983.
- Lei 5/1997, de 22 de xullo, da Administración Local de Galicia.
- O uso das linguas oficiais e do galego na UE, desde o Regulamento 1/1958 ata os acordos de 2009.
- A Carta Europea de Linguas Rexionais e Minoritarias.

A Lei de Normalización Lingüística (1983), na súa disposición adicional, atribúelle á Real Academia Galega a autoridade para establecer a variedade estándar da lingua galega.

En canto a situación do galego en Internet, está entre as 40 linguas máis usadas na rede, segundo Pérez, R. (2023, 9 de marzo). De entre todos os contidos dispoñibles en Internet menos dun 0.1 % está en galego. O éuscaro ten datos similares e o catalán é lixeiramente

superior a ambos, segundo datos do portal web W3techs.com no seu apartado de estatísticas de uso de lingua de contido nos sitios web, como se pode ver no seguinte gráfico:



Uso do galego en páxinas web a 18 de xuño de 2024, segundo w3techs.com.

Conforme o estudo de 2022 realizado polos investigadores franceses por Alain Calvet e Louis-Jean Calvet para o Ministerio de Cultura de Francia, *Barómetro das Linguas*, o galego sitúase no posto 37 de 634 linguas analizadas entre os idiomas máis infuñtes, tendo en conta factores como actividade dos idiomas na Wikipedia, o fluxo de traducións, a cantidade de premios literarios internacionais, a vehicularidade e a súa integración dentro do sistema universitario, entre outros. Onde acada a maior puntuación neste estudo é na súa presenza na Internet como na Wikipedia, ou Galipedia, que é como se coñece comunmente á esta plataforma en lingua galega, que comezou a súa andaina en 2003.

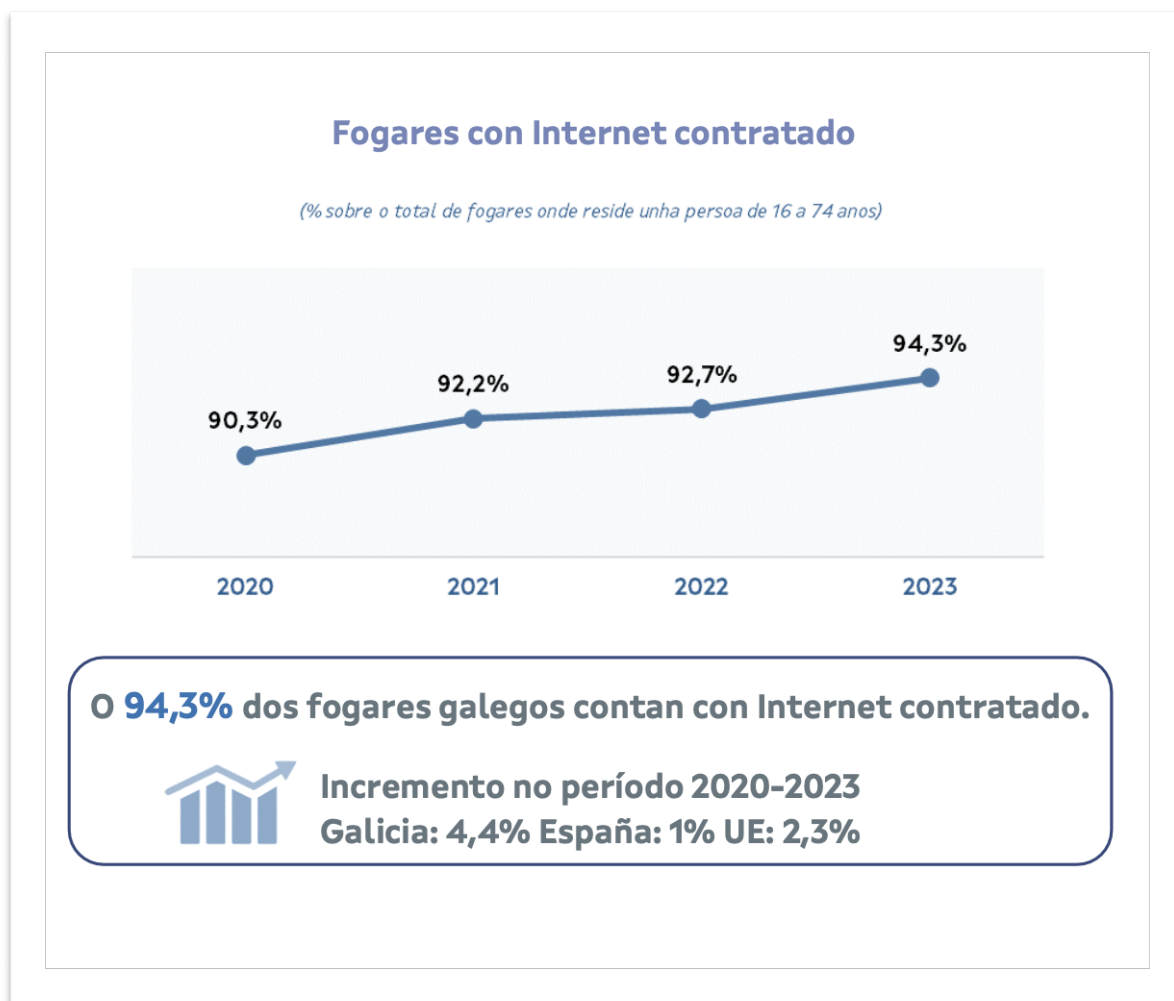
A lingua propia de Galicia conta con presenza nas redes sociais, con aportacións e publicación de contido en TikTok, Youtube ou Instagram. En Galicia existen creadoras de contido con miles de seguidores que publican e divulgan en galego, e que están a dotar de prestixio á rede no noso idioma, como confirmaba o lingüista do Instituto da Lingua Galega (ILG) Xosé Luís Regueira para Galicia Confidencial en entrevista en decembro de 2023: "o Tik Tok favorece un pouco máis o amosarlle aos nosos mozos e mozas a variedade dialectal tan rica que temos aquí en Galicia, favorecendo ao mesmo tempo unha mellor valorización da mesma por parte de todos os falantes".

A inmediatez das redes sociais fan que o acceso sexa universal; todo o mundo pode facer unha publicación en galego e chegar a un público que o demande, aínda que non haxa seguridade de que os algoritmos de busca fagan posible que esta información estea a chegar da forma correcta a todas as usuarias.

Uso do galego nas TIC

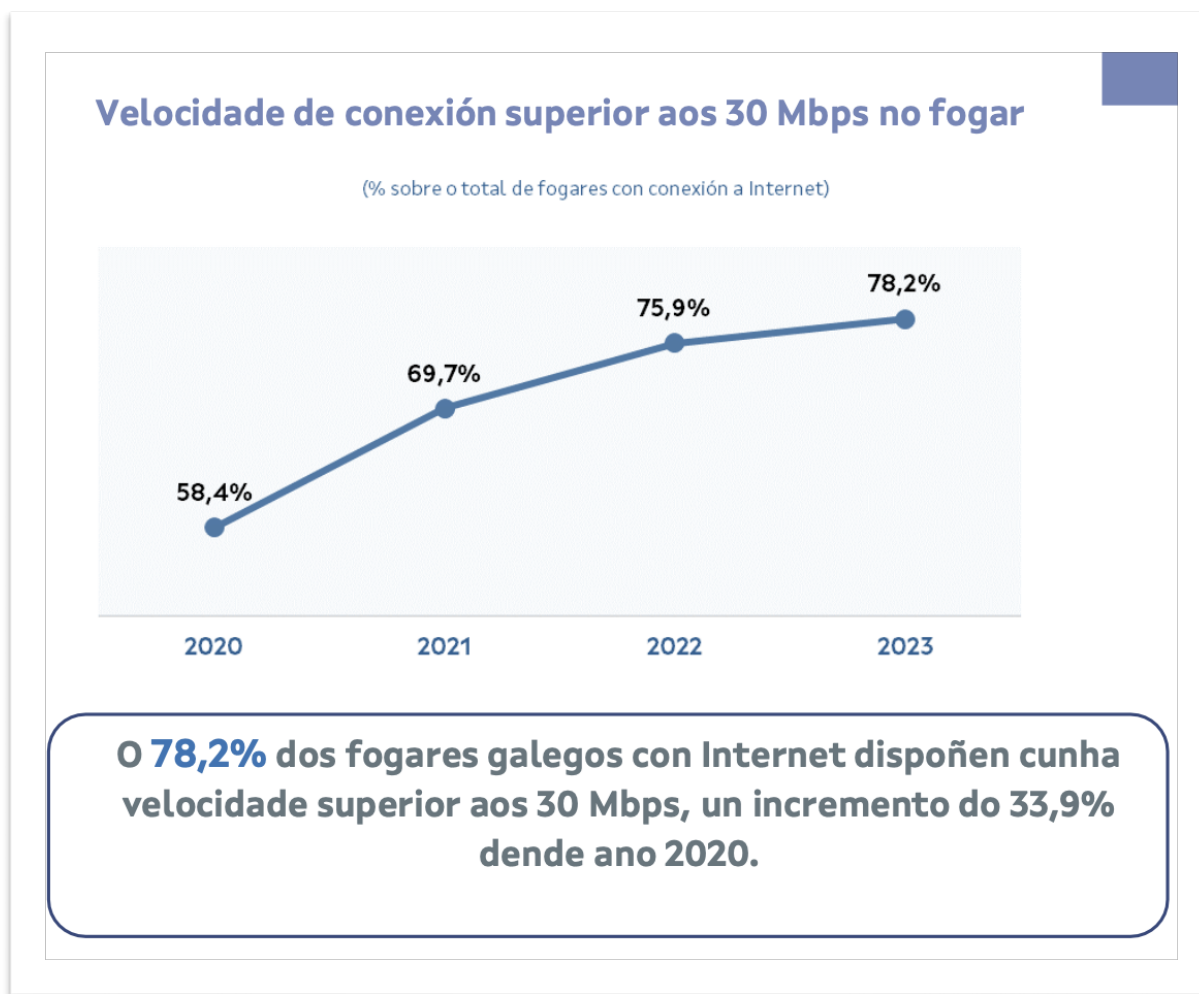
As chamadas Tecnoloxías da Información e a Comunicación son os recursos e ferramentas que utilizamos para procesar, administrar e distribuír a información a través de elementos tecnolóxicos. En Galicia, segundo o estudo de 2023 da Xunta de Galicia, *Galicia Dixital. O Hipersector TIC en Galicia*, no ano 2022 tiñamos arredor do 1.7% das empresas do sector TIC español, con preto de 29.000 traballadores, o que situaba a Galicia na quinta comunidade autónoma con maior número de compañías neste sector. O crecemento das organizacións do sector das Tecnoloxías da Información e a Comunicación considérase unha peza clave do desenvolvemento económico de Galicia debido ó seu potencial e futuro dixital.

Isto non quere dicir que non haxa un uso importante dos recursos tecnolóxicos tanto nos fogares coma nas empresas galegas. Nos fogares, o crecemento do uso da rede é máis grande que no resto de Europa, por exemplo.



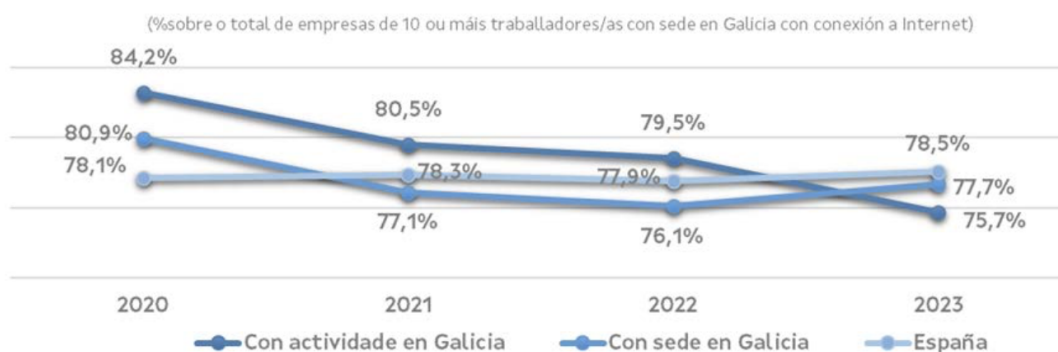
Evolución da porcentaxe de fogares con Internet contratado. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe sobre a sociedade da información nos fogares galegos (2020).

O mesmo se pode dicir da presenza de banda larga nas conexións dos fogares.

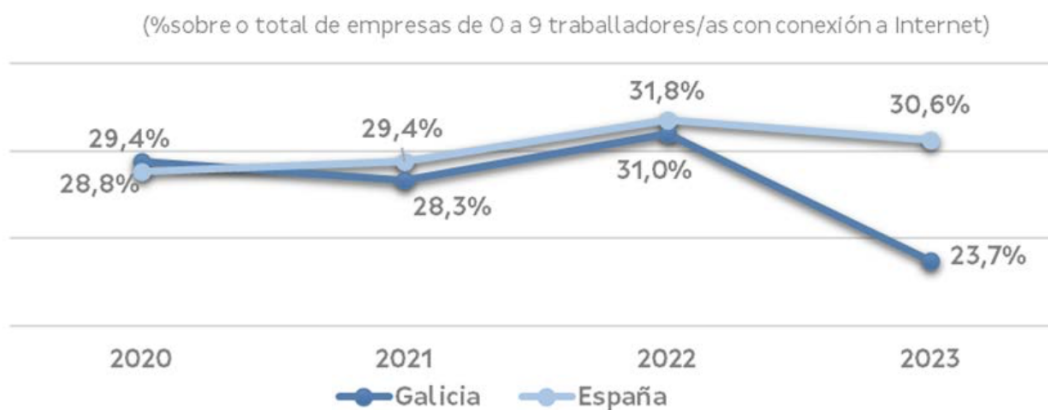


Evolución porcentual da velocidade de conexión dos fogares galegos. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe sobre a sociedade da información nos fogares galegos (2020).

Atópanse datos similares no que ten que ver coa adopción de novas tecnoloxías entre as empresas, observando distintos indicadores coma a presenza de equipamento, a existencia ou non de páxina web das empresas ou o uso de redes sociais para se manter en comunicación co público.

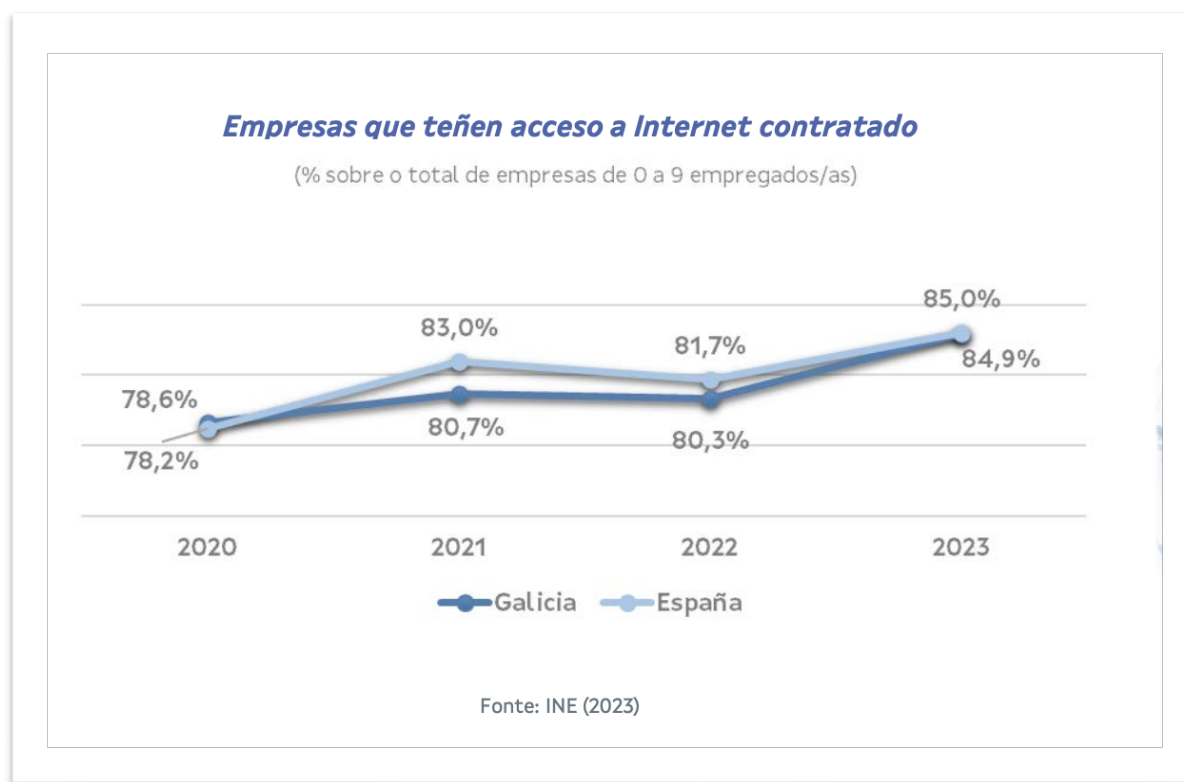
Dispoñibilidade de Páxina Web, Empresas de 10 ou máis empregados, evolución do indicador

Presenza de páxina web en empresas de máis de dez empregadas. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe das empresas TIC de Galicia 2023. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA).

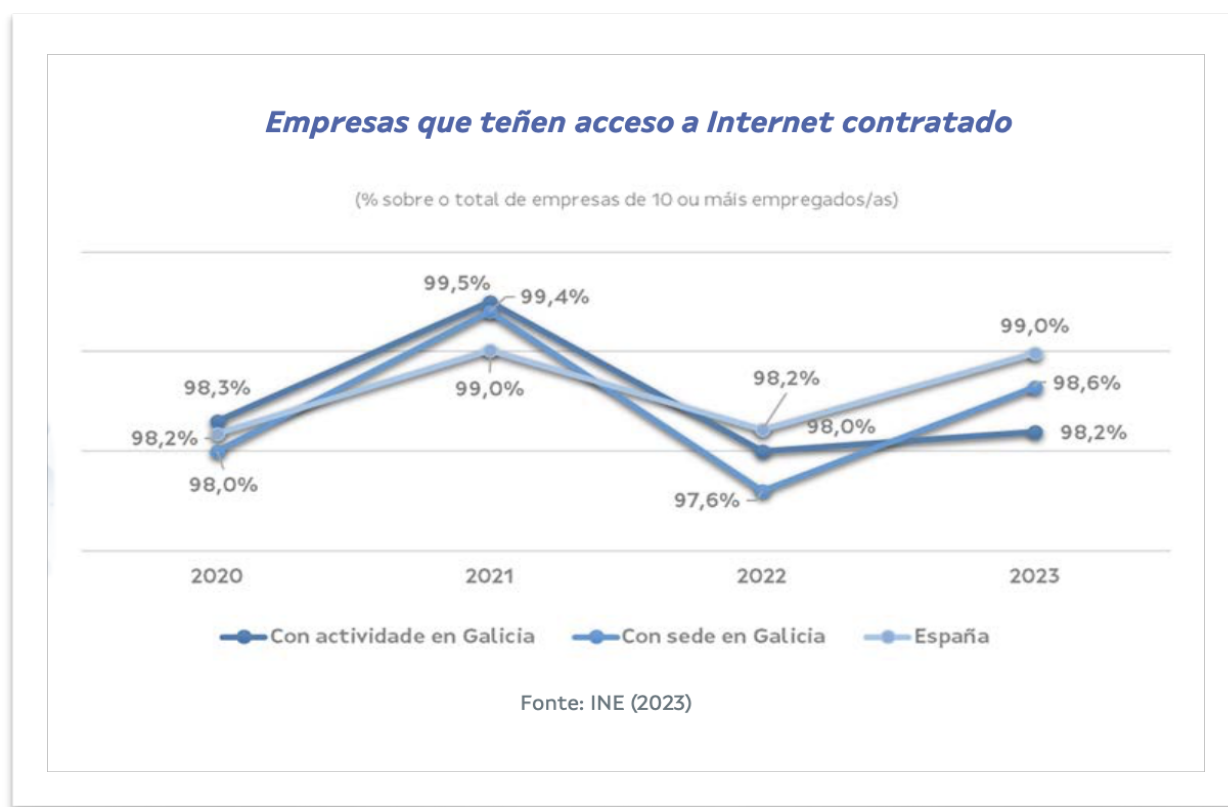
Dispoñibilidade de Páxina Web. Empresas de 0 a 9 empregados, evolución do indicador

Presenza de páxina web en empresas de ata dez empregadas. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe das empresas TIC de Galicia 2023. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA).

En canto ao uso no traballo diario das empresas, sen limitacións á comunicación externa, pódese ver que a evolución é moi positiva e que cada vez se aprecia máis coma unha parte básica do equipamento de traballo.



Presenza de conexión web en empresas de ata dez empregadas. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe das empresas TIC de Galicia 2023. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA).

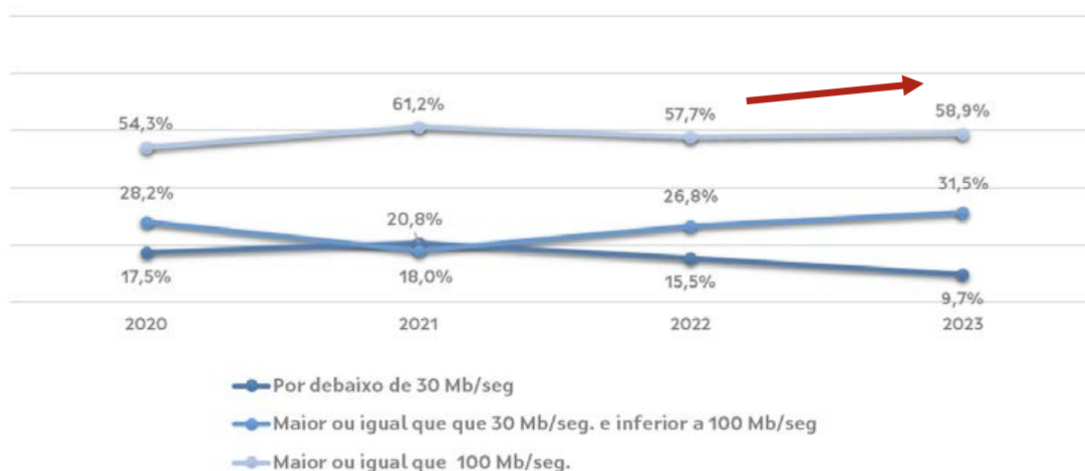


Presenza de conexión web en empresas de máis de dez empregadas. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe das empresas TIC de Galicia 2023. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA).

De igual xeito que nos fogares, a capacidade destas conexións vaise facendo maior a medida que se dispón de máis servizos na nube, sistemas de uso remoto ou se permite ás traballadoras desempeñar as súas tarefas de xeito deslocalizado, ben sexa algúns días á semana ou de maneira permanente.

Evolución da velocidade contratada: 0 a 9 empregados/as

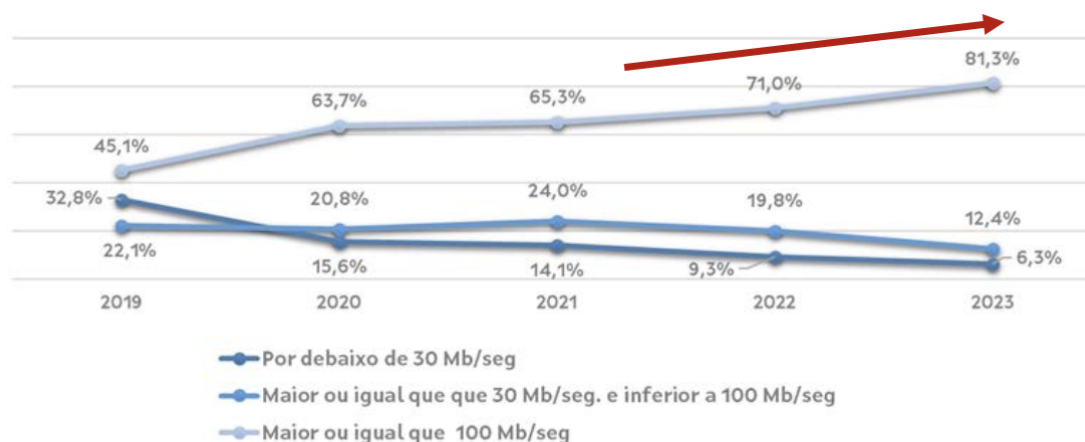
(% sobre o total de empresas de 0 a 9 empregados/as con conexión a Internet e banda larga fixa)



Fonte: INE (2023)

Evolución da velocidade contratada: 10 ou máis empregados/as

(% sobre o total de empresas de 10 ou máis empregados/as con sede e actividade en Galicia con conexión a Internet e banda larga fixa)



Fonte: INE (2023)

Velocidade de conexión web en empresas. Gráfico tirado do Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). Informe das empresas TIC de Galicia 2023. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA).

Iniciativas empresariais

As empresas do sector TIC reducen os seus esforzos en desenvolver tecnoloxías específicas para linguas minoritarias, conforme o recollido no informe de 2022 do *AtlanTic* da Uvigo para o *European Language Equality Report on the Galician Language*. Neste informe conclúese que existe un número moi reducido de recursos, produtos e servizos para a lingua galega, polo que é complicado que se vexa reflexado no seu uso nas TIC, sobre todo no referente ao big data e a intelixencia artificial. Aínda que destaca a existencia nos últimos anos de programas desenvolvidos para o uso do galego por parte das TIC como a web de consultas da Real Academia Galega, o tradutor Gaio, ou proxectos coma nós.gal, grupos de traballo de tradución de software libre, como proxecto Trasno, adicado á localización de software, Agasol (Asociación Galega de Software Libre), Mancomún, que busca a difusión de ferramentas baseadas en software libre en ámbitos de administración e empresa, etc, sigue a chamar a atención sobre a falta de recursos á hora de aplicar estas tecnoloxías,

Proxecto Nós

O proxecto Nós (nos.gal), lanzado en 2020, foi promovido pola Xunta de Galicia e executado pola Universidade de Santiago de Compostela a través de dúas entidades punteiras de investigación en intelixencia artificial e tecnoloxías da linguaxe -o ILG (Instituto da Lingua Galega) e o CITIUS (Centro Singular de Investigación en Tecnoloxías Intelixentes)-. Pretende promover a presenza dixital do galego coa creación de ferramentas e recursos de alta calidade para que non quede excluído dos novos desenvolvementos TIC. Unha das primeiras accións estivo a recollida de voces en galego e a súa validación, para facer un repositorio de modelos válidos orais na lingua de Galicia. Neste momento xa conta con algunhas ferramentas e demostradores, pero aínda están en fase de desenvolvemento. En particular, o proxecto céntrase na intelixencia artificial, a lingüística computacional e a xeración de recursos e ferramentas que faciliten a interacción en galego coas tecnoloxías dixitais, contribuíndo así á preservación e promoción da lingua galega no ámbito dixital.

O Proxecto Nós xorde nun contexto no que a maioría das ferramentas tecnolóxicas e dixitais dispoñibles están dominadas por linguas maioritarias, como o inglés ou o castelán. O galego, como outras linguas minorizadas, enfróntase ao risco de quedar relegado ou excluído das novas dinámicas de comunicación dixital. A USC, a través desta iniciativa, busca situar o galego á vangarda do progreso tecnolóxico, facendo que sexa unha lingua plenamente integrada nas plataformas dixitais, tanto a nivel de uso común como no ámbito académico e profesional.

Un dos aspectos centrais do proxecto é o desenvolvemento de recursos e aplicacións que permitan a integración do galego en asistentes virtuais, sistemas de recoñecemento de voz, tradución automática, procesadores de texto, e outras ferramentas de IA. Isto inclúe a creación de grandes bases de datos lingüísticas, corpus en galego, e algoritmos de aprendizaxe automática que poidan comprender e xerar linguaxe natural en galego, algo fundamental para garantir que a lingua poida empregarse en tecnoloxías como asistentes de voz (tipo Siri ou Alexa), tradutores automáticos ou sistemas de recomendación de texto.

Así mesmo, o proxecto Nós xa ten en marcha unha constelación de utilidades enfocadas a face máis doada a interacción entre as máquinas e os humanos, que se compón de tres ferramentas básicas: un tradutor neuronal multilingüe, un recoñecedor de voz que entende o galego oral e unha aplicación de síntese de voz que a partir dun texto é quen de ler en galego.

Nova Economía da Lingua

Neste sentido, cómpre salientar o proxecto Nova Economía da Lingua. Este PERTE tenta desenvolver modelos multilingües de texto, voz e tradución automática que dean resposta ás necesidades da realidade social e estean en liña coa tecnoloxía actual e está articulado arredor de cinco eixes principais, dos que destacan:

Coñecemento en español e linguas cooficiais: Desenvolveranse corpus lingüísticos con textos variados (novelas, teatro, cine, noticias, cancións, linguaxe oral e de signos) e apoiarase o proxecto LEIA para crear unha base de datos da lingua española coas achegas das comunidades autónomas.

Intelixencia Artificial: Investiranse máis de 330 millóns de euros en proxectos como a creación dunha Rede de Excelencia en Intelixencia Artificial e Programas de Linguas accesibles, co obxectivo de desenvolver ferramentas dixitais para persoas con discapacidade, así como outras accións do Plan Nacional de Tecnoloxía da Linguaxe.

Industrias culturais: Apoiarase o desenvolvemento e mellora da produción audiovisual e do sector dos videoxogos dentro do Plan España, Hub Audiovisual de Europa, e dixitalizaranse contidos de museos e patrimonio cultural. Este apartado contará con case 70 millóns de euros.

Balance da situación

O galego, como outras linguas minorizadas, enfróntase a unha serie de dificultades significativas no ámbito dixital, especialmente a curto prazo. A pesar dos esforzos por integralo nas novas tecnoloxías, hai diversos factores que limitan o seu uso e fan complicada a súa presenza efectiva na rede. Estes desafíos, de carácter técnico, económico e social, son cruciais de analizar se se quere avanzar cara un maior desempeño nos espazos dixitais e evitar a súa exclusión.

Un dos problemas máis evidentes é a escaseza de ferramentas tecnolóxicas dispoñibles en galego. O desenvolvemento de software, aplicacións móbiles, sistemas operativos e dispositivos dixitais está dominado por linguas globais como o inglés e o castelán, o que provoca que moitas das principais plataformas non ofrezan soporte en galego. As grandes compañías tecnolóxicas, como Google, Apple ou Microsoft, adoitan priorizar a creación de versións para linguas con maior número de falantes e con maior valor comercial, relegando linguas como o galego a un segundo plano.

Escaso desenvolvemento tecnolóxico en galego

O desenvolvemento de tecnoloxías para o galego require un investimento considerable en investigación, infraestruturas e persoal cualificado. A curto prazo, a falta de financiamento suficiente e o apoio institucional limitado representan un obstáculo. A creación de ferramentas como tradutores automáticos ou sistemas de recoñecemento de voz en galego é un proceso complexo e custoso, que require grandes bases de datos lingüísticas e algoritmos adaptados.

Aínda que existen proxectos que traballan para mellorar a presenza do galego nas novas tecnoloxías, estas iniciativas adoitan depender de financiamento público ou privado que pode non ser suficiente. Ademais, a colaboración co sector tecnolóxico privado, que adoita estar centrado en linguas con maior proxección comercial, non sempre é doada.

Isto significa que ferramentas esenciais como os correctores ortográficos, tradutores automáticos, sistemas de recoñecemento de voz ou asistentes virtuais non están dispoñibles en galego ou teñen un desenvolvemento moi limitado. Sen estas ferramentas, as falantes de galego non poden interactuar de maneira eficaz coas tecnoloxías dixitais na súa lingua, o que reduce a súa motivación para empregala na rede e forza a un uso crecente doutras linguas como o castelán ou o inglés.

Invisibilidade do contido nos motores de busca

Outra dificultade importante é a invisibilidade do galego nos motores de busca e nas plataformas de redes sociais. O contido en galego tende a ter menos visibilidade en buscadores como Google, que priorizan resultados en linguas máis usadas ou con maior relevancia global, como se verá neste informe. Isto provoca que o contido en galego teña dificultades para ser atopado polas usuarias, o que, á súa vez, desincentiva a creación de novos contidos.

Nas redes sociais, plataformas como Facebook, Twitter ou Instagram non sempre ofrecen opcións para configurar o idioma en galego, nin para crear contidos optimizados para a súa difusión nesta lingua. A ausencia de soporte en galego limita as posibilidades de interacción entre as falantes, e isto afecta negativamente á normalización da lingua no mundo dixital.

Tamén é un factor a percepción persoal do uso da súa lingua no ámbito dixital. Moitas veces, as usuarias optan polo castelán ou o inglés por consideralos máis "útiles" ou "eficaces" nas tecnoloxías dixitais. Esta actitude está influída pola omnipresenza destas linguas nas plataformas dixitais e pola falta de visibilidade do galego.

Ademais, a falta de demanda percibida por parte das usuarias tamén afecta ás decisións das empresas tecnolóxicas. Se as compañías non ven un interese claro por parte das galegofalantes en usar a súa lingua nas tecnoloxías, é probable que non invistan recursos en adaptalas para o galego. Isto crea un ciclo vicioso no que a pouca oferta e a escasa demanda se retroalimentan, dificultando o avance a curto prazo.

A pesar de que existen páxinas web, blogs e redes que promoven o uso da lingua, a produción de contido en galego na rede é limitada en comparación con outras linguas. Isto fai que as galegofalantes teñan menos opcións dispoñibles en galego á hora de consumir contidos, o que impulsa outras linguas. Aínda que iniciativas como a Galipedia (a Wikipedia en galego) ou a CRTVG ofrecen contido en galego, o volume total segue sendo reducido.

Lingua “non grata” na publicidade

A non consideración do galego como lingua publicitaria por Google e outros operadores da rede ten diversas consecuencias negativas, tanto para a lingua en si mesma como para a economía, a cultura e a identidade galega. Esta exclusión contribúe á súa marxinación nos espazos de comunicación masiva. As empresas adoitan optar por linguas dominantes, como o castelán ou o inglés, ao non ter incentivos nin obrigas para empregar o galego. Isto reduce a presenza da lingua nos anuncios en liña, o que provoca que a poboación galegofalante non reciba contidos publicitarios na súa lingua habitual. Como resultado, o galego perde visibilidade, especialmente entre as novas xeracións, que están máis expostas ao mundo dixital.

Desde un punto de vista económico, a falta de publicidade en galego significa unha perda de oportunidade para conectar de maneira máis efectiva co público local. As campañas en galego poderían crear un vínculo máis forte coas persoas galegofalantes, xerando maior fidelidade e identificación coa marca. Ao non recoñecelo, pérdese esta vantaxe competitiva, o que afecta directamente ás empresas galegas que poderían beneficiarse do márketing dixital en lingua galega.

No ámbito cultural e social, esta exclusión reforza a idea de que o galego é unha lingua secundaria ou inadecuada para contextos modernos como a Internet, minando o seu prestixio. Isto impacta negativamente na percepción que a sociedade ten da súa propia lingua e cultura, debilitando o proceso de normalización lingüística.

Os buscadores



O mercado: historia, evolución e uso actual

Os buscadores de Internet son ferramentas dixitais que permiten ás usuarias atopar información na web de forma rápida e eficaz. Funcionan como sistemas de recuperación de información que buscan, organizan e mostran contido dispoñible en sitios web de todo o mundo. Estes sistemas son esenciais para navegar polo vasto e complexo universo de datos que conforma Internet, xa que permiten acceder a información relevante con só introducir unhas palabras clave.

Os buscadores empregan programas automáticos, chamados robots ou arañas (en inglés, *crawlers* ou *spiders*), que recorren a web de forma continua. Estes programas "visitan" as páxinas web dispoñibles, seguindo ligazóns e explorando os contidos novos ou actualizados. Ao facer isto, o buscador vai recompilando toda a información das páxinas que atopan para saber que contido está dispoñible en cada sitio.

Unha vez que o robot rastrea unha páxina web, a información recollida almacénase nunha base de datos masiva que o buscador emprega para indexar os contidos. Este proceso de indexación implica organizar e clasificar os datos, asociando palabras clave e outros metadatos a cada páxina. Este paso é crucial, xa que permite ao buscador decidir que páxinas son relevantes para certas consultas e como deben mostrarse nos resultados.

Cando un usuario realiza unha consulta, o buscador examina a súa base de datos indexada e recupera as páxinas que considera máis relevantes. Para isto, emprega complexos algoritmos que teñen en conta múltiples factores, como a relevancia da páxina segundo as palabras clave, a calidade do contido, a popularidade do sitio, ou o número de ligazóns externas que apuntan a esa páxina. Despois, o buscador ordena os resultados e preséntallos á usuaria nunha lista xerarquizada.

Tipos de buscadores

Existen diferentes tipos de buscadores de internet, cada un cun enfoque distinto sobre como compilan e presentan a información. Os xeralistas son os máis coñecidos e utilizados, e permiten buscar calquera tipo de información na web. O exemplo máis destacado é Google, pero tamén existen outros como Bing ou Yahoo. Estes buscadores cobren unha ampla variedade de temas e fontes, e adoitan ter algoritmos moi avanzados que melloran a precisión dos resultados. Ademais, empregan técnicas de aprendizaxe automática e intelixencia artificial para optimizar a experiencia de busca.

Pero tamén existen buscadores enfocados en categorías específicas de información, como imaxes, noticias, vídeos ou produtos de compra. Por exemplo, Google Scholar, centrado en contidos académicos. Estes buscadores permiten unha experiencia máis dirixida a usuarias que buscan resultados nun campo particular.

Os metabuscadores, pola súa banda, envían a consulta a varios buscadores xerais ao mesmo tempo e compilan os resultados. Exemplos de metabuscadores son Dogpile ou Metacrawler. A vantaxe destes sistemas é que permiten obter resultados de distintas fontes, o que pode ampliar a variedade e a calidade da información recuperada.

Os resultados que ofrece un buscador non son aleatorios, senón que están determinados por unha serie de factores. Conta, por exemplo, a relevancia das palabras clave, a calidade do contido, a credibilidade da web na que se atopa a información, a localización e as preferencias persoais, entre outros factores públicos ou propietarios que constrúen os complexos algoritmos usados.

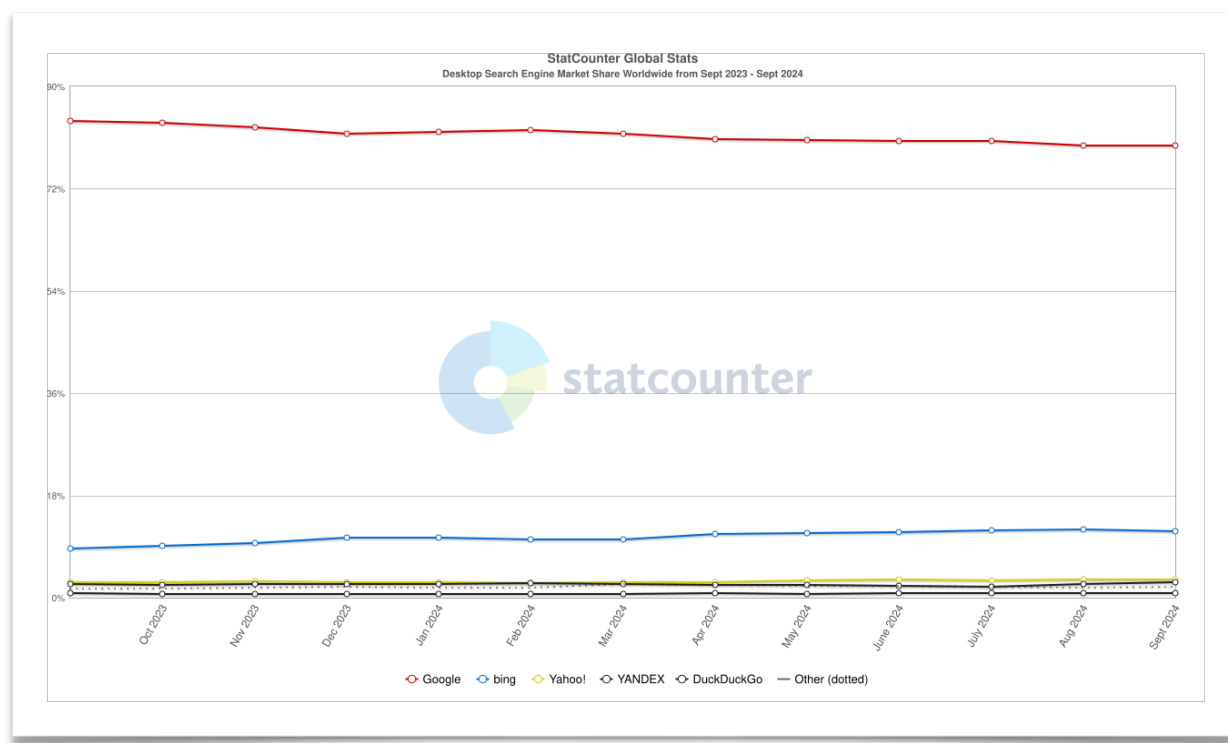
A relevancia económica dos buscadores

As buscas son unha das principais ferramentas para acceder a información, produtos, servizos e recursos en liña. Desde o punto de vista económico teñen un impacto enorme, xa que son o principal medio a través do cal as empresas captan clientes, especialmente no eido do comercio electrónico.

A publicidade nas buscas, como a que ofrece Google a través de Google Ads, é unha das formas máis efectivas de márketing dixital. As empresas invisten grandes cantidades de diñeiro para asegurarse de que os seus anuncios aparecen nos primeiros resultados de busca, o que incrementa a súa visibilidade e, potencialmente, as súas vendas.

Ademais, desempeñan un papel crucial na distribución de coñecemento e na democratización do acceso á información. A capacidade de buscar de forma rápida e sinxela contido de calquera parte do mundo transformou a maneira na que as persoas aprenden, investigan e comunican. No ámbito educativo, por exemplo, permiten o acceso a unha enorme cantidade de recursos académicos e científicos, ampliando o alcance do coñecemento.

Google: o buscador dominante



Comparativa de uso de buscadores entre setembro de 2023 e setembro de 2024, na que se pode apreciar a posición dominante de Google.

A primeira versión de Google foi lanzada en agosto de 1996, nunha páxina web propiedade de Stanford. Era un proxecto tan grande que utilizaba case a metade do ancho de banda da rede da universidade. Finalmente, en setembro de 1997, rexistraron o dominio google.com.

A finais de 1998 o buscador de Google contiña 60 millóns de páxinas e xa rivalizaba cos mellores buscadores do momento. En agosto de 2004 saíu a bolsa polo que o valor da empresa aumentou moi rapidamente. Comezou entón a comprar outras compañías e a desenvolver servizos adicionais (correo electrónico, redes sociais, etc.).



A páxina principal de buscar conta cunha versión en galego desde 2001.

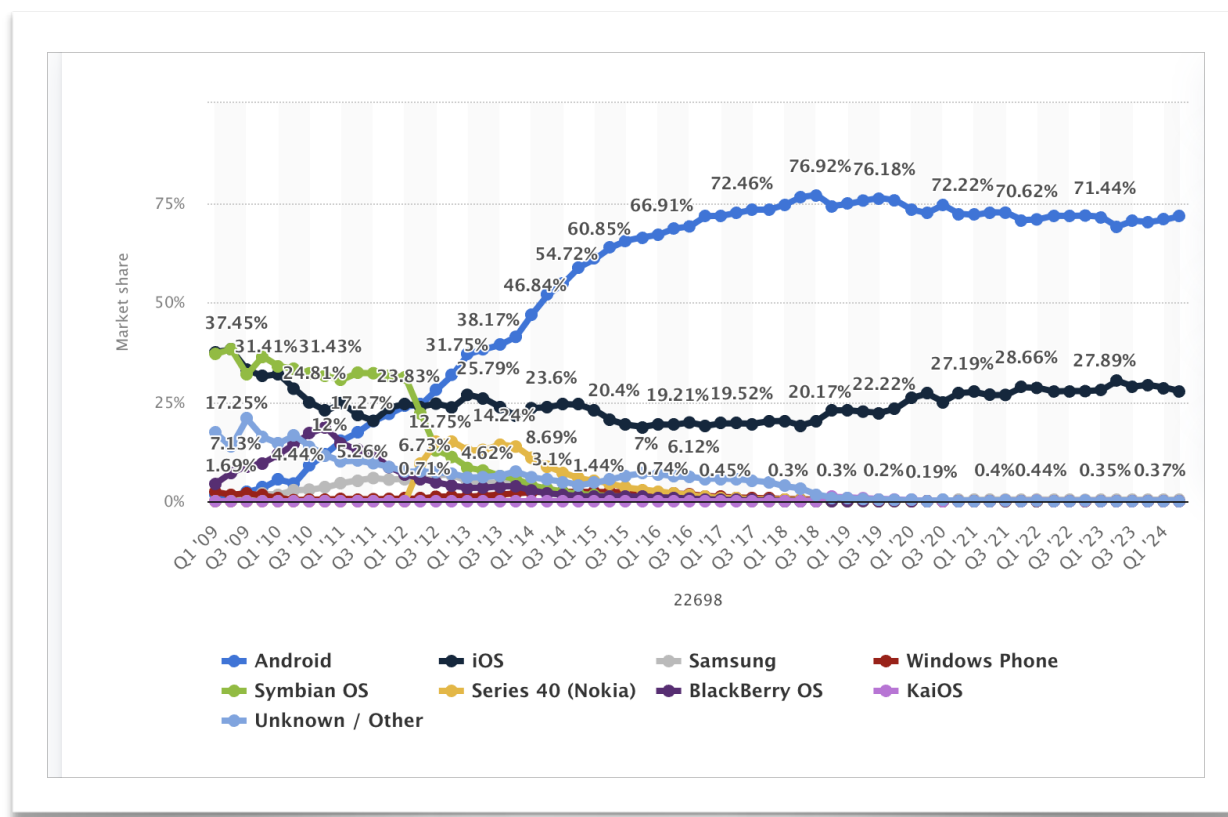
Google é unha das compañías tecnolóxicas máis influentes do mundo, recoñecida principalmente polo seu motor de busca, pero tamén polo seu amplo conxunto de produtos e servizos dixitais. Fundada en 1998 por Larry Page e Sergey Brin, Google comezou como un simple buscador de Internet, pero co paso do tempo expandiuse a moitos outros campos, incluíndo a publicidade en liña, os servizos en nube, os dispositivos móbiles e a intelixencia artificial.

O núcleo de Google segue sendo o seu motor de busca, que domina o mercado global con máis dun 92% de cota de uso, segundo datos de StatCounter (2024). Cada día, Google procesa arredor de 8.500 millóns de buscas a nivel mundial, facilitando o acceso á información para millóns de usuarias en cuestión de milisegundos. Este éxito baséase nos algoritmos avanzados que Google utiliza para ofrecer resultados relevantes e personalizados, tendo en conta a localización, o historial de buscas e outros factores.

Os outros negocios e ferramentas

Ademais do motor de busca, Google é unha potencia na publicidade dixital. A través de Google Ads, a plataforma publicitaria de Google, as empresas poden anunciarse de forma eficaz dirixíndose aos usuarios en función das súas buscas e comportamento en liña. En 2023, os ingresos publicitarios de Google foron de aproximadamente 224.000 millóns de dólares, representando case o 80% dos ingresos totais da compañía. Este modelo publicitario está centrado en anuncios de pago por clic e anuncios en liña, que aparecen non só nas buscas, senón tamén en sitios web que forman parte da rede de publicidade de Google.

Google tamén é responsable do sistema operativo Android, que ten máis do 70% do mercado global de teléfonos intelixentes, segundo Statista (2023). Android ofrece ás usuarias unha ampla gama de aplicacións a través da tenda de Google Play, que conta con máis de 2,6 millóns de apps dispoñibles. A través deste ecosistema móbil, Google mantén unha enorme presenza nos dispositivos de todo o mundo, conectando miles de millóns de persoas cos seus servizos dixitais.



Evolución dos sistemas operativos móbiles desde o 2009 ata o 2024, recollido por Statista.

Outro dos grandes negocios de Google é a súa plataforma de servizos na nube, Google Cloud. Este servizo ofrece ferramentas de infraestrutura e almacenamento a empresas e institucións, competindo con outras plataformas como Amazon Web Services (AWS) e Microsoft Azure. En 2023, Google Cloud representou aproximadamente 32.000 millóns de dólares dos ingresos da compañía. Aparte dos servizos de nube, Google tamén ofrece ferramentas de produtividade como Google Drive, Google Docs, Google Sheets e Google Meet, amplamente utilizadas tanto no sector educativo como empresarial.

Nos últimos anos, Google fixo grandes avances no campo da intelixencia artificial (IA). A súa división de investigación, Google DeepMind, é pioneira en IA, desenvolvendo tecnoloxías que van dende o recoñecemento de voz e imaxe ata a tradución automática e a conducción autónoma de vehículos. Os algoritmos de IA tamén son centrais nos seus servizos, como Google Assistant, que utiliza a aprendizaxe automática para mellorar as respostas a preguntas e ordes dos usuarios.

Trabas xudiciais

Google foi obxecto de múltiples sentenzas e investigacións xudiciais relacionadas coas súas prácticas monopolísticas, especialmente nos mercados de busca en liña e publicidade dixital. Estas sentenzas refírense ao control absoluto ou dominante dun mercado por parte dunha única empresa, o que pode limitar a competencia e afectar os consumidores. Moitas delas alegan que Google ten unha posición dominante que lle permite impor regras que prexudican a competidores e usuarios.

Unha das acusacións máis recorrentes é o abuso de posición dominante no mercado dos motores de busca. Google controla máis do 90% das buscas en internet a nivel global, o que fai que a competencia doutros motores como Bing ou Yahoo sexa case irrelevante. Este dominio permite a Google captar a maior parte dos ingresos por publicidade dixital, sobre todo a través da súa plataforma Google Ads, e controlar as condicións para anunciantes e editores.

Pola súa banda, a Comisión Europea multou a Google por favorecer o seu propio servizo de comparación de prezos (Google Shopping) nos resultados das buscas, prexudicando a competencia, en 2018. Tamén se cuestionaron os contratos entre Google e fabricantes de dispositivos móbiles, que obrigaban a preinstalar as súas aplicacións.

Estas sentenzas e acusacións contra Google culminan coa posible recomendación de fraccionar a compañía para ceder parte dos datos das usuarias a empresas coas que compite. Esta recomendación fíxose pública o pasado 8 de outubro de 2024 e aínda non está resolta.

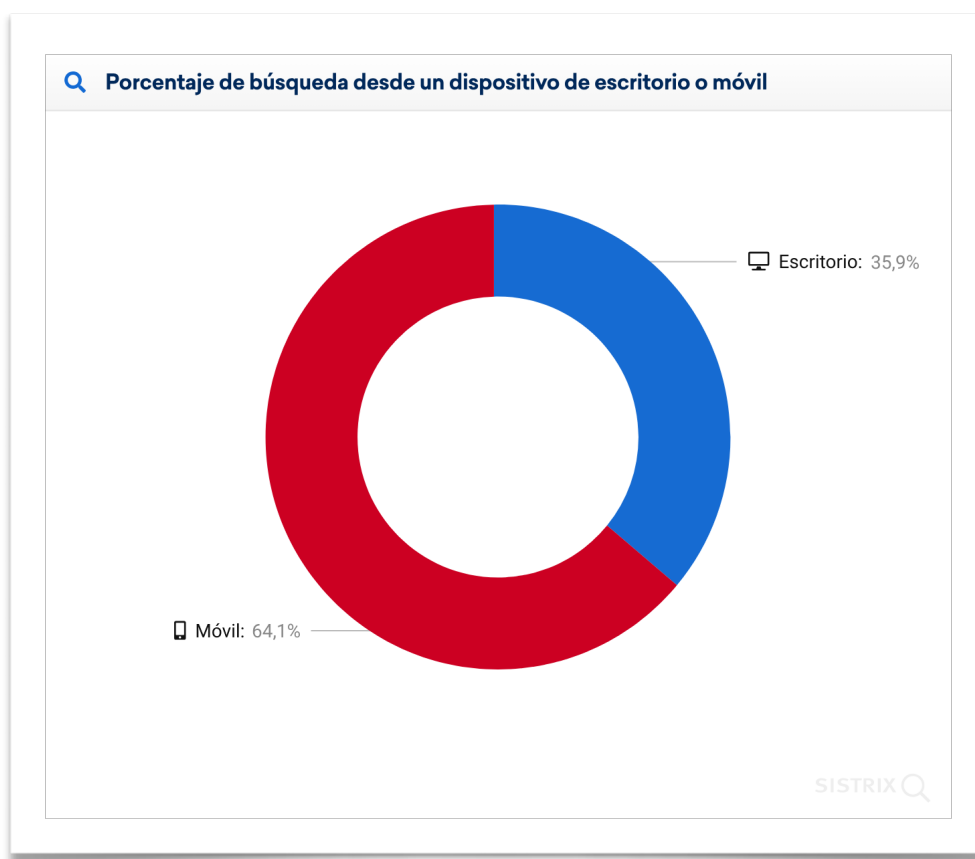
Os buscadores en dispositivos móbiles

O número de búsquedas diarias realizadas desde un teléfono móbil continúa crescendo, mentres as feitas desde ordenadores de mesa ou portátiles -*desktop*- están estancándose.

En 2021 Google cambiou o seu índice de búsquedas de Mobile First a Mobile Only. As versións para escritorio das webs deixarán de rastrearse e só se revisarán as versión móbiles. Actualmente máis de dous terzos das búsquedas realízanse desde dispositivos móbiles.

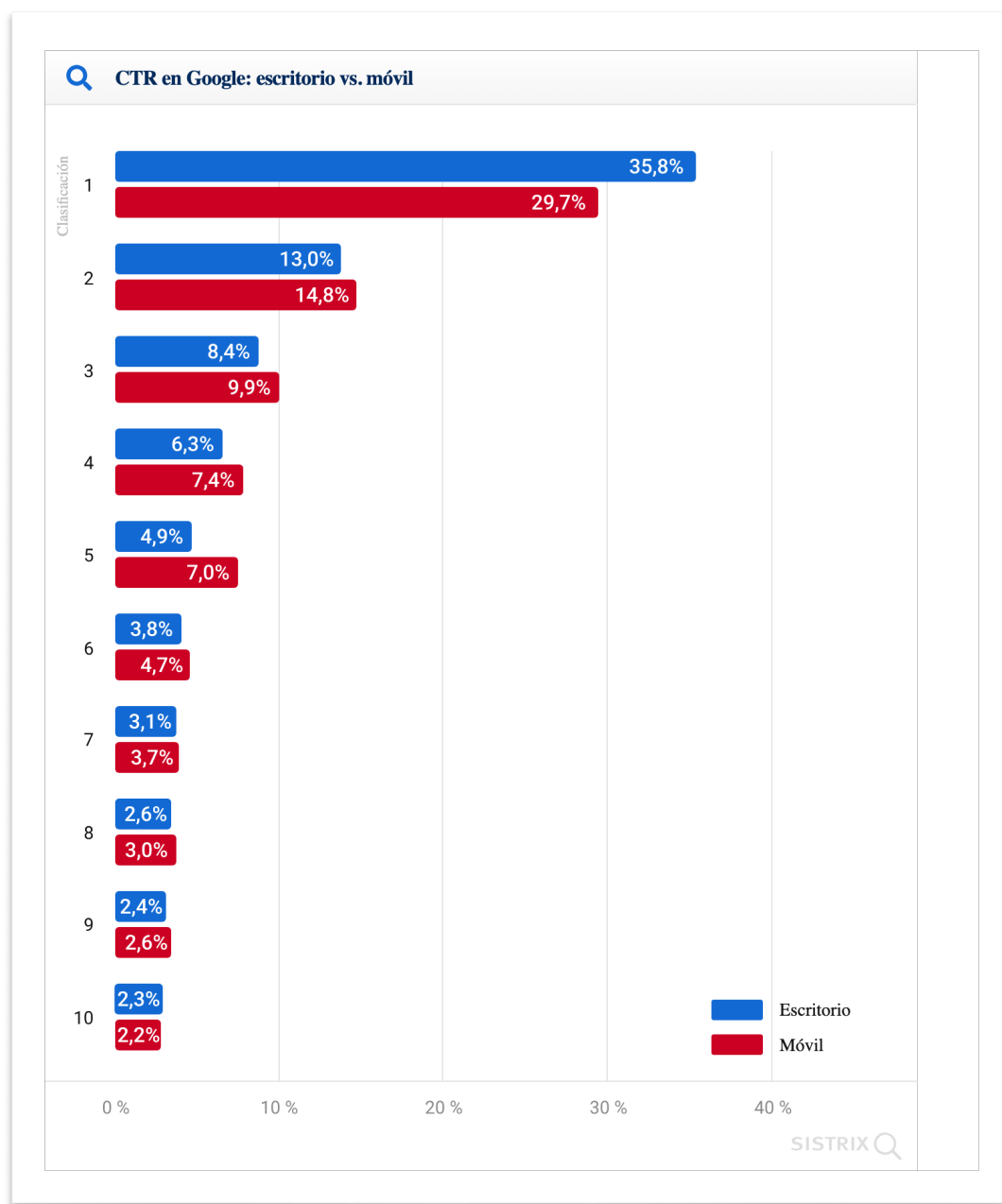
Nunha análise de buscas de 84.637.700 palabras clave só un 27% de búsquedas realizadas desde *desktop* prevalecen por riba das de móbil. Para estas palabras clave, as consultas realizadas desde dispositivos de escritorio conforman un 55% do total, mentres que desde os dispositivos móbiles estas alcanzan un 45% como máximo, segundo datos ofrecidos por Sistrix (2024).

Doutra parte, a proporción de termos nos que as búsquedas desde móbil teñen peso segundo a palabra clave é máis do dobre: casi un 57% de todas esas palabras clave son feitas desde móbiles. Ao redor do 15% de todas as palabras clave rexistran unha distribución máis ou menos equitativa en todos os dispositivos.



Porcentaxe de buscas desde móbil e escritorio, segundo Sistrix (2024).

O comportamento das usuarias tamén é distinto dependendo do dispositivo. Se están buscando desde o ordenador é moito máis probable que fagan clic directamente no primeiro resultado orgánico. Alí é onde acaba máis do 35% de todos os clics desde este dispositivo. No terreo móbil, menos do 30% dos clics aterriza no primeiro resultado orgánico. Ademais, nos resultados orgánicos faise clic máis a miúdo en móbil que en escritorio. De entre todas as palabras clave, o 82'5% das búsquedas desde o ordenador culminan nun clic, mentres que desde o móbil rózase o 85%.



Efectividade da publicidade de Google en escritorio e móbil, segundo Sistrix (2024).

Buscas e intelixencia artificial: novas perspectivas

A intelixencia artificial é un campo da informática que se centra no desenvolvemento de sistemas capaces de realizar tarefas que normalmente requiren habilidades humanas. Estas tarefas inclúen cuestións como o recoñecemento de voz, a toma de decisións, a aprendizaxe automática, a comprensión da linguaxe natural, a resolución de problemas e o recoñecemento de imaxes.

A IA baséase en algoritmos e modelos que permiten ás máquinas analizar datos, aprender deles e mellorar o seu rendemento sen intervención humana constante. Existen diferentes tipos de IA, pero un dos máis destacados é a aprendizaxe automática, que permite ás máquinas "aprender" a partir de grandes cantidades de elementos e identificar padróns ou tendencias. Dentro desta, a aprendizaxe profunda emprega redes neuronais, inspiradas no funcionamento do cerebro humano, para mellorar a capacidade de análise de información.

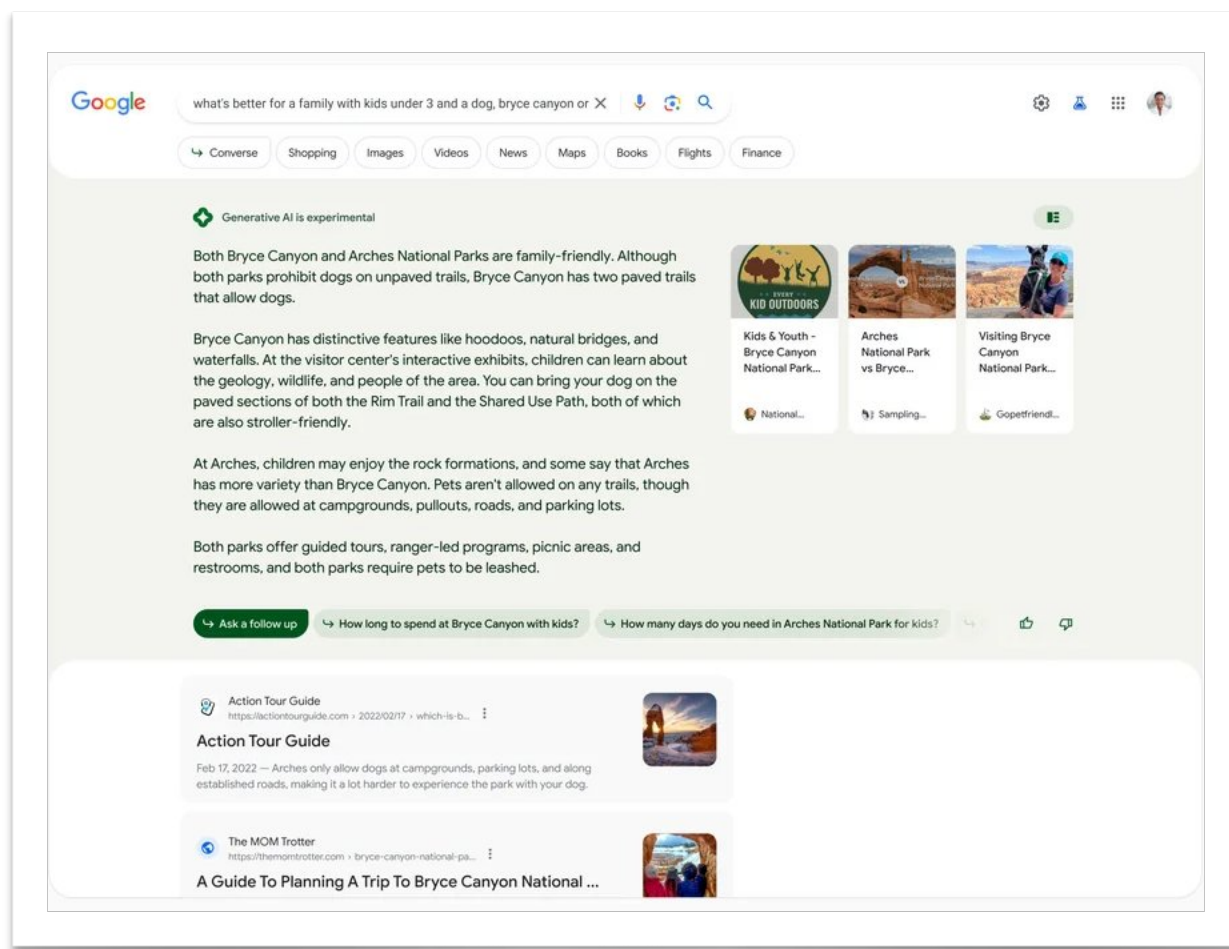
A IA xa se emprega en moitos aspectos da vida diaria, dende asistentes virtuais como Siri ou Google Assistant ata sistemas de recomendación en plataformas de streaming ou de comercio electrónico. Tamén é esencial en campos como a medicina, coa análise de imaxes diagnósticas, e o automóbil, cos vehículos autónomos.

Aínda que a IA ofrece grandes vantaxes, como a automatización de procesos ou a mellora da eficiencia, tamén xera debates sobre cuestións éticas, como o impacto no emprego ou o uso indebido de datos persoais.

Google e os resultados con intelixencia artificial

O buscador de Google é a porta de entrada a internet de millóns de persoas. Os cambios aí non son rápidos, aínda que en febreiro a consultora Gartner predecía que os buscadores perderían un 25% de usuarios en favor de chatbots e outros axentes da IA en 2026. Unha parte das búsquedas baséase en palabras clave. Agora esas palabras convertiranse nunha serie de valores que permitan ao algoritmo calcular como responder mellor a unha pregunta.

Durante a conferencia anual de desenvolvedoras de Google I/O en 2024 a empresa anunciou a súa aposta pola IA para mellorar os resultados de búsqueda grazas á ferramenta *AI snapshot*. Dende hai unhas semanas Google xa proba en EE UU estas *AI Overviews* (resumes con IA) para algunhas búsquedas: parágrafos xerados por IA que tratan de responder á búsqueda realizada. Estes resultados son opcionais, é dicir, as usuarias terán que actualalos.



Un exemplo de resultados de Google con aportacións de intelixencia artificial.

A chegada da IA ás búsquedas non elimina os enlaces. Google ofrece, debaixo dos resumos, os enlaces de onde se extrae a información, engade outros e crea unha lapela nova chamada 'web', onde as usuarias poderán ver só enlaces sen ningún tipo de IA.

Google non ofrece resumos en todas as búsquedas. Nas probas feitas parece máis probable que a IA apareza en búsquedas raras ou preguntas onde antes os enlaces non daban moitos resultados, e menos nas de carácter lucrativo para Google, como por exemplo "mellor cafeteira".

As búsquedas vinculadas a novas ou última hora son máis difíciles de automatizar. A IA necesita tempo para elaborar esos resumos e vai serlle difícil resumir un partido ou un bombardeo mentres está ocorrendo ou xusto ao rematar. Os medios son, porén, un dos grupos máis preocupados con estes cambios; os cales xa empezaron a xerar polémicas entre medios e buscadores. Nos primeiros días de *AI Overviews*, as usuarias encontraron unhas cantas preguntas nas que Google daba resultados erróneos ou ridículos. Nun caso máis grave polas súas connotacións, Perplexity contaba en aberto unha noticia exclusiva que Forbes publicara detrás do seu muro de pago.

Os cambios reais dependerán de como as usuarias van aceptando as novidades de Google, que é, ao final, o único que ten todas as métricas necesarias para medir a magnitude real de toda esta nova realidade.

OpenAI e a súa incidencia nos resultados orgánicos

OpenAI é unha empresa tecnolóxica dedicada á investigación e desenvolvemento de intelixencia artificial (IA). Fundada en 2015, a compañía traballa en diversos proxectos relacionados coa aprendizaxe automática, o procesamento da linguaxe natural e outras áreas da IA, ofrecendo ferramentas e modelos que melloran a capacidade das máquinas para comprender, aprender e xerar contido.

Un dos produtos máis coñecidos de OpenAI é o ChatGPT (Generative Pre-trained Transformer), unha serie de modelos de IA especializados na comprensión e xeración de texto. Esta IA xerativa é quen de realizar tarefas complexas, como responder preguntas, redactar textos e ata escribir código, grazas ao seu adestramento en enormes cantidades de datos. Este tipo de tecnoloxía ten aplicacións en campos como a asistencia virtual, o desenvolvemento de software, a educación e moitos outros.

Desde a aparición de ChatGPT, en novembro de 2022, o ruxerruxe sobre a fin do buscador tradicional é constante. OpenAI, ten en marcha proxectos con Microsoft para, entre outras cousas, mellorar o seu propio motor, Bing. Recentemente anunciaron que están a probar un buscador chamado SearchGPT, un prototipo con novas funcións de búsqueda de IA con acceso en tempo real a información de Internet. Esta nova ferramenta funcionaría coa tecnoloxía GPT-4 da empresa e conta coa colaboración de varios medios (The Wall Street Journal, The Associated Press, Vox Media e a revista Atlantic). Polo momento non hai data de lanzamento e só será accesible para 10.000 usuarias seleccionadas. Xa hai no mercado buscadores que funcionan con IA, como Perplexity, pero a súa cota é pequena.

E que pasa con Galicia?

A AESIA, Axencia Española de Supervisión da Intelixencia Artificial, é un organismo recentemente creado polo goberno de España que ten sede na Coruña. O seu principal obxectivo é supervisar o uso da intelixencia artificial, garantindo que se empregue de maneira ética, segura e respectuosa cos dereitos fundamentais das persoas. Encargarase de vixiar que o uso da IA se axuste á normativa europea, en particular ao Regulamento de Intelixencia Artificial da Unión Europea, que busca establecer un marco común para o desenvolvemento e aplicación de IA de forma segura. Entre as súas funcións clave está a avaliación de riscos, a protección dos dereitos dos cidadáns ante posibles abusos ou fallos nos sistemas de IA, así como promover un desenvolvemento tecnolóxico responsable e inclusivo. Está por ver a influencia do emprazamento desta axencia na Coruña no desenvolvemento da IA en Galicia.

A Xunta, pola súa parte, ten en marcha unha Estratexia Galega de Intelixencia Artificial 2030 que empeza co proxecto de lei de intelixencia artificial e que busca promover o desenvolvemento e uso ético e sostible da intelixencia artificial (IA) en Galicia, co obxectivo de mellorar a competitividade, a innovación e a calidade de vida da cidadanía. Esta estratexia céntrase en fomentar a investigación e o talento en IA, apoiar a transformación dixital das empresas, especialmente das pequenas e medianas, e mellorar os servizos públicos mediante o uso de tecnoloxías avanzadas. Ademais, promove o uso da IA en sectores clave como a sanidade, a educación e a administración pública, garantindo sempre a seguridade, privacidade e a inclusión social.

O pulo da IA non queda reducido ás iniciativas estatais ou rexionais. hai tamén asociacións nadas para impulsala e para moldear o seu crecemento cara a un uso máis social e responsable, como é o caso de Anfaia, unha Asociación Nacional Faro para la Aceleración de la Inteligencia Artificial, recentemente creada e con sede na Coruña.

A detección do problema nas linguas minorizadas

A partir do terceiro trimestre de 2022, numerosas internautas catalanofalantes observaron que o contido en catalán experimentara unha perda considerable de visibilidade nos resultados de busca web, mesmo cando o entorno de navegación estaba configurado para dar preferencia ao contido en catalán. A consecuencia era que, en igualdade de condicións, o catalán quedaba relegado nas buscas moi por debaixo da posición que lle correspondería pola cantidade de falantes.

Isto fixo que se puxera o foco tamén no galego, vasco e outras linguas minorizadas, onde se detectaron os mesmos fenómenos nos resultados das buscas. A revisión de logs e estatísticas de ferramentas como Analytics, desenvolvida polo mesmo Google, confirmou con datos o que, en primeira instancia, parecía unha impresión illada.

Para contrastar esta percepción, a *Aliança pola presenza dixital do catalán*, da que forma parte a Fundació PuntCat elaborou un estudo do que PuntoGal participou como entidade observadora. Esta iniciativa está impulsada por varias entidades para fortalecer e ampliar o uso do catalán no ámbito dixital, onde se inclúen asociacións, empresas e institucións que traballan conxuntamente para promover o catalán e fomentar que as tecnoloxías emerxentes soporten esta lingua. As principais conclusións do informe elaborado pola Aliança foron: que o 66'5% (máis de dous terzos) dos lugares web víronse (e vense) afectados polo problema e perderon tránsito nas súas versións en catalán. Ademais, hai unha forte correlación (unha media do 80%) entre o tránsito catalá e o castelán, o que non significa que a versión en castelán gaña unha páxina vista por cada unha que o catalán perde. Esta correlación é moito máis feble (0'25) entre o catalán e o inglés. Todo o proceso está documentado no [Informe sobre la pèrdua de visibilitat dels continguts en llengua catalana als resultats de cerca web](#).

Google comunicou estes cambios de xeito xeral nun post no seu blog de buscas.

Cómo gestiona la Búsqueda de Google las búsquedas en varios idiomas

Enviar comentarios

Viernes, 8 de septiembre del 2023

En muchos países y zonas del mundo, las personas suelen hablar y hacer búsquedas en más de un idioma. Para ofrecer una mejor experiencia, Google utiliza varios métodos para determinar automáticamente cuál es el idioma más adecuado y mostrar los resultados de la búsqueda en ese idioma.

Cómo determina Google automáticamente el idioma de los resultados de búsqueda

Es comprensible que algunos usuarios crean que la Búsqueda de Google solo muestra resultados que coincidan con su configuración de idioma, pero esta función no es tan útil como parece.

Los navegadores, los dispositivos móviles y los ordenadores tienen su propia configuración de idioma. Esto significa que es posible que el usuario tenga configurado el navegador en un idioma y su dispositivo móvil u ordenador en otro.

La Búsqueda de Google también tiene una configuración de idioma, pero se trata del [idioma de visualización](#), que es el idioma que se usa en los botones y en el texto del menú que aparece alrededor de los resultados de búsqueda. Los resultados de búsqueda no tienen por qué coincidir con el idioma de visualización. Y esto es por una buena razón. Aproximadamente la mitad de las personas que buscan con Google son multilingües y, a menudo, buscan en un idioma que no coincide con su configuración.

La Búsqueda de Google tiene en cuenta todos estos ajustes y otros factores para [determinar automáticamente en qué idiomas](#) sería más útil mostrar resultados. De esta forma, los usuarios que hacen búsquedas en varios idiomas no tienen que cambiar continuamente uno o más ajustes de idioma para obtener resultados en los diferentes idiomas que utilizan.

Por ejemplo, una persona de Francia puede buscar en francés, inglés o árabe y esperar obtener resultados en el idioma adecuado. Del mismo modo, debido a la dificultad para escribir con algunos teclados, un usuario de la India puede hacer búsquedas en hindi con

Desde aquela, algunhas entidades, coma a *Aliança per la presència digital del català* están implementando solucións para monitorizar a distribución lingüística dos resultados das buscas e publicando os resultados destas observacións.

O problema non afecta por igual a todos os sitios web. O informe examina diversos perfís comúns de sitios web, mesmo algúns que non se viron afectados de ningunha maneira. Moitos dos afectados son as organizacións en catalán máis visitadas e relevantes como as administracións, universidades ou medios de comunicación, así como sectores empresariais que publican os seus contidos web en catalán.

Este estudo chega cando a aparición dos chatbots da IA están cambiando a forma na que as súas usuarias buscan e interactúan co contido dixital, facendo imprescindible establecer o lugar do contido orixinal nestes idiomas non tan maioritarios antes de que os chatbots tomen o relevo da búsqueda web no hábitos das persoas.

Antes de que isto ocorrera, o contido en catalán de sitios web multilingües si que se podía encontrar nas primeiras posicións de búsqueda cando o mesmo buscador interpretaba que esta era a lingua na que a usuaria facía a consulta, xa fora polas configuracións e preferencias do usuario ou pola lingua da consulta.

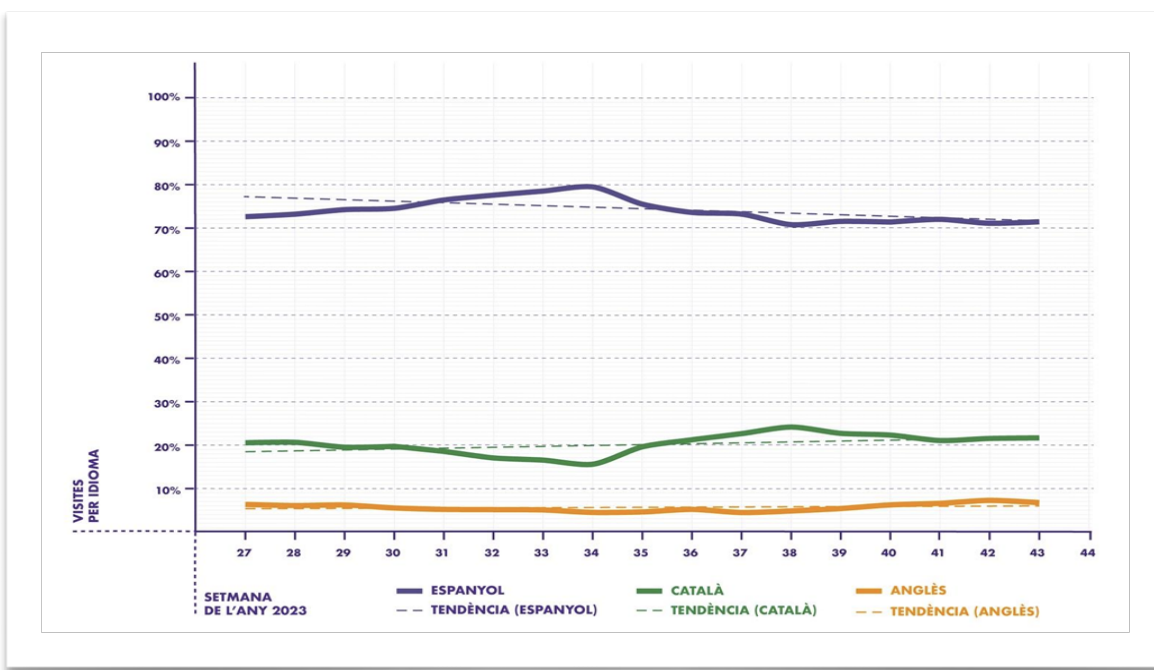
O exemplo do catalán

No estudo comparáronse as visitas ás versións en catalán procedentes dos buscadores coas visitas ás versións nos outros idiomas que cada sitio web ofrece, centrándose no tráfico procedente do buscador de Google. Nalgúns casos, o equipo da Fundación .cat puido acceder directamente ás ferramentas de análise de tráfico. Noutros, as súas propietarias fixeron a extracción de datos de acordo coas súas especificacións e facilitaron os resultados.

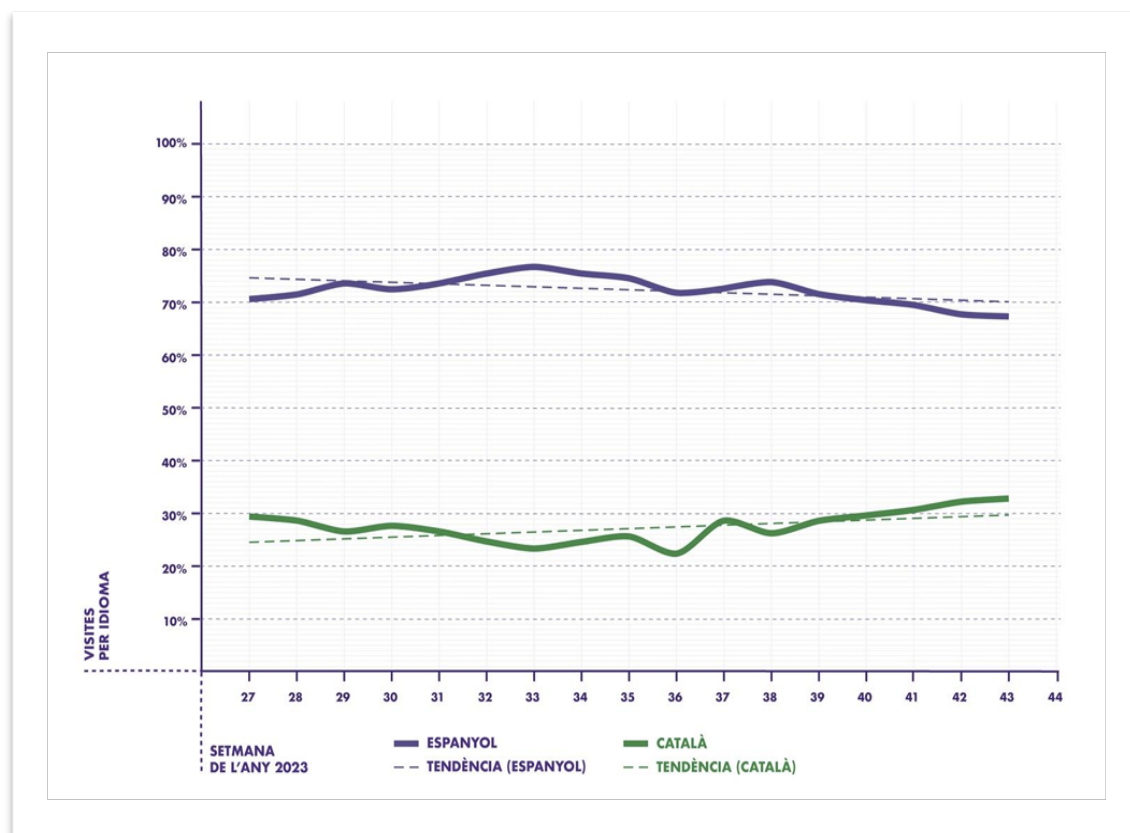
Para este estudo analizouse o tráfico web orgánico procedente de buscadores de 639 sitios web multilingües que inclúen o catalán e unha ou máis linguas. Todas as webs proceden de entidades e organismos dos sectores público, académico, mediático e empresarial cataláns que accederon a colaborar co estudo da *Aliança pola presenza dixital do catalán*, a petición da Fundación .cat que realizou a recolección de datos e que se encargou da elaboración do informe, coa axuda de todos os membros da *Aliança*.

Constatouse efectivamente que o cambio de tendencia do tráfico aportado desde o buscador de Google (no que se centra o estudo) detéctase de forma xeneralizada durante a primavera de 2022 e persiste ata a 2023. Segundo o estudo, o 66'5% dos sitios web analizados víronse afectados polo problema, perdendo así tráfico nas súas versións en catalán. Ademais, os buscadores deixaron de respectar as preferencias lingüísticas explícitas das usuarias. O contido en catalán perdeu visibilidade nos resultados das búsquedas con independencia da configuración de idioma do dispositivo, o navegador e o perfil de usuaria.

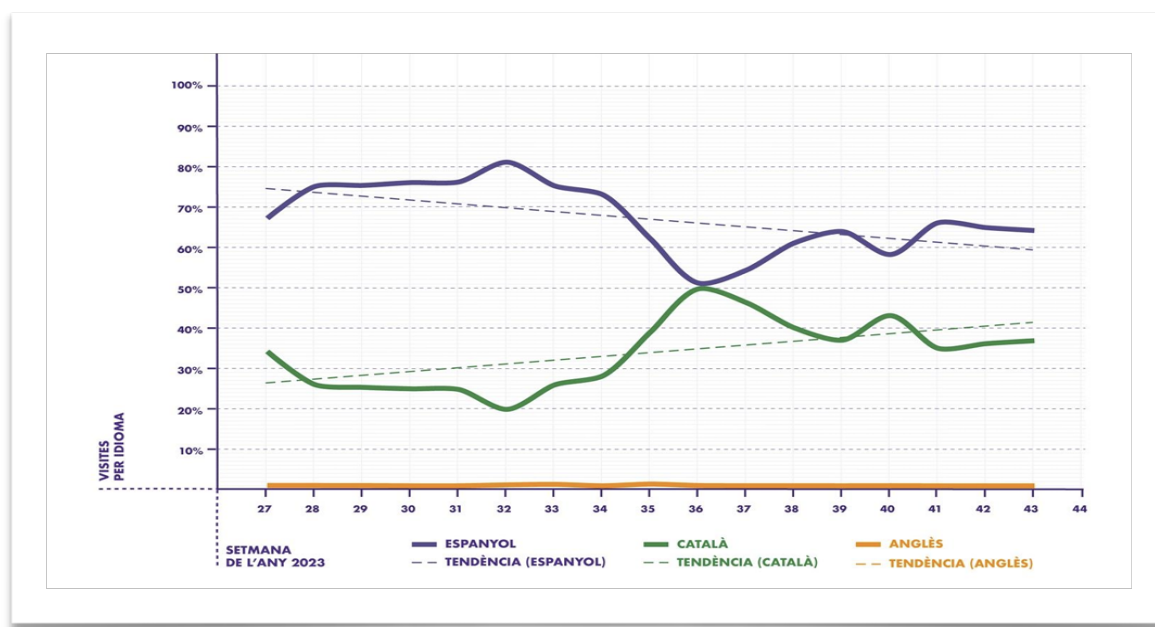
A continuación se poden ver algúns gráficos que reflexan a tendencia da presenza do catalá en distintas webs con versións multilingües. Todos estes datos poden ser consultados na [versión do estudo de maio de 2024](#):



Resultados dunha institución educativa superior, Aliança Digital de Catalunya. (2024). Informe de posicionament estratéxico 2024.



Resultados dunha empresa minorista, que vende ao público en tendas físicas e online, Aliança Digital de Catalunya. (2024). Informe de posicionament estratègic 2024.



Resultados dun medio de comunicación con versión en català, castelán e inglés. Aliança Digital de Catalunya. (2024). Informe de posicionament estratègic 2024.

A conclusión principal deste estudo é que o cambio de algoritmo (*core update*) que Google aplicou a finais de agosto de 2023 reverteu a tendencia á baixa da visibilidade do contido en catalán, que agora mostra unha recuperación cun punto claro de inflexión na data mencionada. En concreto, obsérvase a finais de febreiro de 2024 un incremento de 15 puntos porcentuais respecto da situación inicial (ano 2021) e de 22 puntos porcentuais respecto aos peores momentos da crise de visibilidade (finais de 2022 – mediados de 2023), que chegaron a ser 18 puntos porcentuais no primeiro momento do cambio de tendencia. Na análise individual dos sitios web máis relevantes obsérvanse variacións no tempo que nalgúns casos son considerables.

A consultora Human Level, baixo a dirección de Fernando Maciá, levou a cabo un estudo no que se explora se as webs con dominio .cat obteñen un mellor posicionamento en buscadores fronte a webs similares pero con dominios .es e .com. Esta cuestión é relevante para contidos dirixidos a usuarios de linguas minorizadas con dominio propio, como é o caso do .gal.

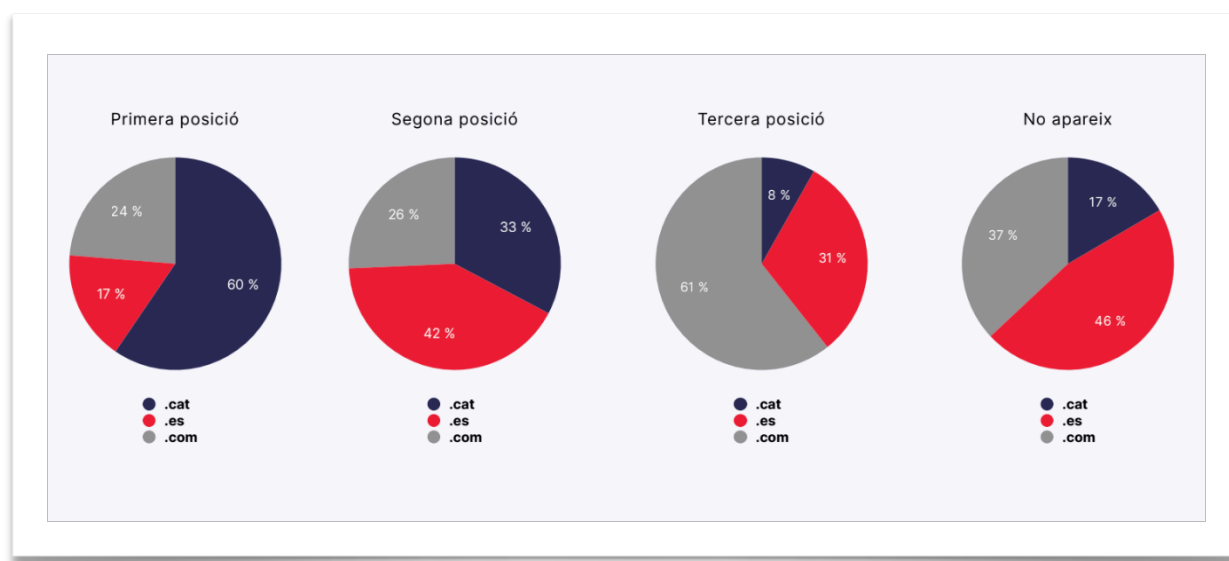
Para medir estes resultados utilizáronse webs de contido idéntico pero con TLDs distintos. O contido foi neutral para evitar sesgos de clasificación e se incluíronse diferentes escenarios de busca (idiomas, localizacións e buscadores).

As conclusións do estudo foron que baixo condicións controladas, as webs .cat acadaron posicións superiores, especialmente na primeira posición. Isto suxire que o algoritmo de Google podería favorecer TLD menos comúns como .cat para diversificar os resultados.

Igualmente, os sitios .cat foron indexados con máis frecuencia que os .com e .es, indicando que Google ten máis probabilidade de incluír dominios .cat. Esta vantaxe mantívose sen depender do idioma ou ubicación da busca, o que suxire que Google non se basea unicamente nestas variables ao clasificar as webs.

O estudo aventura posibles razóns para estas vantaxes: que Google pode priorizar TLD diversos nos resultados, que o algoritmo pode favorecer TLD menos comúns ou que Google podería favorecer `cat` para buscas catalás.

Estas conclusións teñen claras consecuencias no SEO: Para públicos catalanfalantes, o .cat ofrece vantaxe SEO. Ademais, a calidade do contido é crucial, o contido relevante é esencial para o éxito a longo prazo.



Posicionamento dos resultados segundo o dominio das webs, tirado do estudo da Fundació .cat. (2023). Estudi SEO

O estudo evidencia vantaxes significativas para o `.cat` en SEO. Non obstante, é importante avaliar os beneficios fronte a outros factores como a marca e a experiencia de usuario. Para organizacións que se dirixen a un público catalanfalante, o uso do `.cat` é recomendable.

O caso vasco

Segundo os datos analizados polo Observatorio de PUNTUEUS en 2022, o 14% das páxinas web do País Vasco teñen algún contido en éuscaro. Ademais, entre os sitios web que utilizan o dominio .EUS, o 75% publica contido en éuscaro, polo que, se se compara con dominios como .COM ou .ES, obsérvase que a presenza do éuscaro é moito maior.

No País Vasco, as responsables da Wikipedia en éuscaro tamén detectaron un cambio no algoritmo de Google. O número total de visitas presenta unha tendencia á alza, pero con un descenso nalgúns ámbitos.

Segundo o presidente de *Euskal Wikilarien Kultur Elkarte* (EWKE), Luistxo Fernández, “estamos preocupadas porque temos probas de que alimentamos ben a Wikipedia en éuscaro, imos ben no aumento de contidos e calidade, e con todo, desde o ano pasado, estamos perdendo tráfico nalgúns ámbitos concretos. Estamos convencidas de que a clave está nas buscas derivadas desde Google”.

Na mostra analizada por PUNTUEUS tamén se identificou esta discriminación do éuscaro en beneficio do castelán no servizo de busca de Google, motivo polo cal se produciu un descenso importante en termos de visibilidade do éuscaro en Internet.

Segundo Josu Waliño, director da Fundación PUNTUEUS, Google responde ás buscas de usuarias vascas con resultados en castelán. Isto ten graves consecuencias porque exclúe o éuscaro e leva os usuarios a navegar en castelán, reducindo a visibilidade do noso idioma en Internet.

En setembro de 2023 fíxose patente que houbo modificacións nos algoritmos de Google, xa que o éuscaro recuperou parte da súa visibilidade nas buscas perante o castelán. PUNTUEUS puxo a disposición de Google os resultados do seu informe e grazas a modificación do algoritmo: os resultados pasaron do 50% no mes de xullo a un 73% no mes de setembro de 2023. Os informes internos de PuntuEus, non publicados, indican que non houbo cambios dignos de mención nas vagas do estudo que lanzaron nos primeiros meses do 2024.

O posicionamento



O posicionamento é o lugar que unha marca ou produto ocupa, en relación con outros, na mente ou na atención do público. Unha empresa pode usar diferentes métodos e ferramentas para promocionar o seu produto e adxudicarlle unha serie de atributos que poden ser simbólicos, relacionados coa imaxe, a proposta de valor, a vinculación con determinados valores, etc.

Por suposto, o que lle interesa ás empresas é que as súas marcas ou produtos se interpreten coma positivos, amables, desexables e memorables na percepción do público. Idealmente, esta posición representará unha vantaxe competitiva con produtos similares no momento da escolla ou da compra.

Este posicionamento, na rede, tradúcese na presenza nun lugar destacado nas buscas. Arredor deste concepto hai toda unha industria que busca as estratexias e tácticas para ocupar estes primeiros resultados.

A importancia socioeconómica das buscas

As buscas en internet teñen unha importancia crecente na nosa vida cotiá, tanto desde o punto de vista persoal como económico. No caso do galego, revisten un impacto significativo na preservación e promoción do idioma, así como na economía local e rexional.

Cando as usuarias fan buscas en galego, están demandando contido neste idioma, o que motiva ás creadoras de contido a produci-lo. Isto non só enriquece o ecosistema de contido en galego, senón que tamén promove a súa diversidade e calidade. A existencia deste material na Internet fai que máis persoas poidan acceder a información relevante na súa lingua materna, fomentando así a súa normalización.

Esta demanda tamén impulsa ás empresas tecnolóxicas a ofrecer servizos e produtos en galego, incentivando a empresas como Google, Microsoft ou Apple a mellorar os seus produtos e servizos na nosa lingua, o que facilita o seu uso diario e reforza a presenza do galego no ámbito dixital.

As buscas en galego son un indicativo de que o idioma está vivo, non só na comunicación oral e escrita tradicional, senón tamén nas novas tecnoloxías. Isto axuda a garantir que o galego non quede relegado a un segundo plano e que siga sendo unha lingua funcional na vida moderna a medida que a sociedade avanza cara a unha maior dixitalización.

En termos económicos, as buscas en galego teñen un impacto directo na economía local e rexional. Por un lado, a demanda de contido en galego pode impulsar o desenvolvemento de empresas e proxectos tecnolóxicos locais que ofrezan servizos na nosa lingua. Por exemplo, o desenvolvemento de aplicacións, páxinas web e plataformas dixitais en galego pode xerar emprego para desenvolvedoras, deseñadoras e outras profesionais do sector tecnolóxico, e dar lugar a un mercado publicitario en galego, o que pode garantir a sostibilidade de moitos proxectos de creación de contidos xornalísticos, de tempo libre, literarios, multimedia, etc.

Ao mesmo tempo, as escolas, institutos e universidades poden beneficiarse dunha maior dispoñibilidade de recursos e materiais educativos en galego. Isto facilita o ensino e aprendizaxe do idioma, así como a súa integración en diferentes áreas do coñecemento. A dispoñibilidade de recursos educativos en galego na Internet tamén permite ao estudantado acceder a información e materiais de calidade na súa lingua materna, o que mellora a súa experiencia educativa e promove o uso do galego. A medida que se incrementa, tamén se abre a porta a novas investigacións e estudos en diferentes disciplinas. A existencia de contido académico e científico en galego é fundamental para que a nosa lingua teña unha presenza significativa no mundo da investigación e a innovación.

O SEO

O SEO ou *Search Engine Optimization* é un conxunto de estratexias de optimización dunha web para mellorar a súa posición nos resultados dunha busca.

Como se mencionou anteriormente, os motores de busca son para moitos usuarios a porta de entrada na Internet, o primeiro paso para acceder a un produto, un servizo ou un contido na rede. Tendo en conta os mecanismos de busca de cada un dos motores, estas técnicas de optimización permiten aparecer entre os primeiros postos nos resultados para determinados termos de busca.

Segundo Backlinko, unha consultora de tráfico web, só un 9% das usuarias chegan ao final da primeira páxina de resultados e un 23% das veces tendemos a seguir as suxestión de autocompletado de Google. Este mesmo estudo indica que a maior parte das persoas que buscan na rede non pasan da primeira páxina de resultados, e os tres primeiros resultados levan máis dos 50% dos clics. O primeiro resultado orgánico de Google ten dez veces máis posibilidades de ser picado que outro situado nunha páxina na posición 10.

Actualmente o buscador máis utilizado do mundo é Google, obtendo un 92% de toda as buscas mensuais, mentres que si cambiamos a dispositivos móbiles acada o 99% das mesmas. Convén recordar que o buscador pertence á mesma empresa que produce o sistema operativo de moitos dos dispositivos móbiles, e que o motor de busca ven instalado por defecto en moitas das aplicacións. Deste xeito, o SEO está condicionado polos algoritmos e o xeito de organizar as buscas de Google, que non fai pública a totalidade da información que usa para recoller e ordenar as buscas.

Sabendo isto, hai que lembrar que para Google é importante manter o equilibrio entre a relevancia das buscas, a comercialización de espazo publicitario, a fiabilidade e a rapidez á hora de amosar os resultados, de xeito que sexa útil a fidelización e familiarización coa interface de usuaria, o xeito de funcionamento, etc. As expectativas son moi altas. Como exemplo, recordemos que en máis de vinte e cinco anos de existencia do buscador, a súa páxina de inicio aínda mantén o mesmo formato: unha caixa de busca nun fondo branco e os botóns de “busca” e “vou ter sorte”.

Con todo, tendo en conta o anterior, a misión dunha boa estratexia de SEO é posicionar unha páxina web nas primeiras opcións, canto máis arriba nos resultados mellor. Para isto hai que ter en conta unha serie de factores. Entre as recomendacións das especialistas podemos atopar as seguintes claves:

- Xerar contido con intencionalidade de búsqueda baseándose en responder a preguntas das usuarias arredor do contido que, lóxicamente, ofrecemos no sitio web.
- Empregar as palabras clave que atinxen a esas búsquedas: procurar as palabras máis usadas polas persoas usuarias e empregalas para o contido na web destacándoas nos títulos, subtítulos e nas etiquetas. Estas palabras clave poden ser atemporais

(evergreen), estacionais e estar conformadas por unha (*short tail*) ou máis palabras (*long tail*).

- Coidar o aspecto da web: A páxina debe adaptarse a móbiles e tabletas (isto é, ser *responsive*), pois os buscadores penalizan ás webs que non o fan.
- Crear un deseño lixeiro: Páxinas que carguen rápido, con contidos audiovisuais en tamaños reducidos e que non teñan contidos duplicados para facilitar o rastrexo por *spiders*.
- Títulos sinxelos, usar puntos para enumerar, letra grosa e ligazóns internas e externas que os buscadores valoran pola súa autoridade (é dicir, páxinas con tráfico e un dominio seguro con anos de percorrido).
- Publicidade: Evitar publicidade intrusiva na propia páxina.
- Ligazóns externas: Consegui que outras plataformas liguen o sitio para refrendar a súa relevancia como fonte citada.
- Participación: Escribir en webs e foros indicando a nosa web como referencia, se procede.
- Redes sociais: Ter unha estratexia para crear unha comunidade fiel. Se ben non está demostrado que afecte directamente ao SEO, si supón unha fonte de entrada de tráfico que incide na relevancia da web.
- Aumentar o tráfico cara a web utilizando estratexias de márketing por correo electrónico (newsletters) ou mesmo de xeito *offline* con campañas publicitarias ou outras accións (márketing na rúa, de *guerrilla*...).
- Seguir as especificacións ou recomendacións técnicas de Google:
 - Diseñar un sitio ben estruturado, onde non sexa preciso acceder a máis de dous submenús para obter a información.
 - Adaptabilidade a calquera dispositivo (*desktop* ou móbil).
 - Ter un mapa do sitio actualizado.
 - Empregar metaetiquetas e *titles* para explicar o contido da páxina.
 - URL simples e amigables, feitas de xeito semántico, como www.punto.gal/seo-en-galego fronte a www.punto.gal/12-739w.
 - Facer un seguimento da indexación do contido coa ferramenta Google Search Console.

O SEM

Como xa se falou no punto anterior, as páxinas de resultados dos buscadores na web, nomeadamente as de Google, ofertan espazo de pago no que se poden anunciar produtos, servizos ou contidos. SEM son as siglas en inglés de *Search Engine Marketing*, márketing en buscadores e é unha técnica para optimizar a compra destes espazos comerciais. Non se trata de aparecer organicamente nos resultados de busca, senón que, dependendo da busca, Google amosará un anuncio determinado (que poderá aparecer ou non segundo a relevancia que estime o algoritmo, non polo que se paga por el) xunto cos resultados orgánicos.

A diferenza do SEO (*Search Engine Optimization*), que se centra en mellorar o posicionamento orgánico dun sitio web, o SEM implica o uso de anuncios pagados para aparecer nas primeiras posicións dos resultados de busca. Estes anuncios, que normalmente se ven na parte superior ou lateral das páxinas de resultados de motores de busca como Google, son etiquetados como "anuncio" ou "patrocinado" para distinguirse dos resultados orgánicos.

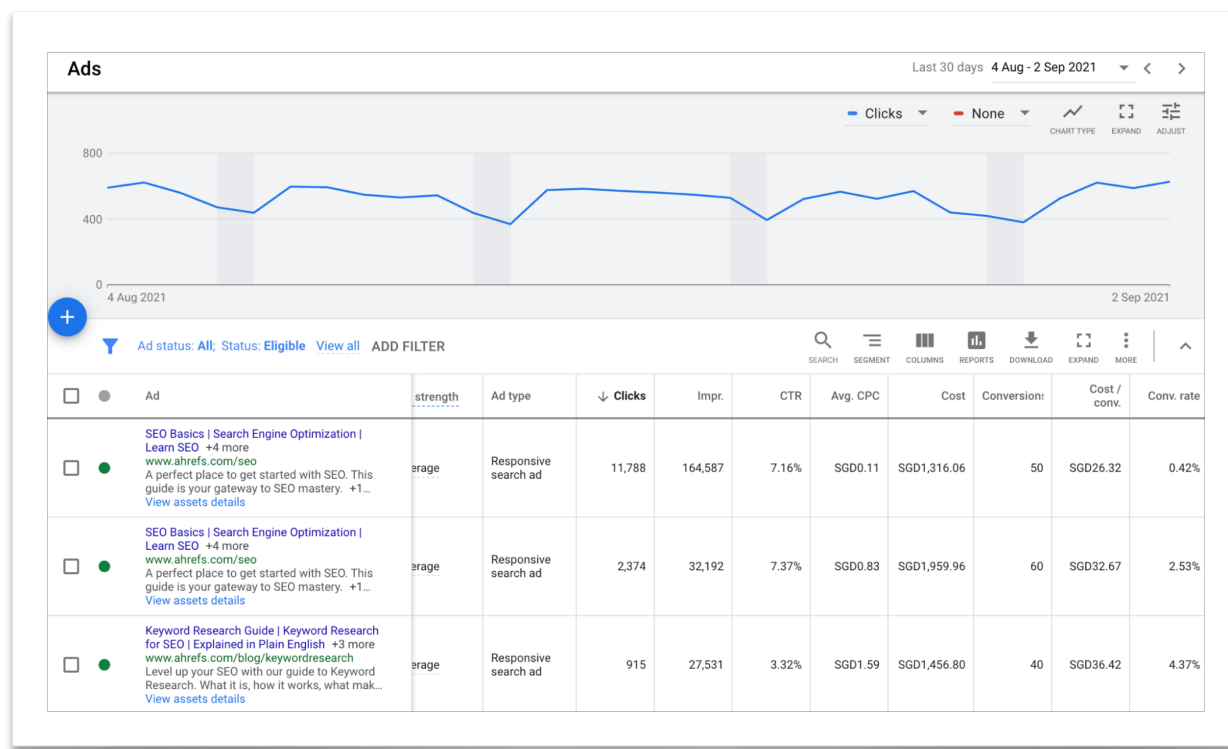
Un dos métodos máis comúns de SEM é a publicidade a través de Google Ads. Neste modelo, as anunciantes ofertan poxas por palabras clave relevantes para o seu negocio, idealmente, as que alguén introduciría no motor de búsqueda cando quere atopar ese produto ou servizo. Cando unha usuaria busca esas palabras, o anuncio pode aparecer nos resultados. Se a persoa fai clic no anuncio, a anunciante paga unha cantidade determinada ao motor de busca, nun modelo coñecido como pago por clic (PPC).

Os anuncios compiten por aparecer na parte superior dos resultados de busca. A posición do anuncio non depende só do prezo que o anunciante está disposto a pagar por clic, senón tamén doutros factores, como a relevancia e calidade do anuncio, medida a través dun índice de calidade chamado *Quality Score*.

O *Quality Score* avalía varios aspectos, como a relevancia das palabras clave elixidas, a calidade da páxina de destino onde se chegará ao facer clic e a taxa de clics esperada, é dicir, a probabilidade de que as usuarias fagan clic no anuncio.

Cando o anuncio aparece nos resultados de busca e un usuario fai clic nel, págase unha cantidade que varía en función da competencia pola palabra clave. Este método permite controlar os custos publicitarios, xa que o pagamento só se realiza cando hai interacción co anuncio.

O sistema PPC de Google Ads é moi eficaz, xa que permite ás empresas chegar ao seu público obxectivo cando están buscando activamente produtos ou servizos similares aos que ofrecen, maximizando así a probabilidade de conversións.



Un exemplo de pantalla de seguimento de campaña de Google Ads, no que se ve a evolución dos clics conseguidos.

Teoricamente, o SEM ofrece tráfico e potenciais clientes máis a curto prazo que o SEO, que require máis traballo e tempo de optimización, xa que implica que os motores de búsqueda posicione os contidos de xeito orgánico. O SEM posibilita medir o rendemento dos anuncios en tempo real, permitindo optimizar as campañas para obter mellores resultados. Métricas como o custo por clic (CPC), a taxa de clics (CTR) e o retorno sobre o investimento (ROI) son fundamentais para avaliar o éxito dunha campaña de SEM.

Á hora de executar a estratexia, en Rivas Moreno, S. (2021, 1 de xuño). destácase que, primeiro, hai que determinar unha lista de expresións clave para posicionar, revisando e contrastando coa competencia. Despois, decídese en que URL integrar cada termo, e finalmente inclúense mediante accións concretas. É esencial seleccionar ben as expresións, utilizando ferramentas como Semrush, Ahrefs, ou Keyword Revealer, que ofrecen datos sobre buscas mensuais e permiten coñecer a estratexia da competencia. Así, é posible afinar a busca e ver as expresións que usan empresas do mesmo sector.

Recomendacións para webmasters de Google

Axeitar o contido para as búsquedas é importante, xa que é o que permite que o vexan máis usuarias relevantes. Este proceso denomínase optimización en buscadores e pode facer que máis persoas visiten un sitio no que teñen interese segundo os criterios introducidos no buscador. Se Google ten problemas para interpretar unha páxina, é posible que se estea a desaproveitar unha fonte importante de tráfico.

Google mantén unha serie moi completa de instrucións, recomendacións e cursos para optimizar e configurar os contidos das webs, ás que se pode acceder desde as súas páxinas de contratación de publicidade ou de busca. A continuación faise unha escolma das máis interesantes, sempre adaptadas da súa web, que se actualiza de xeito constante.

Descubrir como ve Google a web

Pódese examinar o sitio coa ferramenta de inspección de URLs ou coa proba de resultados enriquecidos para comprobar como o ve Google. O robot de Google é o rastrexador web que descubre páxinas novas e actualizadas para o índice de Google, como se explicou falando do funcionamento dos buscadores en xeral.

Te damos la bienvenida a Google Search Console

Para empezar, selecciona el tipo de propiedad

Dominio nuevo

- Todas las URL de todos sus subdominios (m., www., etc.)
- Todas las URL, tanto de https como de http
- Que requieran la verificación de DNS

example.com

Introduce un dominio o subdominio

CONTINUAR

Prefijo de la URL

- Solo las URL incluidas en la dirección introducida
- Solo las URL incluidas en el protocolo indicado
- Que permitan varios métodos de verificación

https://www.example.com

Introduce una URL

CONTINUAR

Vista da consola da ferramenta de inspección de webs de Google.

Comprobar as ligazóns

O robot de Google obtén e analiza enlaces, sitemaps e redireccións para ir dunha URL a outra, e trata cada unha coma se fora a única URL que veu do sitio. Para que o robot de Google poida encontrar todas as URLs do sitio, cómpre seguir estes pasos:

- Utilizar os elementos <a> que Google pode rastrexar. Comprobar que se poida acceder a todas as páxinas do sitio mediante un enlace dende outra páxina que tamén se pode encontrar. Asegurarse de que a ligazón de referencia inclúa texto ou, no caso das imaxes, un atributo “alt” que sexa pertinente para a páxina de destino.
- Crear e enviar un *sitemap* do sitio para que o robot de Google poda rastrexalo de xeito máis eficaz. Os *sitemaps* son arquivos nos que se proporciona información sobre as páxinas, os vídeos e outros arquivos do sitio, así como as relacións entre eles.
- No caso das aplicacións de JavaScript que teñan só unha páxina HTML, comprobar que cada pantalla ou que cada parte do contido teña unha URL.

Comprobar como se usa JavaScript

Aínda que Google poida executar JavaScript, cómpre ter en conta certas diferencias e limitacións ao deseñar páxinas e aplicacións para adaptalas á forma na que os rastrexadores acceden ao teu contido e o procesan.

Informar a Google cada vez que o contido cambie

Para asegurar que Google encontre rapidamente páxinas novas ou actualizadas, convén facer o seguinte:

- Enviar sitemaps cada vez que haxa cambios. Os sitemaps son arquivos nos que se proporciona información sobre as páxinas, os vídeos e outros arquivos do sitio, así como as relacións entre eles.
- Solicitar que Google volva rastrexar as URLs.

Se aínda así hai problemas para indexar unha páxina, compre revisar os rexistros do servidor para ver se hai erros.

Non esquecer poñer texto nas páxinas

O robot de Google só pode encontrar contido visible textualmente. Non pode ver o texto que aparece en vídeos. Para asegurarse de que a búsqueda de Google entende o contido da páxina, estes son os pasos:

- **Expresar o contido visual tamén en texto.** Por exemplo, non é ideal publicar unha páxina de categoría de produtos na que se mostran imaxes de camisas sen ningún texto. É recomendable redactar unha explicación textual para cada unha delas.
- **Comprobar que cada páxina teña un título descritivo e unha metadescripción.** Cos títulos únicos e as metadescripcións, Google indica a relevancia das páxinas para as usuarias, o que pode facer que o tráfico de búsqueda aumente.
- **Usar código HTML semántico.** Google pode indexa páxinas HTML, contido en PDF, imaxes e vídeos, pero non indexa contido que necesite complementos, como Java o Silverlight, nin contido que se renderice en canvas. Recomendamos non utilizar complementos no contido e que se usen etiquetas HTML semánticas sempre que se poida.
- **Asegurarse de que o contido de texto sexa accesible en el DOM.** Por exemplo, o contido que se engade a través da propiedade *content* de CSS non forma parte do DOM e a búsqueda de Google ignóralo polo momento. O DOM é o *Document Object Model*, que conecta páxinas web con scripts e está organizado coma unha árbore.
- **Unificar URLs duplicadas.** Unificar páxinas duplicadas, de xeito que o rastrexador de Google teña constancia de cal é a principal ou “canónica”.
- **Notificar versións localizadas da páxina.** Avisar a Google de que hai varias versións do contido en distintas linguas, etiquetadas correctamente.
- **Facer que unha páxina sexa recoñecible en AMP, Accelerated Mobile Pages,** que facilita e acelera a carga das páxinas para móbil.

Controlar o contido que ve Google

Pódese bloquear o robot de Google de varias formas:

- Para impedir que Google encontre unha páxina, pódese restrinxir o acceso ao contido de modo que só os usuarios rexistrados poidan velo.
- Para impedir que o robot de Google rastrexe unha páxina, pódese crear un arquivo robots.txt.

Hai que ter en conta que este mecanismo non serve para manter unha páxina web fóra de Google. Se o que se quere é que unha páxina web non apareza en Google, hai que usar regras robots *noindex* ou protexer a páxina cun contrasinal.

Para impedir que Google indexe unha páxina, pero permitir o rastrexo, hai unha serie de etiquetas que lle notifican aos rastrexadores o que deben facer con cada páxina.

Se se utilizan varias regras de rastrexo e indexación á vez, é posible que se produza algún conflito entre elas. Google ofrece instrucións para combinalas de xeito adecuado.

Habilitar os resultados enriquecidos do sitio

Os resultados enriquecidos poden incluír estilos, imaxes ou outras funcións interactivas para facer que un sitio web destaque máis nos resultados da busca. Co uso de datos estruturados na páxina, pódese proporcionar a Google información específica sobre o contido. Deste xeito, Google poderá interpretar mellor a páxina e mostrar resultados enriquecidos na busca.

Filtrar resultados para un idioma en particular

O filtro é eficaz para restrinxir os resultados web a un ou varios idiomas elixidos. Non obstante, presenta limitacións. Pode que non funcione correctamente se non se pode determinar con certeza o idioma, como ocorre en páxinas que contén texto en máis dunha lingua, ou en determinadas funcións de busca. Na páxina relativa ao filtro de resultados por idioma detállanse os pasos necesarios para habilitar e utilizar esta función.

Tamén se pode filtrar os resultados por idioma antes de realizar unha busca mediante a páxina de busca avanzada. Ambas opcións ofrecen a flexibilidade de filtrar unha única vez se é necesario.

Ter en conta as persoas que realizan buscas en varios idiomas

Para as editoras en áreas onde se usan varios idiomas, é recomendable crear contido en todos eses idiomas e non soamente no que se entenda amplamente.

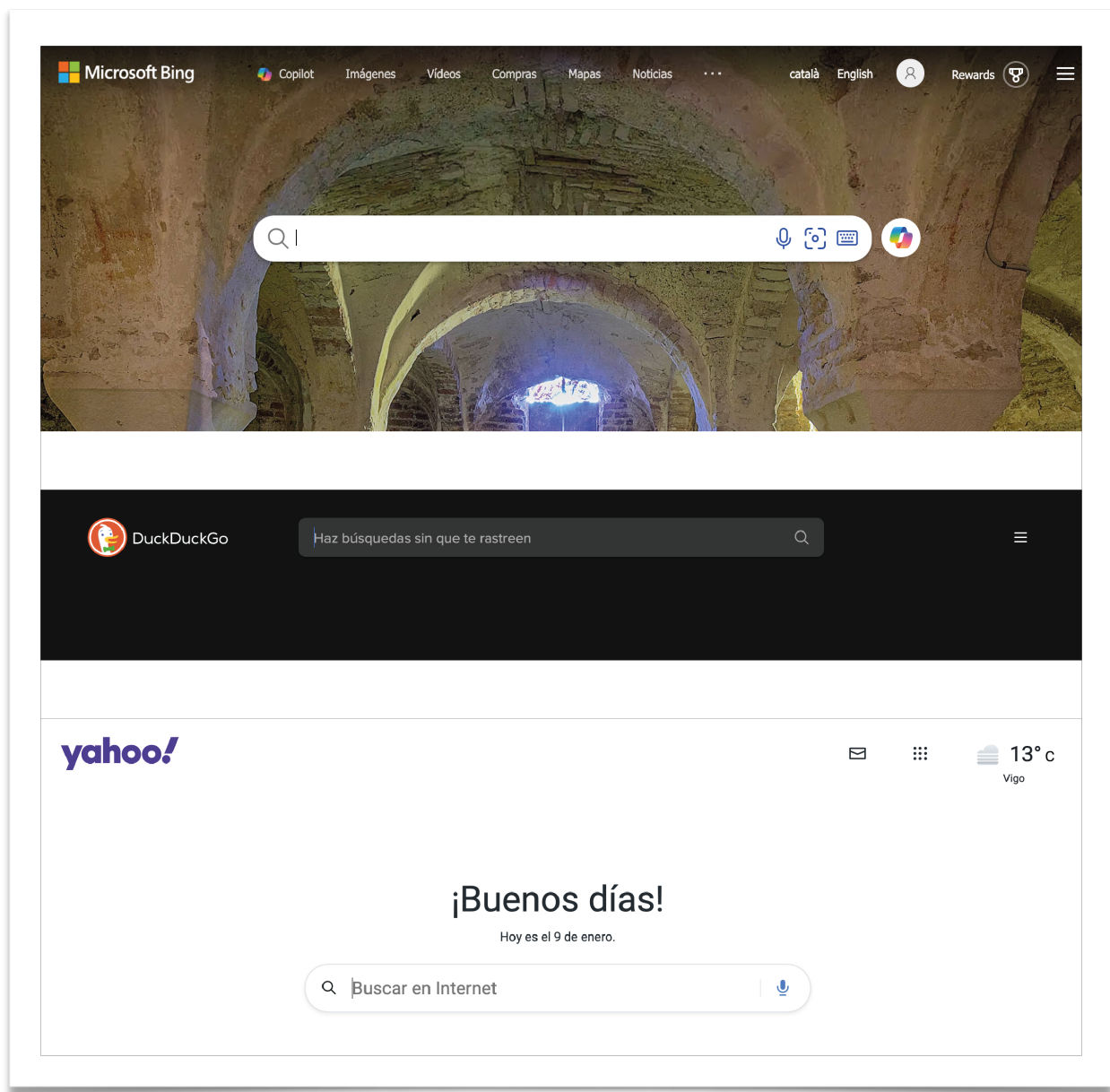
Quen produce contido en varios idiomas tamén debe asegurarse de seguir a guía técnica, tal como se describiu cando se falou de informar a Google das versións en distintas linguas dun mesmo contido.

Manterse informada das novidades nas buscas

Google di que entende que idioma usa a xente a meirande parte do tempo, pero as preferencias de idioma son complexas en todo o mundo. Por iso todos os motores seguen a experimentar con funcións como a busca bilingüe.

Situación nos outros buscadores

Analízase Bing, Yahoo! Search e DuckDuckGo como motores onde poder traballar o posicionamento do galego. O resto de buscadores, de nicho ou máis limitados, non se teñen en conta actualmente neste traballo.



Bing (Microsoft)

Aínda que Bing, o motor de busca desenvolvido por Microsoft, ten unha presenza menos dominante que Google nalgúns mercados, incluíndo o galego, a súa estratexia de mostraxe de resultados en idiomas minoritarios ten evolucionado. Ocupa un lugar relevante no universo dos buscadores en liña, a pesar de non alcanzar a popularidade de Google. Lanzado en 2009 como sucesor de Live Search, Bing buscou ofrecer unha alternativa competitiva con características distintivas, como a integración de tecnoloxía de recoñecemento visual, melloras na busca de imaxes e vídeos, e un deseño atractivo e intuitivo.

Un dos seus puntos fortes é a estreita integración cos produtos de Microsoft, especialmente co navegador Edge e o sistema operativo Windows. Esta interconexión fai que moitas usuarias de Windows usen Bing de forma predeterminada, dándolle unha cota de mercado significativa, especialmente en dispositivos de escritorio. Ademais, Bing tamén fornece resultados para outras plataformas como Yahoo! Search e DuckDuckGo, o que lle permite estar presente de forma indirecta noutros espazos da busca en liña.

Un aspecto clave na súa relevancia é a innovación en intelixencia artificial e a análise de grandes volumes de datos. Bing utiliza algoritmos avanzados para ofrecer resultados personalizados, especialmente en busca de compras, viaxes e noticias. Recentemente, Bing introduciu o seu chatbot potenciado por IA, Bing Chat, que permite interaccións máis naturais co motor de busca, unha tendencia en crecemento no sector.

No entanto, a dispoñibilidade de resultados específicos en galego pode variar dependendo do tema da busca e da popularidade do contido en Internet. Bing ten unha capacidade limitada en comparación con idiomas de maior uso como o inglés ou o castelán, o que pode resultar nunha menor profundidade de resultados ou nunha menor frecuencia de actualización en comparación con outros idiomas.

En termos de interface e usabilidade, Bing adoita presentar opcións para traducir páxinas web ao galego, aínda que o nivel de precisión pode variar.

Yahoo! Search

Yahoo! Search, que foi un dos primeiros motores de busca populares en Internet, ten evolucionado co tempo e, a día de hoxe, depende doutros motores de busca para fornecer os seus resultados. Actualmente, Bing é o provedor principal dos resultados de Yahoo! Search. Isto significa que o comportamento de Yahoo! Search en relación cos resultados das buscas en galego está fortemente influenciado polo rendemento e as capacidades de Bing.

DuckDuckGo

DuckDuckGo, un motor de busca que aposta pola privacidade e a protección de datos, ten unha estratexia distinta á de outros motores de busca, como Google ou Bing, en termos de como xestiona as buscas en idiomas minoritarios, como o galego. Ao igual que Yahoo! Search, depende doutros motores de busca e fontes de datos para ofrecer resultados. Usa en gran medida APIs e fontes de Bing, Yahoo!, Wikipedia e outros sitios de contido libre. Isto significa que a calidade dos resultados en galego está suxeita ás mesmas limitacións que se observan en Bing.

En xeral, para buscas en galego sobre temas amplamente cubertos, como noticias locais, cultura, ou temas históricos, DuckDuckGo é capaz de ofrecer resultados relevantes en galego. Por exemplo, páxinas da Galipedia (a versión en galego de Wikipedia), artigos de medios rexionais como Praza Pública, ou fontes educativas e culturais en galego son comúnmente visibles entre os primeiros resultados.

O posicionamento do galego



O posicionamento do galego con respecto aos buscadores en xeral e a Google en particular ten moito en común con outras linguas cooficiais da península. Tal e como se comentou no punto anterior, Google non publica información na que se singularice o tratamento dunha lingua e non se fale das outras. É un feito que, por número de falantes e por soporte técnico, o catalán está en mellor situación que o éuscaro ou o galego: hai versión catalá dos sistemas operativos máis importantes, tanto para dispositivos móbiles como para sistemas de sobremesa, e dos programas ofimáticos e empresariais máis comúns, así como navegadores e plataformas. Ademais, os sistemas de publicidade de Google (usados para SEM ou para inserción de publicidade sindicada) recoñecen o catalán como lingua de contidos e para a emisión de anuncios, posibilitando a monetización de contidos e servizos en catalán. Non é así no caso do galego, como se menciona máis adiante.

Esta situación se menciona con certa frecuencia nos medios, con referencia tanto aos resultados das buscas, comentando que as páxinas en galego quedan afogadas entre as que están noutras linguas, como nos casos de vacilacións ou deturpacións toponímicas ou as decisións que se toman por parte das plataformas e as redes sociais respecto á lingua.

Isto fai que moitos dos contidos de administracións públicas, como por exemplo, as páxinas da propia Xunta de Galicia, nas buscas de Google, saían por defecto en castelán, aínda que o contido e idiomas principais sexa o galego e as procuras se fagan desde dispositivos configurados nesta lingua.

Situación actual

Hoxe, cando se busca información en galego nos principais motores de busca, os resultados adoitan estar mesturados con contidos en castelán ou outras linguas. Isto ocorre porque os algoritmos de busca non sempre diferencian ben entre o galego e o castelán, debido á súa similitude lingüística e á menor cantidade de contido dispoñible en galego. Ademais, moitos dos contidos xerados polas usuarias en Galicia están en castelán ou presentan versións bilingües, o que provoca que as páxinas en galego teñan menos relevancia segundo os algoritmos de posicionamento SEO.

Así, un dos principais retos para o galego nos buscadores é a presenza de contidos en galego na rede, menor que a doutras linguas. Como resultado, cando unha persoa realiza unha busca en galego, os resultados relevantes son máis limitados en comparación con outras linguas.

Outro desafío é a detección e identificación automática do galego polos buscadores. En ocasións, os algoritmos non identifican correctamente o idioma dun texto, o que pode levar a unha mestura inadecuada de resultados en galego, castelán ou portugués. Esta falta de precisión afecta negativamente á experiencia da usuaria que busca información especificamente en galego, que ten que seleccionar manualmente entre os resultados que se lle ofrecen.

Por outra banda, a día de hoxe tampouco podemos atopar opcións en galego nas plataformas de publicidade SEM de Google Ads, polo que é moi difícil posicionarse no mercado, dado que se unha empresa ou entidade galega non ten a oportunidade de pagar por poñer a súa publicidade en galego, non vai ter a capacidade de obter os mellores resultados nas búsquedas.

Esta situación ocorre con outras ferramentas de contido en rede como Youtube ou redes sociais como Facebook, que poderían axudar a crear marca no mundo dixital e que nin permiten publicidade en galego nin recomendan publicacións afíns na nosa lingua, o que fai perder pertinencia no posicionamento.

A día de hoxe os datos dos que dispoñemos son escasos e subxectivos. Os buscadores non publican os seus criterios para lle conceder relevancia a uns resultados sobre outros, nin comparten os seus posibles plans para incorporar o galego como lingua de pleno dereito nos seus produtos. Precisamente, desta situación parte a necesidade e a pertinencia dun estudo como o que se propón neste informe, tanto para establecer a situación da que partimos como para monitorizar os cambios que poidan vir nun futuro

Recomendacións das expertas reunidas por PuntoGal

No mes de xullo de 2024, PuntoGal convocou un comité de seis expertas do ámbito público e privado con experiencia en diferentes eidos relacionados co SEO (o propio posicionamento, deseño web, comunicación dixital e ámbito académico) para poñer en común reflexións arredor de como mellorar a presenza do galego en internet:

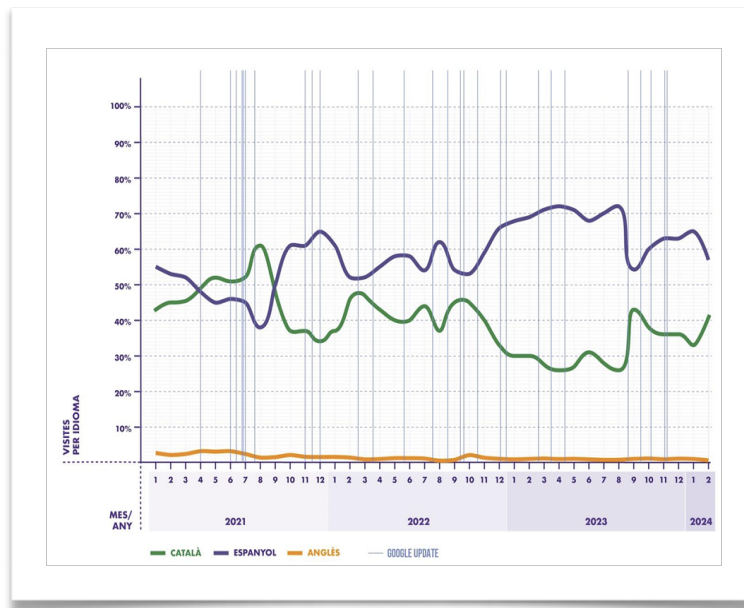
- **Sergio Ramírez**, quen leva traballando en posicionamento web desde 2008, e actualmente é director de márketing en Raiola;
- **Alejandro Arce**, Digital Product Manager en Visual Publinet desde hai 26 anos (onde xestionan máis de 1.000 webs, das que o 20% son ecommerce);
- **Saúl Vérez**, especialista en SEO e programador (ten a súa marca persoal, www.saulverez.com);
- **Bárbara G. Vilariño**, xornalista especializada en comunicación e márketing dixital, vencellada a proxectos nos eidos culturais, deportivos e asociativos a través do colectivo www.davella.gal;
- **Anxo Sánchez**, deseñador web con foco en clientes multiidioma ou só en galego, e promotor da plataforma de emprendemento en galego www.medra.gal;
- **Xabier Rolán**, profesor universitario que estuda como se relacionan as comunidades dixitais.



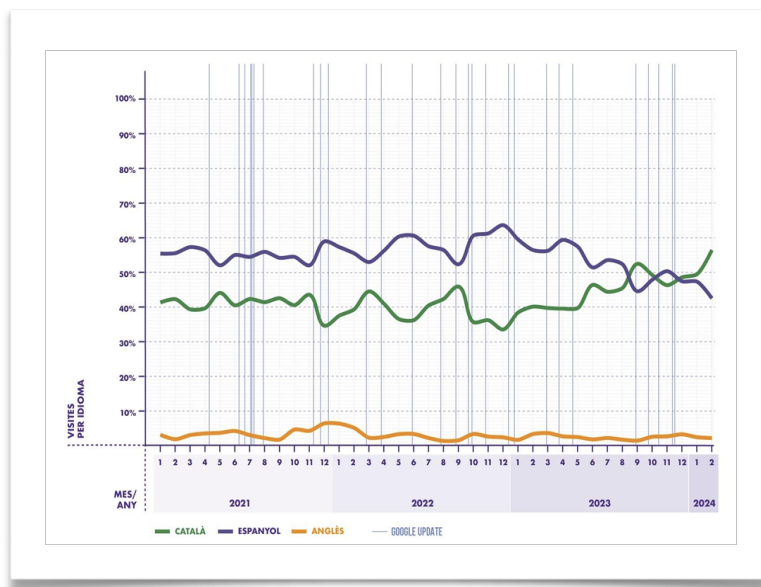
De esquerda a dereita: Sergio Ramírez, Alejandro Arce, Saúl Vérez, Bárbara G. Vilariño, Anxo Sánchez E Xabier Rolán na reunión de expertas en SEO

Previamente a esta reunión, compartíronse achegas da *Aliança per la presència digital del català* a través de dous documentos previos previamente explicados neste informe..

Neste sentido, estes son os dous gráficos máis significativos que se referenciaron de xeito previo á reunión, extraídos do informe da *Aliança per la presència digital del català*. (2024) e nos que pode apreciar as alteracións no idioma dos resultados das buscas na partir dos cambios no algoritmo do buscador:



Este primeiro fala dos resultados de visitas en todas as webs que se tiveron en conta no estudo catalán, con 400 millóns de visitas. Aliança per la presència digital del català. (2024).



E o segundo, no que se exclúen os medios de comunicación e reflexa os resultados de 100 millóns de sesións. Aliança per la presència digital del català. (2024).

Este primeiro encontro forma parte das actividades do proxecto de análise da situación do galego na rede e busca tamén posibles solucións que fagan máis sostible a presenza da nosa lingua en Internet, tanto desde a perspectiva cultural como tecnolóxica e económica.

Aínda tendo en conta que as plataformas publicitarias máis importantes da rede, como a de Google, non consideran o galego como lingua independente, ~~de feito~~ hai unha comunidade que produce contidos de todo tipo en galego, que tamén vende produtos e servizos en galego e que aspira a que as usuarias se acheguen. Por iso esta primeira xuntanza de expertas en SEO e márketing en galego na rede é un primeiro paso para poder formular de xeito explícito os problemas que se atopan desde o punto de vista técnico e comercial, e as posibles solucións que se poden aportar, tanto desde o tecido empresarial como desde as institucións e o mundo académico.

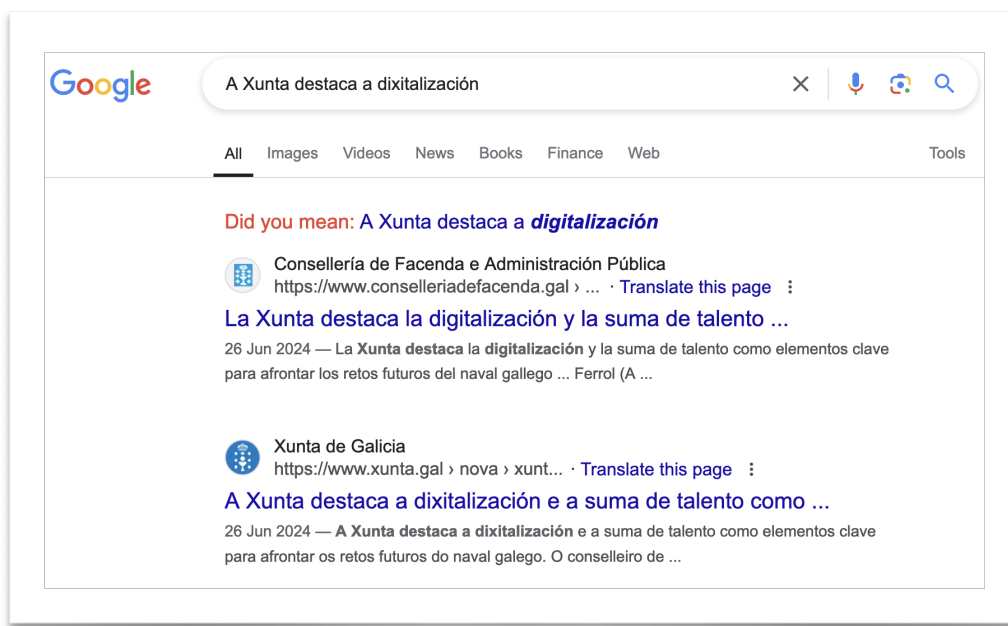
Ao respecto desa comunidade, debateuse na xuntanza que ese grupo de creadoras -e mesmo persoas provedoras de bens e servizos- tiña presenza especialmente nas redes sociais, onde non existe unha conexión coas procuras orgánicas, polo que, sería convinte incentivar dalgún xeito a translación desa conversa e xeración de contidos cara espazos dixitais como blogs ou webs. Deste xeito, aportaríamos volume de material que xustificaría a relevancia do galego en Internet, xa que o que queda en redes sociais case resulta inexistente para Google. De feito, o grupo coincidiu en sinalar isto como unha fonte de riqueza de contidos envexada por outras comunidades, como a vasca, que se podería levar ao resto da Internet.

Esta cuestión entronca coa sostibilidade económica do galego na rede, xa que non é posible facer anuncios no idioma. Polo tanto, se non hai rendibilidade económica, non haberá contido (blogs, foros, webs...) en galego, e doutra banda non haberá especialistas SEO en galego que axuden a posicionar todos estes formatos.

Nas posibles solucións marcáronse como exemplos de socialización as achegas da Mesa pola galeguización de dispositivos, as guías de R de como configurar o teléfono en galego e as guías de PuntoGal de como configurar navegador e redes sociais en galego. Tamén se coincidiu en que estes esforzos individuais, malia que imprescindibles, requiren dun apoio institucional, nomeadamente de Política Lingüística e mesmo vencellalo a consellerías como Economía e Emprego, para darlle máis pulo e velo como unha ferramenta de activación económica no eido dixital. Destacouse que se fan campañas do uso en galego, en abstracto, pero que talvez sería máis útil dirixilas cara “busca en galego”. Propúxose facer boas prácticas con medios que castellanizan novas en galego e a oportunidade de crear alianzas con algo como un selo distintivo de PuntoGal, que supoña unha motivación extra para involucrarse.

Situación actual

Nestes momentos, o algoritmo de Google xa é case como unha rede neuronal, na que mesmo as súas propias enxeñeiras non son quen de identificar as capas que o conforman. Por citar un caso paradigmático, cando se busca unha nota de prensa da Xunta os resultados sempre priorizan o castelán. Noutras procuras do día a día, como poden ser servizos de urxencia (cerralleiro, copias de chaves, reparación...) ou mesmo cuestións técnicas, empréganse o castelán ou o inglés, respectivamente.



Un exemplo de resultados bilingües nos resultados de Google.

A primeira vista, poderíase dicir que hai máis galego en redes sociais que en Google. Segundo os informes do Instituto Galego de Estatística (IGE), o uso do galego en redes non é escaso, como se pode ver no cadro adxunto, tirado do informe de 2019:

Persoas segundo a lingua que utilizan nos medios de comunicación. Galicia. Ano 2018

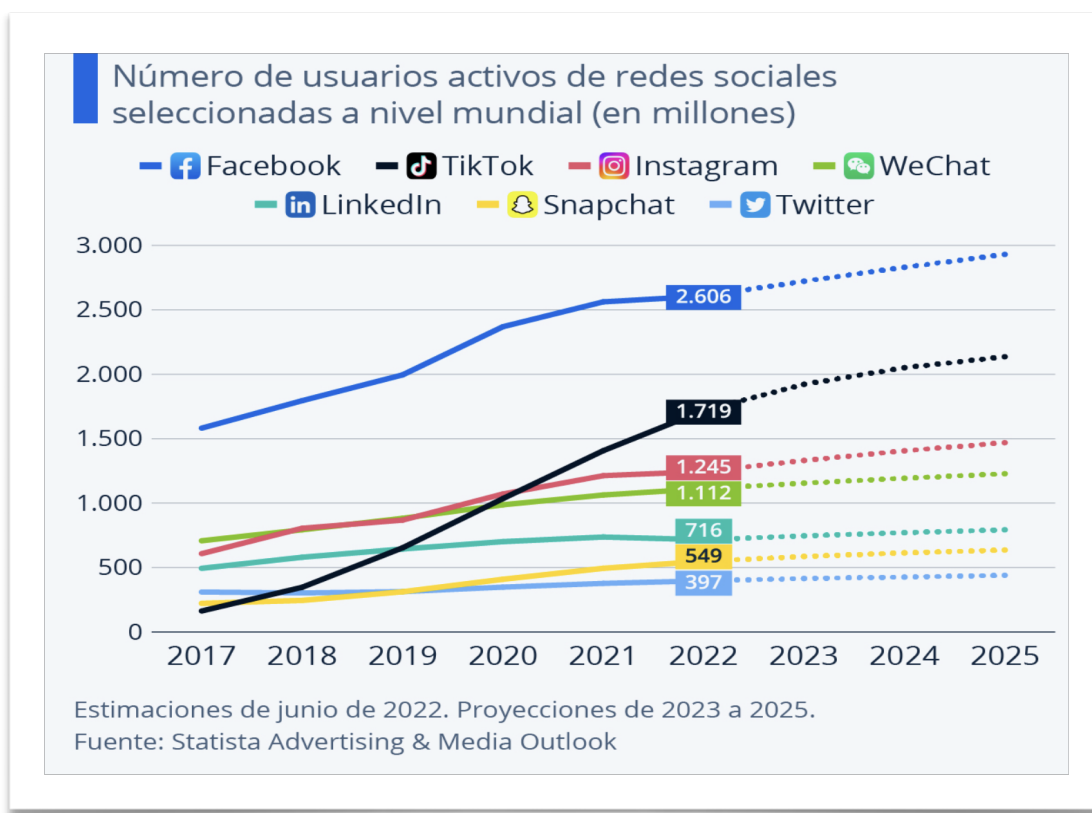
Unidade: porcentaxe (%)

Medios	En galego sempre	Máis galego ca castelán	Máis castelán ca galego	En castelán sempre
Ver a televisión	2,92	14,73	55,75	26,60
Escoitar a radio	3,48	10,78	46,72	39,02
Ler a prensa	1,06	2,90	36,94	59,10
Ler libros	1,19	3,95	40,83	54,03
Navegar por internet	1,29	3,78	25,46	69,47
Redes sociais	8,25	18,03	21,65	52,07

Fonte: IGE. Enquisa estrutural a fogares. Coñecemento e uso do galego ([descarga da táboa](#))

Uso lingüístico no consumo de medios, segundo o informe do IGE (2019).

Neste senso, hai que considerar que TikTok é a canle que emprega a xente máis nova para buscar, informarse e xerar contido, en lugar de Google, polo que sería de grande relevancia contar con datos de uso de idioma das fontes onde se vaian requirir datos para o estudo.



Uso de redes sociais referidos a 2022. Datos tirados de statista.com.

Respecto á escolma de datos, seguramente se requira anonimizar as mostras e, con alta probabilidade, non se poidan manexar cifras absolutas, engadíndose a isto a desaparición dos datos de Universal Analytics de Google. Hai que considerar tamén que os resultados acabarán véndose afectados polas buscas de intelixencia artificial ou o cambio de uso de Google con respecto a Edge, e ese emprego como buscador das propias redes sociais.

Esta problemática debería contrastarse cos datos do ecosistema de webs en galego, número de persoas usuarias que empregan galego en internet, e amosar ese mercado existente para poder ofrecelo a posibles clientes que non estaban a valorar ter a súa páxina en galego. Unha das persoas participantes indicou que perdeu 95% do tráfico SEO ao eliminar o castelán da súa web.

O comité de expertas coincidiu en que manter o galego funcionando en Internet debe deixar de ser voluntariado, e que cómpre xerar foros de debate que unan profesionais e entidades públicas e privadas interesadas, así como trazar alianzas coas comunidades vasca e catalá para adquirir en conxunto ferramentas útiles para a análise e implementación.

Recomendacións das expertas reunidas por PuntoGal

O comité referendou dez propostas arredor de dous eixos: a difusión e socialización do problema e a estandarización dunha metodoloxía de análise e aplicación de solucións.

Recomendacións sobre a difusión e socialización do problema:

1. Cómpre facer pedagogía e indicar os beneficios do emprego do galego en Internet activando distintas fórmulas que atinxan a público xeral e profesionais. Por unha banda, a promoción da lingua a nivel de usuaria, promovendo guías sinxelas do día a día (como configurar en galego ferramentas básicas e cotiás como o navegador, ou enderezo electrónico, redes sociais...) ou campañas de sensibilización e divulgación destas opcións de personalizar a lingua en diversa índole de aplicacións dixitais. Destácanse efemérides como o Día de Internet ou mesmo as Letras Galegas, fitos con sona nos que nunca se traballou en profundidade na presenza do galego no eido dixital. Nesta cuestión poderían contribuír distintos axentes, como os creadores de contido nas redes sociais. No tocante a profesionais que venden os seus servizos desde Galicia e en galego, a problemática é clara e xera unha cadea difícil de quebrar se non se aplican medidas transversais desde un punto de vista global. Se non hai rendibilidade económica, non se xerará contido (blogs, foros, webs...) en galego, e como consecuencia non haberá un ecosistema laboral arredor do galego en internet (especialistas SEO en galego, deseñadores web, redactores...). Resultaría relevante demostrar que se pode ter un negocio en Internet baseado no galego, e mesmo que se pode converter nun nicho no que atopar oportunidades comerciais que non se dan se só se cultiva unha versión en castelán, xa que o público é moito máis específico.
2. Elaborar guías de boas prácticas do uso de galego en Internet, procurando unha motivación extra cun selo distintivo de PuntoGal. Deste xeito se estarían premiando casos exemplificantes que sirvan ao mesmo tempo de multiplicadores do efecto das campañas de sensibilización, e se axudaría a que outras iniciativas poidan despegar ou perfeccionarse grazas á simplificación de ferramentas (segundo comentado na primeira recomendación) e aportación de pautas que evite tanto erros no uso como priorización do castelán por defecto nas versións bilingües. Ao mesmo tempo, un selo distintivo garantiría unha calidade e confianza que reforzaría as distintas iniciativas empresariais.
3. As institucións públicas, unidas de xeito transversal, poden axudar a ver isto como unha ferramenta de activación económica no eido dixital. Tanto a nivel interdepartamental e de distintos ámbitos locais como desde distintas áreas de aplicación, xa que a promoción do galego en Internet atinxe non tan só á política lingüística, tamén a todo o tecido socioeconómico de Galicia. A colaboración transversal entre as institucións públicas permite unha abordaxe integral que supera as fronteiras tradicionais da política lingüística. Cando diferentes departamentos e áreas locais traballan en conxunto, pódese xerar un impacto máis significativo. As políticas interdepartamentais poden fomentar a

creación de contidos en galego, apoiar o comercio electrónico local e impulsar proxectos tecnolóxicos que empreguen a lingua galega. Ademais, a coordinación entre áreas como a educación, a cultura, o turismo e o emprendemento pode crear un ecosistema dixital e inclusivo, onde o galego sexa unha lingua viva e útil para a cidadanía.

4. Incentivar as buscas en galego, apoiándose no rico ecosistema de influentes e creadores nas redes sociais para amosar que, se buscas en galego, vas atopar o que procuras, e que non tes que mudar ao inglés ou o castelán. Estas profesionais teñen unha capacidade única para conectar coas usuarias, especialmente as máis novas, e poden demostrar de maneira efectiva que o galego é unha lingua funcional e moderna. Campañas de sensibilización e promoción nas que se destaque a dispoñibilidade de información relevante en galego poden cambiar os hábitos das usuarias, animándoas a empregar a súa lingua materna. Ademais, ao colaborar con creadoras de contido, tamén se potencia a elaboración de máis contidos en galego, enriquecendo o ecosistema dixital e asegurando que as usuarias atopen respostas axeitadas ás súas procuras sen necesidade de recorrer a outras linguas.
5. É preciso manter un grupo de conversa coa flutuación que deba ser, mediante lista de correos, poñendo en común as achegas que resulten relevantes. Esta medida garante a continuidade e a efectividade das accións emprendidas. Estes grupos permiten a colaboración constante entre os diferentes entidades involucradas, facilitando a posta en común de ideas, experiencias e resultados. A flutuación no grupo non debe ser vista como un problema, senón como unha oportunidade para incorporar novas perspectivas e adaptar as estratexias segundo as necesidades e os cambios no entorno dixital. Ademais, un grupo de conversa activo pode identificar rapidamente problemas e oportunidades, permitindo unha resposta áxil e coordinada que mellore a eficacia das iniciativas para promover o uso do galego en Internet.

Recomendacións sobre a metodoloxía de análise e aplicación de solucións:

6. A escolma de datos debería realizarse empregando un dispositivo móbil GPS, xa que as consultas desde este dispositivo supoñen o 66% das buscas e o contido amosado é distinto á versión escritorio. A situación respecto a un dispositivo de sobremesa supón unha variable do 90%. Nos estudos consultados, ademais, os resultados ofrecidos varían segundo a plataforma utilizada. Os móbiles permiten unha análise máis precisa e contextualizada, xa que o contido amosado pode variar segundo a localización da usuaria e outros factores como o tipo de dispositivo ou o historial de procuras. É importante ter en conta que as experiencias de busca en dispositivos móbiles e de sobremesa son diferentes, polo que a recollida de datos debe reflectir esta diversidade. Utilizar móbiles con GPS permite recoller datos máis exactos sobre os hábitos e preferencias das usuarias, o que é crucial para a creación de estratexias de SEO eficaces que atendan ás especificidades de cada plataforma e dispositivo.

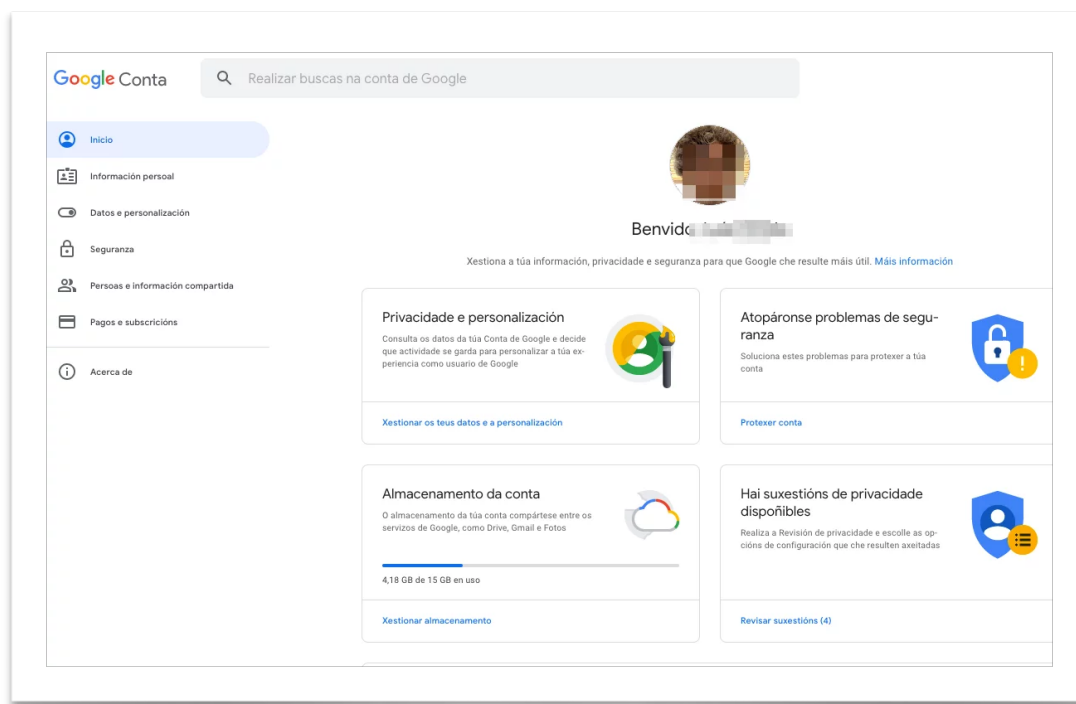
7. Sobre a frecuencia de medición, indicouse que sería máis rendible empregar máis palabras clave e reducir a frecuencia de medición. En vez de realizar medicións constantes con poucas palabras clave, unha estratexia que implique a monitorización dunha ampla variedade de termos pode identificar patróns e oportunidades que poderían pasar desapercibidas nunha análise máis limitada. Esta metodoloxía permite unha mellor adaptación ás variacións estacionais e sectoriais, e proporciona datos máis ricos para a toma de decisións. Ademais, a redución da frecuencia de medición pode diminuír os custos operativos sen sacrificar a calidade da análise, ao centrarse en medicións clave que aporten información relevante.
8. Sobre a selección destas *keywords*, deberían ser varios caixóns segundo intereses sectoriais (festivais, librarías, deportes, servizos públicos como as citas do Sergas...) e estacionalidade ("axudas vivienda", "matrículas universidade"...), poñendo en valor as palabras con intención de compra (produtos galegos gastronómicos, artesanía...), búsquedas locais e as que non teñen tradución en castelán (San Xoán, Samaín...), mais deberían incluírse tamén aquelas formuladas en castelán para captar unha audiencia máis ampla e diversa, especialmente en contextos onde ambas linguas conviven e se complementan. Isto permite unha segmentación eficaz do público obxectivo e a creación de contidos altamente relevantes. As palabras clave con intención de compra son cruciais para atraer tráfico cualificado e converter visitantes en clientes, mentres que as buscas locais poden mellorar a visibilidade de negocios e servizos na área. Ademais, incluír termos sen tradución ao castelán destaca a singularidade cultural de Galicia e fomenta o uso do galego.
9. Sobre os portais bilingües, sería relevante confrontar os resultados en galego e en castelán fronte a mesma procura, e nas búsquedas estacionais, valorar que porcentaxe de idioma ocupa cada posición. Esta análise permite identificar posibles discrepancias na visibilidade e posicionamento de contidos en cada idioma. Ademais, examinar as buscas estacionais e a porcentaxe de idioma en cada posición pode revelar patróns de uso e preferencia lingüística, proporcionando datos valiosos para optimizar a estratexia de SEO e mellorar a accesibilidade e relevancia dos contidos en ambos idiomas. Esta confrontación tamén pode destacar áreas onde o contido en galego necesita reforzarse para competir en igualdade de condicións co castelán.
10. O galego debería existir como opción publicitaria en Google e outras plataformas. Actualmente, a maioría das ferramentas publicitarias non ofrecen soporte completo para o galego, o que limita as oportunidades de chegar ao público galegofalante de maneira efectiva. Integrar o galego nas opcións publicitarias permitiría aos anunciantes segmentar mellor as súas campañas e conectar directamente coa audiencia galega, promovendo produtos, servizos e contidos de forma máis eficaz. Ademais, isto tamén contribuiría á normalización do uso do galego no ámbito dixital, reforzando a súa presenza e relevancia nun entorno cada vez máis globalizado.

Como optimizar os navegadores e buscadores en galego

Google proporciona un conxunto de servizos moi amplo na rede. Ademais do buscador web, conta con correo, *suite* ofimática *online*, servizos de calendario, vídeos (Youtube), etc. Como a autenticación está unificada para todos os servizos, pódese cambiar o idioma para todos ao mesmo tempo, para así configurar o buscador en galego.

Pasos a seguir para configurar os servizos de Google en galego:

1. Abrir o perfil da túa conta.
2. Clicar en “datos e personalización”.
3. En “Lingua”, seleccionar “galego”.
4. Recargara as lapelas que teñas abertas para que se actualicen os cambios.



Pantalla de configuración dos servizos de Google.

No caso dos navegadores, Chrome mostrarase automaticamente no idioma predeterminado do sistema do teu ordenador. Para configurar Windows en galego:

1. Abrir a configuración.
2. Ir á lapela de “idioma e configuración”.
3. Seleccionar “idioma”.

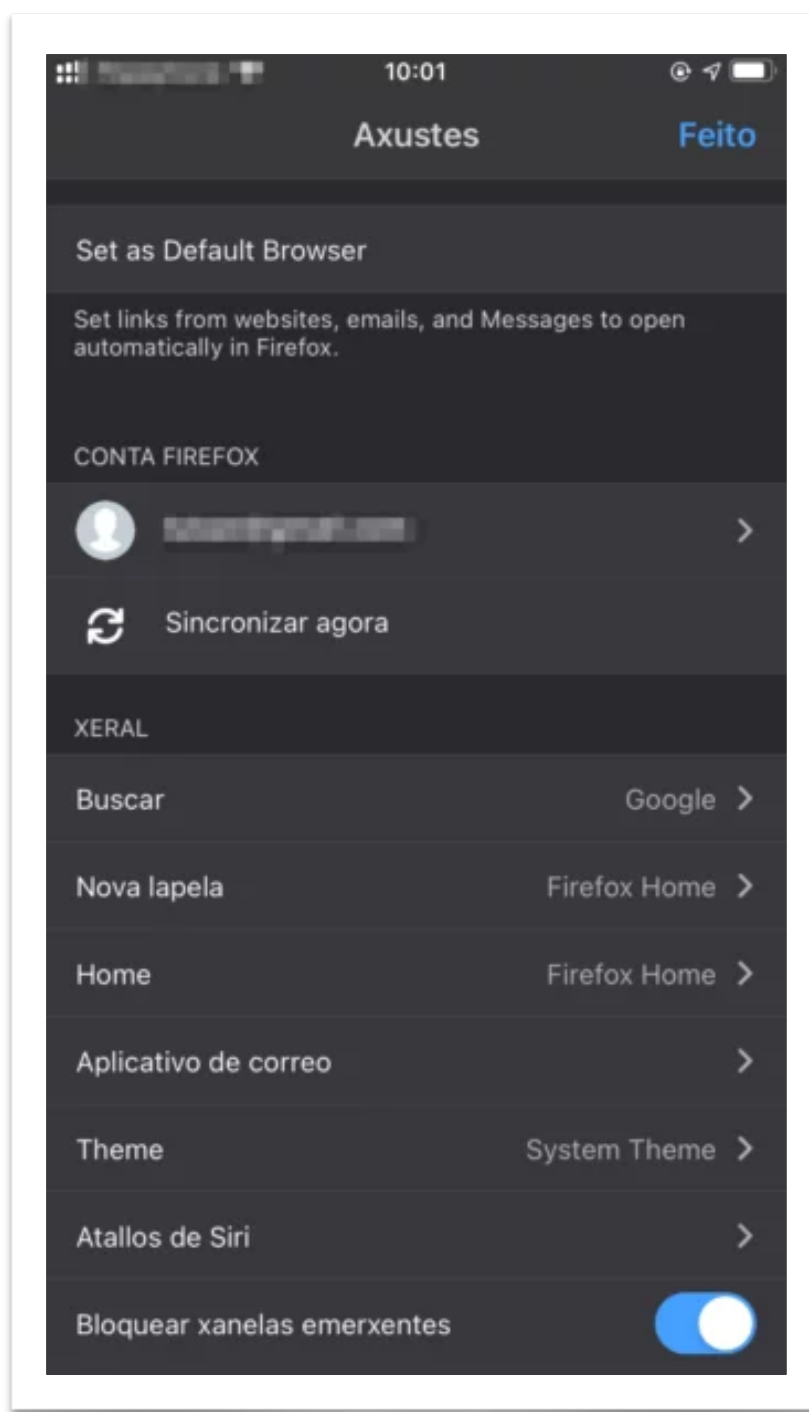
4. En “idiomas preferidos” agregar o idioma galego e darlle a “idioma predeterminado para Windows”.
5. Reiniciar o equipo para gardar os cambios.

Para cambialo dende o propio navegador:

1. Abre Chrome.
2. Arriba á dereita, clicar en nos tres puntos verticais → “Configuración”.
3. Na parte esquerda, clicar en “Idiomas”.
4. En “Idiomas preferidos”, xunto ao idioma que queiras usar, fai clic en “Máis”.
 - a. Se o idioma non aparece, fai clic en “Engadir idiomas” para incluílo na lista.
5. Clica en “Mostrar Google Chrome” neste idioma.
 - a. Esta opción só está dispoñible en Windows.
6. Reinicia Chrome para aplicar os cambios.

En Firefox, a versión en galego está preparada e compilada por Proxecto Trasno. Está dispoñible para Windows, Mac, Linux, IOS e Android. Pasos a seguir:

1. Ir a <https://www.mozilla.org/gl/>.
2. Darlle ao enlace de descarga e gardar o arquivo executable.
3. Darlle permiso para facer modificacións no dispositivo.
4. Aí comezaría a instalación, no caso de teres instalado o Firefox faría unha actualización.
5. Unha vez instalado abrírase o teu navegador Firefox.



Pantalla de configuración de Firefox en Android.

Proxecto de monitorización



Como xa se indicou ao longo do documento, este proxecto ten como obxectivo principal monitorizar o funcionamento das buscas en Google en galego. Ese obxectivo conseguírase, en primeiro lugar, ampliando a comunidade do SEO en galego. E, en segundo lugar, dispoñendo de datos obxectivos propios.

O desenvolvemento dunha rede de equipos propios espallados pola xeografía galega é unha ferramenta con moita potencialidade, non só xa para o caso en concreto que nos ocupa, senón como primeiro paso para monitorizar outras variables e servizos, como plataformas de redes sociais, nas que os cambios son especialmente rápidos, tanto nos algoritmos de selección de contidos coma nos datos que determinan o seu uso: palabras clave, recoñecemento de lingua falada ou ortografías irregulares, etc.

Neste achegamento, PuntoGal fixo unha avaliación da situación actual e das distintas opcións polas que optaron outras comunidades lingüísticas, nomeadamente a catalá, a máis activa neste aspecto. Como consecuencia, optouse por unha dobre vía de traballo.

A primeira acción foi o lanzamento dun piloto coa instalación dun pequeno número de dispositivos, dez, todos eles en distintas ubicacións. Este piloto fíxose mediante un convenio de colaboración coa Fundació PuntCat. Deste xeito puidéronse visualizar as posibilidades do sistema, pero tamén os retos tecnolóxicos e loxísticos aos que necesitará responder o programa de PuntoGal. Un piloto destas características implicou a creación dun corpus de termos e a colaboración co equipo de PuntCat para establecer un marco de exploración dos datos que permita compartilos e comparalos no futuro. Independentemente do modelo polo que opte cada dominio, un obxectivo do proxecto é que distintos territorios contén con información que se poida compartir nunha rede común. O resultado deste piloto pode lerse máis adiante neste informe.

A segunda liña de investigación foi o desenvolvemento do concepto dun dispositivo propio, en colaboración con empresas do ámbito TIC galego. PuntoGal solicitou o esquema detallado dun hardware que cumpra coas necesidades do proxecto desde a perspectiva galega e que se poida integrar nunha rede de datos xeral. As características deste equipamento e o seu plan de implementación están descritas con máis detalle máis adiante neste informe.

As aprendizaxes e as conclusións destas dúas actividades permitirá planificar os seguintes pasos do proxecto nos vindeiros meses.

O proxecto catalán

No ano 2023, detectáronse cambios nas estatísticas de visitas de diversas webs bilingües. Os resultados das buscas, sobre todo en Google, deixaban a impresión de que se estaban a priorizar as versións das webs en español fronte aos mesmos contidos en linguas cooficiais. Pronto algúns webmasters constataron que estas diferenzas se traducían nas métricas das webs, sendo especialmente patentes nos casos en que estas tiñan un mantemento de contidos bilingüe de xeito habitual.

Así foi como diferentes entidades que traballan para a promoción e defensa da lingua catalá uníronse para impulsar un pacto que promova a presenza do seu idioma nos diversos ámbitos do mundo dixital.

Detrás da Alianza pola presenza dixital do catalán están as entidades Amical Wikimedia, Acción Cultural do País Valencià, Fundación.cat, Institut d'Estudis Catalans, Institut Ramon Llull, Obra Cultural Balear, Òmnium Cultural, Plataforma per la Llengua, Softcatalà e WICCAC.

O secretario técnico da *Aliança* é o xornalista Albert Cuesta, quen define que o obxectivo da Alianza é “representar conxuntamente os dereitos lingüísticos da sociedade civil e catalanofalante no ámbito dixital para que sexan respectados polas empresas, plataformas e as organizacións de todo o mundo nos produtos e servizos que ofrecen á cidadanía”. Identifican oito ámbitos dixitais onde a presenza do catalán non é satisfactoria: contido web, aplicacións móbiles, redes sociais, plataformas audiovisuais, automoción, electrodomésticos, asistentes de voz e intelixencia artificial.

En primeira instancia abórdase o problema cos buscadores, sobre todo Google, que penaliza a presenza de páxinas web en catalán nos resultados das buscas, mesmo se a usuaria ten os dispositivos configurados en catalán, afectando negativamente á visibilidade do contido web en lingua catalá.



Sesión de traballo dos equipos de PuntoGal e PuntCat sobre os sensores do programa piloto de monitorización.

Co fin de reverter a situación, a **Aliança** propón tres liñas de actuación: diagnóstico da situación baseado en datos; interlocución coas entidades públicas e privadas que determinan a presenza do catalán e seguimento continuado desta presenza para verificar os efectos das accións de mellora. Ademais, traballarán para detectar o antes posible posibles deterioracións.

O primeiro paso foi a creación dun dispositivo que cuantifique a perda actual de visibilidade do catalán nas buscas web durante os últimos meses e sirva para reclamar solucións ás empresas dixitais e detectar incidencias similares no futuro.

Para tentar identificar as flutuacións no algoritmo, desenvolveron un sistema de monitorización baseado en Raspberry Pi con Chromium en catalán, lanzando peticións horarias de palabras clave e almacenando as primeiras 50 posicións de busca para análise posterior.

Actualmente contan con 60-70 dispositivos, cada un cunha Raspberry Pi e conexión específica, espalladas xeograficamente por territorio español e máis concretamente, centrados nas zonas nas que se fala catalán.



Reunión do equipo de PuntoGal con Roger Serra, da Fundació .cat

A proposta de monitorización de PuntoGal

1
0
0 0
1 1
0 0 1 1
1 1 0 1 0 1 1 0 1 1 1 1
0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 1
1 0 0 1 1 0 0 1 0 0 1 0 0 1 1 0 0 1 0 0
0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 1 1 0 0 1
1 0 1 0 0 1 1 0 1 0 1 0 0 0 0 1 1 0 0 1 0 0 0
0 1 0 1 1 1 0 0 1 0 1 0 0 1 0 0 0 1 1 1 0 1 1 0 1 0 0 1
0 0 0 0 0 0 1 1 0 1 0 1 1 0 1 1 0 0 0 0 0 1 0 0 1 0 1 1 0
1 1 1 0 1 0 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 1 0 1 0 1 0 1
0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 1 0 1 1 0 1 0 0 0 1 0 1 0 0 0
1 1 0 0 0 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0
0 0 1 0 1 1 0 0 1 0 0 1 0 0 1 1 0 0 0 0 1 1 0 1 1 0 0 1 0 0 1

Desde PuntoGal percíbese a necesidade de medir a resposta dos buscadores, especialmente de Google, con respecto ao galego. Para isto, como parte dunha actuación que tamén implica a análise do tráfico das webs máis relevantes, propónse recoller os resultados de buscas orgánicas do xeito máis similar posible a unha navegación humana e natural, pero mantendo o meirande número de variables controladas, de xeito que exista o mínimo nesgo nos resultados analizados.

O proxecto ten como obxectivo monitorizar os resultados das buscas dunha serie de termos no buscador Google en galego, obtendo as primeiras 50 posicións de cada palabra clave indicada. A busca realizarase desde dispositivos propios (hardware específico), cada un cunha IP e localización xeográfica distinta en base a onde vaian finalmente colocados.

Xeraranse históricos de datos con consultas diarias para a súa posterior análise, con intención de ter en conta, entre outras, as variables de localización tal e como as teñen en conta os buscadores, cunha distribución progresiva por toda a xeografía galega, pero sen desbotar os emprazamentos nos que se detecta unha presenza máis alta de falantes de galego tanto na península e nas illas como no resto de Europa e América, nomeadamente do sur.

O alcance do proxecto, de darse as condicións indicadas, pasa pola acumulación de datos sobre estes termos de busca de xeito continuado durante varios anos, arquivando os resultados e as variacións dos mesmos para poder comparalos en varios eixes de traballo.

Por unha parte, a variable xeográfica, que se entende que pode provocar a aparición de distintos resultados, sobre todo en buscas relacionadas con produtos e servizos que teñen máis relevancia local. Pensamos, por exemplo, no comercio, a hostalería ou os servizos de saúde. Esta variable, que en principio debería ser fixa, ao interpretarse tal e como a perciben os buscadores, pode variar a medida que os buscadores afinen as súas ferramentas ou que as fornecedoras de acceso a Internet informen sobre os enderezos IP que están usando ou a relevancia que se lles dea ás *cookies* e outras tecnoloxías de rastrexo.

Tamén a variable temporal e estacional se terá en conta. Sabemos que buscar a vacina da gripe é moito máis común no outono que na primavera, ou que a información deportiva, sobre todo a que ten que ver con retransmisións televisivas ou radiofónicas, é moito máis consultada os venres que os martes, etc. A existencia de datos ao longo de varios anos nos que estas variables fosen consideradas permitiría a comparación entre distintas versións dos algoritmos dos navegadores, o peso que se concede á lingua das buscas e outros datos interesantes.

Ademais de manter unha serie de termos de busca para poder así manter illados unha serie de factores de comparación, propónse tamén unha serie de termos tanto estacionais coma dependendo da actualidade informativa o da relevancia momentánea, para dispor de datos que van máis alá da lectura académica dos datos e, na medida do posible, vinculados ao que pasa noutras áreas de rede, coma as redes sociais.

A relevancia estatística de cada toma de datos, deste xeito, vai medrando a medida que se van acumulando, permitindo ir alimentando un modelo que cada semana ou cada mes permita sacar conclusións ou determine novas variables que considerar. O progreso destas

interpretacións determinará tamén a publicación dos mesmos, tanto para o público xeralista, cando poida ser de interese, coma en ámbitos máis pechados. Entendemos, por exemplo, o especial interese que terán os datos relacionados co idioma dos resultados para a Secretaría Xeral da Lingua.

Tendo en conta que PuntoGal tamén colabora con institucións que están en proceso de estudar os mesmos resultados de busca noutras linguas, estes datos pódense comparar tanto en termos absolutos, coma nas variables que se consideren interesantes ou relevantes á hora de tomar decisións sobre actuacións accións de comunicación, en común ou de xeito illado.

Segundo se explicou neste documento, Google agora mesmo non recoñece o galego coma un idioma en termos de captación publicitaria. Isto quere dicir que non hai una industria do SEO en galego, como si existe en castelán ou en catalán, coas consecuencias que isto ten para as empresas e para os creadores de contido. Todo o traballo de investigación que fagamos para entender como tratan os buscadores os resultados en galego será camiño andado para conformar un sector forte cando chegue o momento e esta posibilidade de mercado sexa realidade.

Os obxectivos específicos que se cubrirán co proxecto son:

- Verificar a atribución autonómica por IP.
- Analizar se as buscas cambian segundo a IP ou localización xeográfica.
- Conectar o sistema a servizos externos para interpolar datos obtidos con información web.

Requisitos cos que deberá contar o proxecto

Identificáronse os seguintes requisitos funcionais:

- Posibilidade de actualizacións remotas do software.
- Acendido e apagado remoto e tamén automático dos equipos.
- Uso de *cookies* e configuración do navegador en galego.
- Principais palabras clave a monitorizar: nomes propios e de produtos que sexan comúns en galego e castelán.
- Recollida de datos nun panel de control para posterior análise, con capacidade de exportación da información, por exemplo, a CSV.

Prototipo: solución aportada

Para a análise desta solución, PuntoGal contou coa colaboración da consultora Sixtema, que aportou a súa experiencia no deseño de redes de sensores para poñer en marcha unha serie de dispositivos que teña capacidade de medrar en número sen perder de vista as diferentes variables de ubicación, intervalos de funcionamento, etc.

Analizadas as necesidades anteriores, considérase oportuno facer un primeiro prototipo baseado en dúas partes ben diferenciadas:

- Contar cun dispositivo físico desde o que se tramiten as peticións de palabras clave ao buscador Google.
- Contar cun software específico que simule estar nunha localización determinada e desde ela facer as peticións de palabras clave ao buscador Google.

A continuación, amósanse as particularidades de cada unha das anteriores opcións, que poderán formar parte do prototipo inicial a desenvolver.

Prototipo: hardware

No relativo ao prototipo de hardware, identifícase a Raspberry PI 4 8GB RAM como o dispositivo óptimo para levar a cabo o proxecto. Trátase dunha placa computadora (SBC) de baixo custo desenvolvida no Reino Unido pola Fundación Raspberry Pi co obxectivo de estimular a ensinanza de ciencias da computación nas escolas. Ten un prezo moi accesible, un tamaño moi manexable e moi baixo consumo eléctrico.

Características do hardware escollido:

- Dispón dunha ampla documentación.
- Sistema operativo Linux completamente funcional.
- Posibilidade de monitorización con Zabbix mediante o axente oficial.
- Custo razoable para o seu gran potencial.

Características técnicas do dispositivo:

- Procesador Quad-Core. Procesador de 64 bits con un rendemento elevado para as tarefas que se precisan levar a cabo.
- Conectividade:
 - 2 portos USB 3.0.
 - 2 portos USB 2.0.
 - gigabit Ethernet.
 - Wi-Fi.
 - Bluetooth.
 - Soporte de almacenamento. Capacidade de usar discos NVMe M.2 para o almacenamento.
 - Compoñente de almacenamento proposto para o prototipo: Kingston NVMe PCIe 250GB.
 - Saída de vídeo dual, con dúas saídas micro HDMI que permiten a conexión a dous monitores en 4K.
 - Sistema Operativo Linux. Permite un control avanzado e a implantación de sistemas de monitorización centralizados como o uso de Zabbix, que nos dará un panel personalizado e moi completo sobre o estado dos dispositivos.
 - Caixa de aluminio con ventilador silencioso. Permite a protección e selado dos compoñentes que forman parte do dispositivo, así como unha constante refrixeración do mesmo cando está en funcionamento.

No tocante á configuración coa que contará o dispositivo (nesta fase de prototipo), enumeramos as seguintes:

- Sistema operativo. O dispositivo disporá de un sistema operativo Linux que aporta seguridade e capacidade de xestión ao proxecto.
- Actualización do sistema. As actualizacións serán xestionadas remotamente desde un servidor central.
- Monitorización dos dispositivos. Os dispositivos disporán de un sistema de monitorización constante, con gráficas en tempo real do estado do hardware e con alertas sobre posibles problemas a curto e medio prazo.

Prototipo: software

O obxectivo de facer o prototipo de software é o de desenvolver un sistema que realice consultas automáticas a Google para monitorizar as palabras clave especificadas, gardar os resultados do top de busca nunha base de datos e proporcionar un panel de administración para xestionar as palabras clave e visualizar os resultados.

Este desenvolvemento non precisaría dun dispositivo físico para o seu funcionamento (a Raspberry Pi mencionada anteriormente) e permitirá simular a consulta a Google desde distintas localizacións.

Así, o sistema previsto permitiría realizar consultas automáticas a Google para monitorizar as palabras clave especificadas, gardar os resultados do top 50 de busca nunha base de datos e proporcionar un panel de administración para xestionar os termos e visualizar os resultados.

As chamadas simulando localizacións a través de exclusivamente software faranse a través dunha API (de pago). O seu uso será compatible coa existencia de dispositivos físicos.

Alcance do proxecto

O alcance do proxecto sería como segue:

- Consultas a Google. Realización de múltiples consultas por hora para a palabra clave introducida a través da API que simula diferentes localizacións.
- Almacenamento de resultados. Gardar os resultados do top 50 de busca en base de datos, tanto os obtidos vía API (software exclusivamente) como os obtidos vía dispositivo de hardware (Raspberry Pi).
- Integración da comunicación co prototipo de hardware.
- Xestión das palabras clave de aplicación no dispositivo.
- Xestión remota para a actualización do sistema no dispositivo.
- Comunicación vía API para o gardado dos resultados obtidos das procuras realizadas.
- Panel de administración.
- Entrada e xestión das palabras clave.
- Visualización dos resultados das consultas.
- Exportación a CSV dos resultados obtidos das consultas.

Tarefas a levar a cabo no desenvolvemento:

Enuméranse a continuación; as distintas tarefas que se identifican como parte do desenvolvemento de software proposto:

- Análise para seleccionar a arquitectura do sistema máis óptima para o desenvolvemento.
- Esquema do modelo de datos.
- Deseño das interfaces de usuaria.
- Sistema automático e xeolocalizado de consultas.
- Paneis de usuaria.
- Mecanismo de *scraping* mediante uso de navegador xeolocalizado.
- Control de frecuencia de consultas e probas.
- Exportación de datos.
- Sistema automático de consultas a través de API.
- Mecanismo de scraping mediante uso de API.
- Probas.
- Xestión remota para actualización do sistema.
- Modificación das palabras clave que serán consultadas en Google a través dos dispositivos físicos.
- Captura da información vía API para a lectura da información dos dispositivos hardware.
- Control de calidade.
- Documentación: Creación de documentación técnica e de usuaria.

Prototipo: termos de busca

Os dispositivos permiten facer unha serie de buscas a intervalos precisos. Para manter os resultados dentro do espectro das buscas feitas por humanos, o ritmo das mesmas debe ser pausado e suxeito a horas determinadas, usando os mesmos intervalos e xeito de busca que se as estivera facendo unha persoa interactuando cun ordenador ou un móbil. En caso contrario, o buscador pode inferir que é un proceso automatizado e cambiar as súas respostas, deixar de responder e mesmo colocar ese enderezo IP nunha lista negra para outros servizos.

Tendo en conta a recomendacións das expertas e a experiencia doutros dominios, decidiuse a inclusión das seguintes palabras clave no grupo de buscas permanentes do dispositivo.

bus coruña
sergas
depor
abono renfe
celta
ikea porto
intermodal vigo
termas ourense
meteogalicia
tvlg
xunta
galiña azul
usc
udc
uvigo
camiño de santiago
galicia
galiza

Ao mesmo tempo consideramos que hai determinadas palabras que podemos ir alternando de xeito estacional e que entendemos que serán máis relevantes, desde o punto de vista estatístico, se as imos repetindo ano a ano durante os períodos nos que os contidos na rede sobre estes temas aumentan. Nos casos elixidos teñen que ver con festividades, prazos administrativos ou prevención de enfermidades estacionais, por exemplo:

vacina gripe
luces vigo
vacuna gripe
magosto
san xoan
matrícula USC
Selectividade Galicia

Estes dous grupos de buscas complementaríanse con termos puntuais que, aínda que non sexan especialmente frutuosos desde o punto de vista estatístico, poden resultar interesantes nun momento histórico ou nun contexto informativo especial, como períodos electorais, probas deportivas, etc.

Fase 2: seguintes pasos

Despois de por en marcha o prototipo de hardware, se porá en marcha o piloto. Varios dispositivos se activarán en ubicacións escollidas para que lancen as buscas preliminares dos primeiros termos de busca. Estas buscas pódense lanzar desde unha API que ten en conta as distintas localizacións.

Así, terminada a fase 1 e analizados os resultados obtidos, poderanse sumar novos dispositivos ao sistema, aproveitando o coñecemento desta fase para mellorar o sistema na súa fase 2.

A base da fase 2 será o realizado na fase 1, sumando os cambios identificados tanto a nivel de software como de hardware, pero engadindo a maiores, previo consenso, as medidas de seguridade que non están presentes no dispositivo na fase 1:

- **Seguridade na conectividade. Medidas a aplicar:**
 - Chaves SSH. Só se permitirá a autenticación con claves SSH para incrementar a seguridade, evitando o uso de claves de usuario.
 - Uso de Fail2ban. Usarase Fail2ban como sistema de bloqueo de acceso para previr ataques de forza bruta nos servizos SSH.
 - Devasa (Firewall). Para protexer os dispositivos a nivel de rede usarase unha devasa que se configurará directamente sobre o sistema operativo.
- **Xestión remota. Medidas a aplicar:**
 - Túnel SSH permanente. A Raspberry Pi manterá un túnel SSH sempre activo cun servidor central, mediante o uso de un servizo en systemd e un túnel SSH.
 - Xestión mediante Ansible. Utilizarase Ansible e o túnel SSH para acceder e xestionar os dispositivos remotamente, permitindo a actualización e configuración de software de forma remota e centralizada.
- **Protección de acceso físico. Medidas a aplicar:**
 - Cifrado de disco. As unidades de almacenamento estarán cifradas para protexer os datos en caso de subtracción do NVME.
 - Inhabilitar acceso local. Desactivarase o acceso local aos usuarios do sistema, impedindo calquera inicio de sesión que non se faga desde o servidor central.
 - Selado de portos. Os portos non utilizados poden ser selados para impedir accesos non desexados, por exemplo os portos USB poderán ser selados con epoxi e incluso fixar o que se usa para a conexión do NVME.

A proba piloto

En que consiste o piloto?

O piloto de monitoraxe do galego dixital é o primeiro paso dun proxecto pioneiro que ten como obxectivo medir os cambios de funcionamento nos algoritmos dos principais motores de busca que poidan afectar aos resultados en galego. A presenza do galego no mundo dixital é crucial para a súa supervivencia e promoción. Con frecuencia, non se dispón de datos sobre como aparecen os contidos en galego en motores de busca como Google.

Grazas a unha rede de sensores distribuídos estratexicamente por todo o territorio de fala galega, a rede do piloto recolle datos sobre a visibilidade dos contidos en galego e permite identificar tendencias, detectar problemas e avaliar a eficacia das distintas estratexias para a promoción do galego na rede.

Este primeiro piloto consta de dez sensores distribuídos pola xeografía galega e por zonas de fala galega fóra da comunidade. Estas áreas fronteirizas son especialmente relevantes porque desde elas os buscadores predominantes na rede non teñen predeterminado o seu funcionamento e se entende que a variable que ten que ver co territorio queda neutralizada, sen ser determinante nun sentido ou noutro.

Os emprazamentos destes dez sensores son:

Santiago de Compostela

A Coruña

Vigo

Verín

Lugo

Ferrol

Monçao

Pobra de Seabra

Forcadela

As Ellas (Badaxoz)

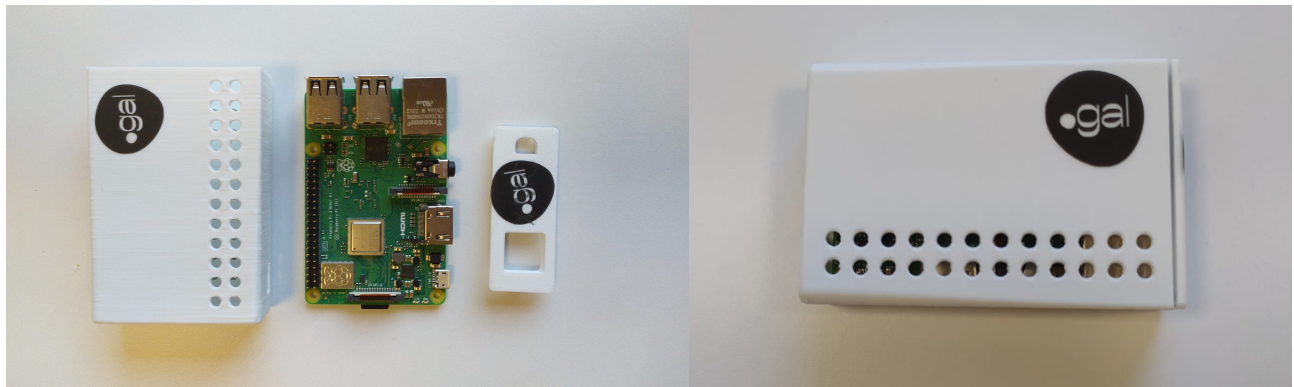
Tui

Como funciona?

Os sensores simulan o funcionamento dun ordenador ou dun móbil autónomos configurados en galego, actualmente cos navegadores Chrome e Firefox co galego como idioma preferente. O dispositivo realiza buscas arbitrarias partindo dunha batería de buscas preestablecida polo equipo de PuntoGal e, simulando o uso humano, realiza 4 buscas por hora durante 11 horas ao día, comunicando aos servidores os 10 primeiros resultados de cada busca. Isto fai que recibamos un gran volume de resultados ao día e poidamos ter unha mostra significativa tras uns meses para analizar os resultados.

Inicialmente, o piloto foi pensado como unha aplicación que puidese descargarse no ordenador e no buscador, e que fose desde esa onde se realizasen as buscas e se comunicasen os resultados, pero houbo un impedimento: non podíamos garantir que a configuración do dispositivo onde se instalase a aplicación fose a correcta. Se o usuario nin sequera ten configuradas as preferencias do seu entorno dixital en galego, non se lle pode pedir a Google que adiviñe que prefere contido en galego.

O dispositivo do piloto é unha Raspberry Pi, un aparello de baixo custo. A súa instalación é moi sinxela: só é preciso conectalo á corrente e darlle acceso á rede. O dispositivo queda en depósito coa persoa ou entidade correspondente e non se pode usar para outros fins. No caso de detectar un funcionamento diferente ao establecido, determinarase se é necesario retirar de forma permanente o dispositivo.



Dúas imaxes dos dispositivos de monitorización do piloto instalados.

Cada sensor consta dunha placa Raspberry Pi modelo 3B+. Executa o sistema operativo Raspbian baseado en Debian 12 (Bookworm). En relación á rede, cada dispositivo está conectado á rede central mediante unha VPN cifrada baseada en WireGuard, que garante unha conexión segura co servidor central. Os sensores só precisan conexión eléctrica e acceso á rede a través de cableado, sen acceder á rede privada do colaborador.

Finalmente, cada sensor dispón de credenciais individuais e privadas para observar e gardar datos na base de datos (PostgreSQL). Deste xeito, cada sensor só pode interactuar cos seus propios datos e non observar os doutros sensores, asegurando así a privacidade dos datos reportados ao sistema.

Comportamento dun sensor

Os sensores despregados realizan todo o proceso de monitoraxe cada 15 minutos, entre as 9:00 e as 22:00. Este proceso consiste en:

1. O sensor consulta ao servidor central os seguintes datos: tipo do dispositivo (móbil ou ordenador), navegador (Chrome ou Firefox) e motor de busca (Google). Execútase unha función na base de datos PostgreSQL onde están definidos os diversos aspectos de configuración para devolver unha modalidade concreta, de xeito aleatorio ou condicionado.
2. Unha vez que o sensor dispón de todos os datos e métricas de configuración, este volve buscar na base de datos do servidor central cal será a busca a executar coa función seguinte_busca(). Esta, igual que as anteriores, executa unha función na base de datos PostgreSQL onde recibe a busca a realizar. Esta función, a diferenza das anteriores, está bastante condicionada. Depende do número de buscas realizadas polo sensor e do número de veces que executou cada unha delas. Se hai buscas moito menos executadas que outras, a función tenderá a devolver esas buscas primeiro para equilibralas.
3. Unha vez o sensor ten toda a información necesaria para construír o navegador, iníciase co axente de usuario correspondente e busca o motor designado. Feito isto, execútase todo o proceso de rastrexo. Primeiro escribe o termo a buscar na barra do buscador e obtén os resultados. Unha vez obtidos, o sensor garda os datos dos 10 primeiros resultados na base de datos do servidor central. Se non pode obter os 10 resultados na primeira páxina do buscador, busca o número de resultados que faltan na segunda páxina. Para cada un dos resultados, gárdase: sensor, hora, navegador, buscador, posición, título, URL, descrición, a busca que se executou para obtelo e o axente de usuario do navegador.

O idioma dos resultados valórase máis tarde por parte do servidor central para non sobrecargar os sensores.

Neste proceso de monitoraxe, determinouse que todos os resultados clasificados como "Novas" non forman parte dos 10 primeiros resultados. Isto débese á posible variabilidade destas novas e a resultados realmente pouco significativos dos datos.

Servidor central

O servidor central é onde se gardan toda a información recollida polos sensores deste piloto. Este servidor, xestionado pola Fundación .cat, está conectado con todos os sensores da rede e xestiona todas as operacións que estes realizan, como indicar que buscas deben realizar, con que motor de busca, con que navegador ou con que tamaño.

A comunicación entre o servidor central e os sensores faise a través da VPN configurada con WireGuard. A comunicación é bidireccional entre o servidor central e cada sensor, pero os sensores non poden comunicarse entre si.

O servidor central tamén executa as seguintes tarefas:

1. Detección do idioma dos resultados
2. Monitoraxe da actividade dos sensores
3. Monitoraxe do estado de conexión dos sensores

Accións necesarias

Para que os sensores funcionen, deben estar conectados á corrente e dispoñer dunha conexión a Internet mediante un cable cara ao router. É frecuente que algún destes dous requisitos non se cumpra, polo que os sensores poden requirir accións manuais por parte do seu anfitrión. Aínda que cada sensor está configurado para reiniciarse automaticamente á unha da madrugada, ás veces o sensor queda bloqueado e deixa de executar ordes. Doutra banda, pode ser que o sensor se desconecte fisicamente da rede ou da corrente.

En calquera destes casos, requirirase que o anfitrión do sensor reinicie o aparello de maneira activa. Para facelo, hai que desconectar o aparello e esperar uns segundos ata que todas as luces se apaguen. Ao conectalo de novo, hai que observar na parte frontal do sensor, onde non está o cable de rede, que aparecen dúas luces diferentes: unha vermella que indica que o sensor ten corrente e está en marcha, e unha amarela que indica que está funcionando correctamente e non ten problemas de hardware.

Existe a posibilidade de que o sensor se dane nalgún momento, xa que as placas Raspberry son delicadas e ás veces, no transporte ou na manipulación destas, poden romperse conectores e pezas aparentemente irrelevantes pero vitais para o seu correcto funcionamento. No caso de que o sensor non estea reportando datos, se ao iniciar o sensor non se visualiza ningunha luz amarela na parte frontal, ao lado da luz vermella da corrente, determinarase que o sensor se estragou e iniciárase o proceso para intercambialo por un novo.



Panel de monitorización da dispoñibilidade de sensores do piloto

Por este motivo, todos os anfitrións dun sensor están suxeitos a ser contactados para realizar reinicios manuais dos sensores. Este contacto realizarase mediante correo electrónico, automático ou manual dependendo do caso, e, se é necesario, mediante chamada telefónica para resolver posibles dúbidas.

É importante ter en conta que, coa distribución descentralizada destes sensores, a súa posta en marcha e o seu mantemento operativo é en si mesma unha tarefa constante: Por distintos motivos, como poden ser a dificultade de acceso aos routers ou ás fontes de alimentación en contextos domésticos, a complicación de actualización de equipos de rede, a presenza de limitacións na conexión por parte de ISPs, etc, o normal é que no momento da conexión dos sensores haxa ata un corenta por cento deles que non funciona no primeiro intento. Aínda que despois da intervención dos técnicos do proxecto se chegue a unha dispoñibilidade total do parque instalado, a monitorización do mesmo obriga a intervencións case diarias para mantelo en funcionamento.

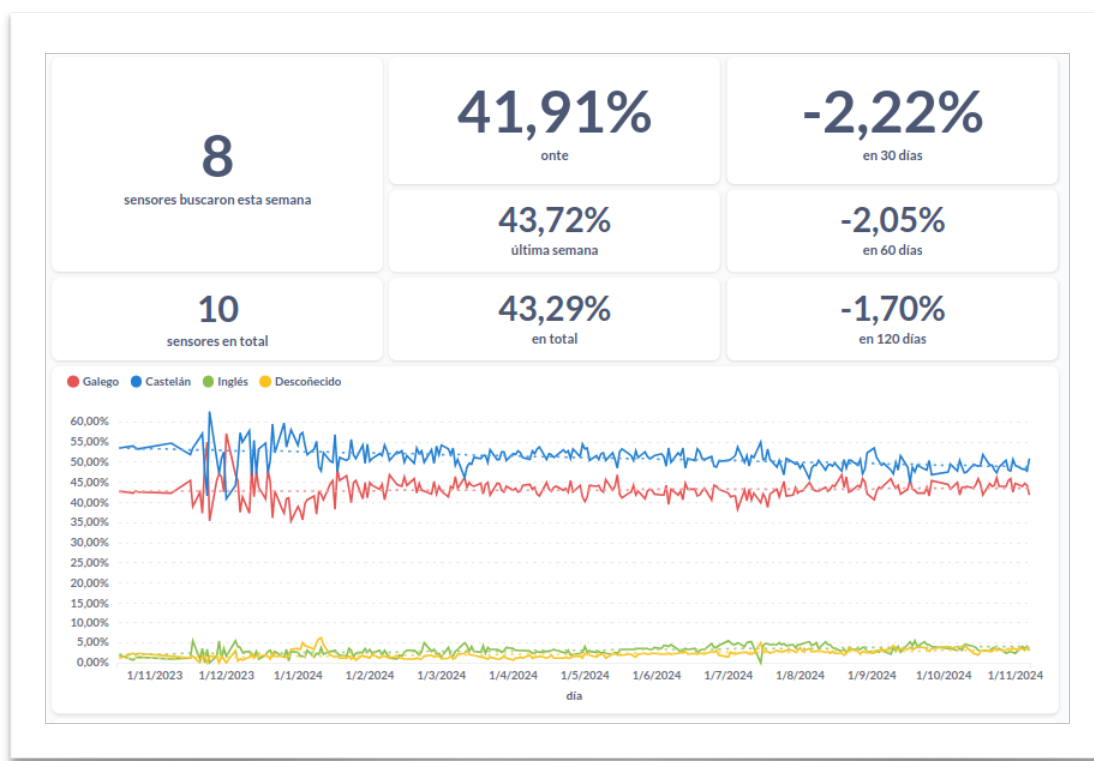
Finalmente, existe a posibilidade de que a rede onde se instale o sensor teña unha devasa que non permita a saída de datos do sensor cara ao servidor central. Nestes casos, a localización considérase inválida e búscase, prioritariamente, unha nova localización. Se non se atopa ningunha, creárase unha excepción dentro da devasa para permitir esta actividade.

Primeiros resultados do piloto

A primeira consideración antes de aventurar ningunha conclusión relacionada co obxecto do estudo é que a proba piloto persegue en realidade coñecer as dimensións do traballo que se quere levar a cabo a medio e longo prazo.

Aínda que nas semanas que os sensores levan instalados se obtivera xa unha cantidade de buscas que pode parecer razoable, o feito é que non é estatisticamente relevante e que resultaría temerario tirar conclusións delas.

A intermitencia das conexións dos sensores, a protocolización da súa instalación e posta en marcha e recollida de datos, así como a adaptación dos termos de busca obrigou a ir modulando os parámetros do sistema de monitorización dos resultados de xeito constante.



Panel de resumo semanal da recollida de datos polos sensores do piloto

O coñecemento obtido deste piloto, en cambio, ten que ver coa planificación a futuro do estudo e a permanencia da monitorización no tempo para poder ir ofrecendo ademais unha perspectiva temporal da evolución dos resultados das buscas, permitindo unha visión dos cambios nos algoritmos de busca e da súas consideración das diferentes linguas.

Neste sentido, só se poderá xulgar a validez do piloto pola relevancia destas consideracións metodolóxicas a medida que avance o proxecto e pola oportunidade de incorporar os datos recollidos ao corpus xeral cando este teña unha validez estatística comprobada.

Próximos pasos e conclusión

A pervivencia dun idioma está vencellada co seu uso e presenza nas redes. E, nese marco, os buscadores son ferramentas centrais. Este informe, primeiro elaborado por PuntoGal sobre o tema, pon de manifesto un panorama no que as grandes tecnolóxicas, especialmente Google, determinan o que ten e o que non ten visibilidade. Esa condicional implica consecuencias para as empresas, as institucións, a cidadanía en xeral; en resumo, para un país e a súa lingua.

Os algoritmos e a intelixencia artificial determinan se un produto aparecerá na primeira páxina de resultados ou se, simplemente, será invisible. Os efectos económicos e de visibilidade para o idioma son facilmente entendibles. O estado da arte que PuntoGal pon de manifesto nas páxinas anteriores conduce cara unha conclusión principal: a necesidade de monitorizar o sistema, de coñecer o funcionamento dos buscadores, de fornecer guías para as usuarias, de empregar todas as ferramentas ao noso alcance para influír na medida do posible no que aparece nas pantallas de teléfonos e ordenadores.

Máis concretamente, precísase:

1. Un monitoreo constante dos resultados de Google e doutros buscadores principais. Esta tarefa permitirá coñecer o comportamento das usuarias e do buscador a partir dunha serie de termos elixidos e actualizados de forma permanente. A información pode xerar estratexias de traballo, servir de fonte de información e presión ante o buscador e, finalmente, incorporar ao galego a un rede de linguas minorizadas que traballen conxuntamente neste ámbito
2. A creación de grupos de traballo do ámbito do SEO en galego. A elaboración deste documento permitiu dar un primeiro paso, pero é necesaria a existencia dunha entidade ou grupo de traballo permanente arredor do SEO no noso idioma. Se non existe unha masa crítica arredor do posicionamento do galego, o progreso nesta materia será artificial.
3. O apoio decidido do país ao uso do galego na rede, sexa nas páxinas web, sexa nas redes sociais ou nas aplicacións. Este apoio necesita da cidadanía, das empresas, das institucións e, moi especialmente, da administración pública. Esta última debe favorecer iniciativas no ámbito do galego na rede de forma decidida, promover e liderar unha enérxica estratexia con políticas valentes.

En PuntoGal existe unha fonda preocupación polo idioma e, máis concretamente, pola súa supervivencia e evolución no ámbito tecnolóxico. É por iso que se puxo en marcha este proxecto, do que este informe é tan só o primeiro paso. Nos últimos meses, o equipo do dominio mantivo diversas reunións de traballo con actores implicados na materia e deu os primeiros pasos para a creación dunha rede propia de monitorización das buscas. Esta rede é agora aínda un piloto, con dez dispositivos recentemente instalados en varias ubicacións relevantes. De momento, os datos non son estatisticamente significativos, pero si permitiron abordar e coñecer a problemática, a loxística e a planificación para a creación desta rede de sensores.

Nos vindeiros meses, o dominio despregará máis dispositivos en novas localizacións xeográficas, irá refinando e ampliando os termos de busca e publicará os primeiros informes cos datos recollidos polos equipos. Para este traballo, implicará a outras organizacións vencelladas coa lingua e a tecnoloxía. A importancia de contar cun mecanismo propio de análise das buscas de Google e outros buscadores non reside unicamente na posta á disposición da cidadanía de datos propios pensados e traballados desde Galicia e o galego; tamén é un primeiro paso para dimensionar unha rede de captación de información ampliable a outros campos. Dito doutra forma, a rede de sensores constituirá un primeiro elo no ámbito dos dereitos dixitais lingüísticos. A monitorización permítenos establecer o punto de partida de distintas problemáticas que, a continuación, poderán ser abordadas polos actores sociais.

Unha vez rematada a primeira parte do traballo, este documento de estado da cuestión, o equipo de PuntoGal desenvolverá un calendario de traballo para 2025 que incluírá os seguintes fitos:

- Finalización do piloto e análise dos datos: Revisión e procesado dos datos fornecidos polos sensores xa instalados en 2024 para unha primeira exportación de información. Se ben non constituirá aínda un informe estatisticamente relevante, si achegará tendencias e amosará o camiño polo que transitar na solución tecnolóxica que finalmente se adopte
- Selección do dispositivo que PuntoGal creará para a posta en marcha da Rede de Monitorización do Galego Dixital (RMGD). Isto farase tendo en conta o resultado tanto do piloto como da asesoría en materia técnica coa que está a traballar o dominio. O dispositivo final ten que ser quen de facer este traballo de forma eficiente, autónoma, ter capacidade de aportar datos para outro tipo de informes e ser escalable en canto a prezo para poder multiplicar o número de dispositivos instalados
- Fabricación e instalación nas localizacións seleccionadas do primeiro grupo de equipos propios, elaboración dun calendario de implantación e dun mapa de localizacións
- Inicio da captación de datos dos equipos instalados e desenvolvemento dunha web ou landing para amosar os resultados, unha vez se dispoña deles
- Presentación do primeiro informe completo cos dispositivos instalados no primeiro ano de vida da rede, 2025.
- Organización dun encontro profesional amplo arredor do SEO en galego para seguir detectando as necesidades neste ámbito en Galicia e xerar dinámicas de traballo e puntos de encontro entre profesionais
- Integración dos datos do galego na doutros idiomas minorizados que están traballando nesta cuestión, o que permitirá ter unha visión máis global e comparada dos datos

ANEXO

Ferramentas TIC en galego



Neste anexo cítanse as principais ferramentas dispoñibles. En caso de que, por novidade ou descoñecemento, haxa algunha omisión, agradecemos a comunicación das mesmas a administracion@dominio.gal para a súa inclusión en sucesivas edicións deste informe.

- **Web e APP da Real Academia Galega**

O dicionario da RAG pódese atopar tanto na propia web da Real Academia Galega coma na súa APP correspondente. A maiores, a RAG na súa web ofrece unha *Guía de nomes galego* que recolle a forma estándar de máis de 1.500 nomes que forman parte da tradición antroponímica galega.

- **O Portal das Palabras**

O Portal das Palabras é un espazo creado en 2013 pola Real Academia Galega e a Fundación Barrié para difundir o léxico e a lexicografía do galego. Desde o seu nacemento pode accederse desde el ao Dicionario da Real Academia Galega e mantén as seccións Palabra do día, Mira que din, Setestrelo, Palabras asinadas e Pregúntalle á Academia. Actualmente conta ademais cunha sección de xogos.

- **Ximnasio Léxico**

Ximnasio Léxico é a aplicación de xogos baseada no Dicionario da Real Academia Galega e que está dispoñible tanto para Android como para iOS. Esta ferramenta está concibida para reforzar o vocabulario normativo.

- Tradutor Gaio da Xunta

O tradutor Gaio é unha ferramenta de código libre, baseada nunha plataforma de código aberto, que permite a tradución de texto plano e tamén a tradución inmediata de páxinas web entre o galego e os idiomas castelán, catalán, inglés, francés e portugués. Conta con APP para Android e iOS.

- Proxecto Nós (nos.gal)

O proxecto Nós (nos.gal), un proxecto da Xunta de Galicia cuxa execución foi encomendada á Universidade de Santiago de Compostela, a través de dúas entidades punteiras de investigación en Intelixencia Artificial e Tecnoloxías da linguaxe: o ILG (Instituto da Lingua Galega) e o CITIUS (Centro Singular de Investigación en Tecnoloxías Intelixentes). Pretende promover a presenza dixital do galego, para que non quede excluído dos novos desenvolvementos TIC coa creación de ferramentas e recursos de alta calidade. Nunha das súas primeiras accións estivo a recollida de voces en galego e a súa validación, para facer un repositorio de modelos válidos orais na lingua de Galicia.

Actualmente conta con varias ferramentas (aínda en fase de desenvolvemento) como un repositorio en GitHub, Zenodo e Hugging Face; e varios demostradores como un tradutor, un TTS (Text-to-Speech) e un ASR (Automatic Speech Recognition).

Recentemente presentaron Carballo, un modelo de linguaxe de gran escala e alta calidade en galego, que permitirá o desenvolvemento de ferramentas tecnolóxicas e sistemas intelixentes creados especificamente para o galego.

- Proxecto Trasno

O Proxecto Trasno é unha comunidade aberta dos proxectos de localización de software libre ao galego. No seu paquete de traducións encóntranse: WordPress, VLC, Gimp, LibreOffice, Firefox e Blender.

- agalega.gal

agalega.gal é unha plataforma de contidos audiovisuais en galego. Está dispoñible tanto en web como en Android e iOS. Conta con máis de 1.400 contidos en aberto e gratuítos e vén acompañada de dúas apps: AGalega Audio, con todos os servizos da Radio Galega integrados, como directos, podcasts e contidos exclusivos; e Xabarín, a app con máis de 1.400 contidos para a cativada.

- **DiGaTIC**

Proxecto desenvolvido dende a Asociación de Enxeñeiros de Telecomunicacións de Galicia (AETG) que pretende pór a disposición da comunidade tecnolóxica un amplo léxico técnico en galego que facilite a utilización e a propagación da lingua galega dentro deste ámbito.

- **O Portal da Lingua Galega**

O Portal da Lingua Galega é a páxina web da Secretaría Xeral da Lingua, un espazo con ferramentas, información e contidos útiles e prácticos relacionados coa lingua galega. Dentro dos recursos para traballar en galego contan con TERGAL (un banco de termos galegos recomendados), manuais de linguaxe xurídica e administrativa, varios paquetes ofimáticos, Galgo (un corrector de galego), etc.

ANEXO

Guías PuntoGal

1
0
0 0
1 1
0 0 1 1
1 1 0 1 0 1 1 0 1 1 1 1
0 0 1 0 0 0 1 0 1 0 0 0 1 0 1
1 0 0 1 1 0 0 1 0 0 1 1 1 0 0 0 1 0 0
0 1 0 1 0 0 1 0 0 1 1 1 0 0 1 1 0 0 1 1
1 0 1 0 0 1 1 0 1 0 0 0 0 0 1 1 0 0 1 0 0 0
0 1 0 1 1 1 0 0 1 0 1 0 0 1 0 0 0 1 1 1 0 1 1 0 1 0 0 1
0 0 0 0 0 0 0 1 1 0 1 0 1 1 0 1 1 0 0 0 0 0 1 0 0 1 0 1 1 0
1 1 1 0 1 0 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 1 0 1 0 1
0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 1 0 1 1 0 1 0 0 0 1 0 1 0 0 0
1 1 0 0 0 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0
0 0 1 0 1 1 0 0 1 0 0 1 0 0 1 1 0 0 0 0 1 1 0 1 1 0 0 1 0 0 1

PuntoGal, desde o seu compromiso coa lingua e con uso do galego na rede e nas TIC, van adicando un esforzo especial a poñer en valor iniciativas que axudan a facer máis doado para o público en xeral ou para grupos sociais ou profesionais máis pechados.

velaquí.gal: ferramenta que ofrecen PuntoGal e Dinahosting para o intercambio de arquivos ou imaxes ata 3GB de tamaño. O sistema permite compartir ficheiros en múltiples formatos así como a subida de maneira simultánea de moitos documentos na mesma solicitude. Toda a interface está en galego.

i.gal: ferramenta que ofrecen PuntoGal, Dinahosting e Agasol para acurtar e personalizar ligazóns e facer un seguimento das mesmas. Toda a interface está en galego.

Guías para adaptar determinadas ferramentas ao uso en galego, que se van publicando e renovando periodicamente na web dominio.gal/guias:

- Dicionario castelán-galego de Rinoceronte Editora.
- Usar IOS en galego.
- Configurar Telegram en galego.
- Configurar Android en galego.
- Configurar Google en galego.
- Usar Firefox en galego en IOS.
- Configurar LibreOffice en galego en Windows.
- Instalar Firefox en galego en MAC.
- Configurar Twitter en galego.
- Configurar LibreOffice en galego en OS X.
- Configurar Microsoft Office en galego en Windows.
- Instalar Firefox en galego en Windows.
- Configurar Windows en galego.
- Usar Firefox en galego en Android.
- Configurar Microsoft Office en galego en Mac.
- Configurar Facebook en galego.
- Instalar e configurar o plugin Opentrad-Puntogal para WordPress.

Do mesmo xeito, e como parte do traballo no ámbito da ciberseguridade, PuntoGal ven de localizar algunhas caixas de ferramentas pensadas para un público profesional. Este traballo faise en colaboración coa Globa Cyber Alliance e están publicadas para PeMEs, xornalistas e traballadores de ONGs. Estas caixas de ferramentas pódense consultar en dominio.gal/ciberseguridade-ga/

Bibliografía e referencias



Aliança per la presència digital del català. (2024). *Informe sobre la pèrdua de visibilitat dels continguts en llengua catalana als resultats de cerca web.* https://aliancadigital.cat/wp-content/uploads/2024/06/informe_posicionament_2024_CAT.pdf

AMP Project. (n.d.). *Facer que unha páxina se poida descubrir en AMP.* <https://amp.dev/documentation/guides-and-tutorials/optimize-and-measure/discovery?referrer=ampproject.org>

Arrizado, F. (7 de xuño de 2023). *O galego afoga no océano Google.* NÓS Diario . <https://www.nosdiario.gal/articulo/lingua/galego-afoga-oceano-google/20190619202455081014.html>

Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA). (n.d.). *Lei de Intelixencia Artificial.* https://amtega.xunta.gal/gl/lei_IA

Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA). (n.d.). *Creación do comité para a Intelixencia Artificial de Galicia.* <https://amtega.xunta.gal/gl/node/9861>

Backlinko (20 de setembro de 2024). *Sobre o comportamento dos usuarios nas páxinas de busca.* <https://backlinko.com/google-user-behavior>

Bello, E. (5 de abril de 2024). *Buscadores de Internet: Qué son y los más utilizados.* IEBS. <https://www.iebschool.com/blog/buscadores-alternativos-a-google-business-tech-tecnologia>

Beus, J. (12 de marzo de 2021). *La proporción de búsquedas móviles es mayor de lo que crees: Qué necesitas saber.* SISTRIX. <https://www.sistrix.es/blog/la-proporcion-de-busquedas-moviles-es-mayor-de-lo-que-crees-que-necesitas-saber>

CITIUS. (18 de xuño de 2024). *Nace na USC un novo 'Carballo', o primeiro gran modelo de linguaxe da historia para o galego.* <https://citius.gal/gl/news/news/carballo-the-first-large-scale-language-model-in-history-for-galician-is-born-at-usc>

Consello de Europa. (2009). *Carta Europea de Linguas Rexionais e Minoritarias.* <https://rm.coe.int/spainecrml6-en/1680b1b32d>

Cyberclick. (2024). *¿Qué es el SEO? Definición, cómo funciona y guía completa de posicionamiento en buscadores.* <https://www.cyberclick.es/que-es/seo>

Cyberclick. (xaneiro de 2024). *¿Qué es el SEM? Cómo funciona el marketing en buscadores.* <https://www.cyberclick.es/sem>

David McCabe. (8 de outubro de 2024). *U.S. Weighs Forcing Google to Break Off Parts of the Company.* The New York Times. <https://www.nytimes.com/2024/10/08/technology/google-search-antitrust.html>

EFE. (26 de xullo de 2024). *OpenAI anuncia un buscador con intelixencia artificial, o seu ataque definitivo contra Google.* elDiario.es. https://www.eldiario.es/tecnologia/openai-anuncia-searchgpt-prototipo-buscador-basado-inteligencia-artificial_1_11550370.html

Europa Press. (2024, 4 de outubro). *Google Lens agora comprende os vídeos e admite preguntas en voz alta.* <https://www.europapress.es/portaltic/internet/noticia-movil-vence-ordenador-trafico-web-20210723130206.html>

Fundació puntCAT. (2023). *Estudi sobre SEO en webs catalanes 2023.* https://domini.cat/wp-content/uploads/2023/04/Estudi_SEO.pdf

Google (n.d.). *Como se xeran automaticamente os resultados.* <https://www.google.com/intl/es/search/howsearchworks/how-search-works/ranking-results>

Google. (2023). *Cómo gestiona la Búsqueda de Google las búsquedas en varios idiomas.* <https://developers.google.com/search/blog/2023/09/multilingual-searches?hl=es>

Google. (n.d.). *Combinar regras de indexación en buscadores.* <https://developers.google.com/search/docs/crawling-indexing/robots-meta-tag?hl=es#combining>

Google. (n.d.). *Como crear e enviar sitemaps.* <https://developers.google.com/search/docs/crawling-indexing/sitemaps/build-sitemap?hl=es>

Google. (n.d.). *Como estruturar os datos para que Google atope o máis relevante das páxinas.* <https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data?hl=es>

Google. (n.d.). *Como funciona a busca bilingüe en Google.* <https://blog.google/intl/en-in/company-news/inside-google/google-for-india-2022-product-announcements/?hl=es-419>

Google. (n.d.). *Como solicitar que Google rastrexe a web de novo.* <https://developers.google.com/search/docs/crawling-indexing/ask-google-to-recrawl?hl=es>

Google. (n.d.). *Como usar o filtro de idioma.* <https://support.google.com/websearch/answer/13485060?hl=es-419>

Google. (n.d.). *Controlar o que se comparte con Google.* <https://developers.google.com/search/docs/crawling-indexing/control-what-you-share?hl=es>

Google. (n.d.). *Documentación esencial sobre a busca.* <https://developers.google.com/search/docs/essentials?hl=es>

Google. (n.d.). *Etiquetas que lle indican a Google o que debe facer cunha páxina.* <https://developers.google.com/search/docs/crawling-indexing/robots-meta-tag?hl=es#noindex>

Google. (n.d.). *Notificar de versións noutras linguas dunha mesma páxina.* <https://developers.google.com/search/docs/specialty/international/localized-versions?hl=es>

Google. (n.d.). *Unificar páxinas duplicadas.* <https://developers.google.com/search/docs/crawling-indexing/consolidate-duplicate-urls?hl=es>

Haßler, G. (2006). *A Fala: normalización tardía e identidade cultural.* Revista de Estudios Extremeños, 62(2), 823-836.

Instituto Galego de Estatística (IGE). (2023). *Enquisa estrutural a fogares: Coñecemento e uso do galego.* https://www.ige.gal/web/mostrar_actividade_estadistica.jsp?codigo=0206004

J. N. (2020, 11 de agosto). *Por que o posicionamento en web dunha empresa é agora máis importante que nunca.* El País. <https://elpais.com/economia/estar-donde-estes/2020-08-11/por-que-el-posicionamiento-en-web-de-una-empresa-es-ahora-mas-importante-que-nunca.html>

Limón, R. (19 de maio de 2024). *La nueva inteligencia artificial cava la tumba de los buscadores y asistentes tradicionales*. El País . <https://elpais.com/tecnologia/2024-05-19/la-nueva-inteligencia-artificial-cava-la-tumba-de-los-buscadores-y-asistentes-tradicionales.html>

Lluvia Digital. (n.d.). *Posicionamiento web: Guía completa sobre SEO*. Lluvia Digital. <https://lluviadigital.com/blog/posicionamiento-web/>

Luzes. (2023, 4 de febreiro). *Cando o GPS vai a “Sangenjo”: unha campaña esixe a Google e Apple que os seus mapas usen a toponimia oficial en galego*. <https://luzes.gal/04/02/2023/luzes-publico/cando-o-gps-vai-a-sangenjo-unha-campana-esixe-a-google-e-a-apple-que-os-seus-mapas-usen-a-toponimia-oficial-en-galego/>

Monteagudo, H. (2024, 25 de setembro). *Novo informe do Consello de Europa*. Real Academia Galega. <https://academia.gal/-/novo-informe-do-consello-de-europa-1>

Mozilla Developer Network. (n.d.). *Que é o DOM e como se usa*. https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model

Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). (2023). *Informe das empresas TIC de Galicia 2023*. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA). https://www.osimga.gal/sites/w_osimga/files/documentos/2023_informe_empresas_0.pdf

Observatorio da Sociedade da Información e a Modernización de Galicia (OSIMGA). (2020). *Informe sobre a sociedade da información nos fogares galegos 2020*. Axencia para a Modernización Tecnolóxica de Galicia (AMTEGA). https://www.osimga.gal/sites/w_osimga/files/documentos/informe_fogares_2020.pdf

Observatorio da Sociedade da Información e a Modernización de Galicia. (2023). *Galicia dixital: O Hipersector TIC en Galicia*. <https://www.osimga.gal/gl/informes/galicia-dixital-hipersector-tic-en-galicia-edicion-2023>

Pérez, J. (25 de xuño de 2024). *Como cambiará Google coa IA, segundo os SEOs que traballan para mellorar os resultados*. <https://elpais.com/tecnologia/2024-06-25/como-cambiara-google-con-la-ia-segun-los-seos-que-trabajan-para-mejorar-los-resultados.html>

Pérez, M. (30 de xullo de 2020). *Rei Zentolo, contra o algoritmo de Google: "as principais ferramentas publicitarias toman o galego como unha errata"*. Praza.gal. <https://praza.gal/ciencia-e-tecnoloxia/rei-zentolo-contra-o-algoritmo-de-google-as-principais-ferramentas-publicitarias-toman-o-galego-como-unha-errata>

Pérez, R. (2023, 9 de marzo). O galego ascende ao posto 37 no barómetro das linguas do mundo. *Nós Diario*. <https://www.nosdiario.gal/articulo/lingua/galego-ascende-ao-posto-37-barometro-das-linguas-do-mundo/20230309123656164012.html>

Precado, A. (19 de decembro de 2023). *Cal é o futuro da lingua galega na Internet...? Vale para as redes sociais ou as buscas en Google?* Galicia Confidencial. <https://www.galiciaconfidencial.com/noticia/248956-cal-futuro-lingua-galega-internet-vale-redes-sociais-ou-buscas-google>

Proxecto Nós. (13 de xullo de 2023). *Galego, catalán, éuscaro e valenciano colaboran dende o PERTE 'Nova Economía da Lingua' para afianzarse na sociedade e na economía dixital*. Proxecto Nós. <https://nos.gal/gl/proxecto-nos/novas/galego-catalan-euscaro-valenciano-colaboran-perte-economia-lingua-afianzarse-sociedade-economia-dixital>

Proxecto Nós. (5 de xuño de 2023). *O proxecto lanza as primeiras ferramentas de interacción persoas e dispositivos en galego*. Proxecto Nós. <https://nos.gal/gl/proxecto-nos/novas/proxecto-lanza-primeiras-ferramentas-interaccion-persoas-dispositivos-galego>

Ramírez, J. M., & García, C. (28 de febreiro de 2022). *Report on the Galician Language*. atlanTTic. https://european-language-equality.eu/wp-content/uploads/2022/03/ELE_Deliverable_D1.15_Language_Report_Galician_.pdf

Rivas Moreno, S. (2021, 1 de xuño). *Como aparecer nos primeiros postos de Google? As novas claves do posicionamento SEO e SEM*. Cinco Días. https://cincodias.elpais.com/cincodias/2021/05/31/extras/1622456193_053046.html

Rubal, M. (25 de setembro de 2018). *Google, unha historia de éxito*. La Vanguardia. <https://www.lavanguardia.com/vida/junior-report/20180920/451931192581/google-stanford-buscador-page-brine-20-anos-historia.html>

Secretaría Xeral de Política Lingüística. (1958-2009). *O uso das linguas oficiais e do galego na UE, desde o Regulamento 1/1958 ata os acordos de 2009.* https://www.lingua.gal/c/document_library/get_file?folderId=1647062&name=DLFE-9291.pdf

StatCounter. (2023). *Comparativa de uso de buscadores entre setembro de 2023 e setembro de 2024.* <https://gs.statcounter.com/search-engine-market-share/desktop/worldwide>

Statista (23 de setembro de 2024). *Cota de mercado global dos sistemas operativos móbiles dende 2009.* <https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/>

Vaquero, J. R. (1997). *Recuperación da información en Internet: motores e outros axentes de busca.* <https://www.iberid.eu/ojs/index.php/scire/article/view/1078/1060>

Wikipedia (n.d.). *Mercadotecnia en motores de busca.* https://es.wikipedia.org/wiki/Mercadotecnia_en_motores_de_b%C3%BAqueda

Zapiplay. (1 de xuño de 2021). *A proporción de buscas móbiles é maior do que pensas: o que precisas saber.* <https://zapiplay.com/la-proporcion-de-busquedas-moviles-es-mayor-de-lo-que-crees-que-necesitas-saber>

•gal



SECRETARÍA
XERAL
DA LINGUA