

Primer va ser el punt. Ara toca posar-hi l'accent

El català i la cultura catalana en temps d'intel·ligència artificial

Joan Abellà — director d'Accent Obert

El català és la llengua d'ús més freqüent per a menys d'un terç de la població de Catalunya. I, malgrat això, en l'entorn digital es troba en una posició comparativament bona —millor que moltes altres llengües en situació similar. Aquesta paradoxa —retrocés social i alhora oportunitat digital— vertebrava la conversa més urgent que podem tenir avui sobre la llengua. I la finestra per actuar-hi és ara.

Mirar les dades de cara

L'Enquesta d'usos lingüístics publicada l'any passat, amb dades del 2023, és clara: el català és la llengua d'ús més freqüent per a menys d'un terç de la població de Catalunya, el 32,6 %. El 2003 era el 46 %. En vint anys hem perdut tretze punts. Entre aquells que s'identifiquen amb el català com a llengua pròpia, el percentatge ha caigut al 30 % —la xifra més baixa dels darrers vint anys.

El castellà, mentrestant, és l'idioma habitual del 46,5 % de la població i s'imposa pràcticament en tots els àmbits de relació quotidiana: amb els amics, amb els veïns, a la feina, als comerços, amb el metge. En tots, excepte un. L'únic espai on el català resisteix com a llengua majoritària és la comunicació amb la Generalitat i amb l'administració local.

Només el 15 % de la població parla exclusivament en català amb els amics. El 31 % ho fa exclusivament en castellà. I més de la meitat dels catalanoparlants reconeix que canvia de llengua per evitar problemes. Per evitar la fricció.

Una dada per damunt de totes: segons el Centre d'Estudis d'Opinió, parlar català ha esdevingut el primer motiu de discriminació a Catalunya. La primera causa. Per davant del color de la pell, de la classe social, del gènere. Una causa de conflicte.

La paradoxa digital

I aquí ve la paradoxa que vertebra el que vull explicar. Enmig d'aquest retrocés social, en l'entorn digital el català es troba en una posició comparativament bona. Millor que moltes altres llengües en situació similar.

Tenim més de 1.400 creadors digitals que publiquen regularment en català a les xarxes; el 2025 van acumular 2.800 milions de visualitzacions. Hi ha 118.000 dominis .cat actius. Tots els sistemes operatius principals estan disponibles en català. El Projecte AINA ha posat el país al mapa de la recerca en IA i llengua. La Viquipèdia en català és la tretzena més important d'Internet. L'Informe Llista, que analitza més de 1.400 projectes digitals en català el 2024 i el 2025, confirma una cosa que cal dir alt: les publicacions van créixer un 31,3 % i les visualitzacions un 71,7 %. El consum creix gairebé el doble de ràpid que la producció. El català digital atrau més del que produeix.

I encara n'hi ha una de millor. Els dotze principals xatbots comercials d'IA generativa —ChatGPT, Gemini, Claude, Grok, Copilot i companyia— ja saben dialogar en català per escrit. Tres d'ells, fins i tot amb veu. Fa tres anys, cap. Avui, tots.

La dada que ho talla tot: el 27 %

Només un de cada cinc dispositius connectats als territoris de parla catalana està configurat amb el català com a idioma preferent. De vint-i-cinc milions de dispositius, cinc. I aquesta és la xifra que les empreses tecnològiques miren quan decideixen si val la pena invertir en suport al català.

A Accent Obert acabem d'elaborar el primer benchmark sistemàtic sobre com tracten el català els quaranta sistemes d'IA més utilitzats del món, comparant-ho amb l'anglès, el castellà i el danès. Sis tipus de tasques. Una pregunta de fons: quina qualitat rep, exactament, un parlant de català quan demana ajuda a un sistema d'IA?

La resposta:

Els grans sistemes d'IA funcionen, de mitjana, un 27 % pitjor en català que en anglès. Vint-i-set per cent menys de qualitat.

Quan un català li demana a ChatGPT que li redacti un escrit, quan li demana a Claude que li resumeixi un document, quan li demana a Gemini que li expliqui un concepte, rep, de mitjana, un servei un 27 % inferior al que rebria si li parlés en anglès.

I la bretxa no és uniforme. Quan demanem a un model que tradueixi un text al català, perd més de quaranta punts respecte de l'anglès. Quaranta. La traducció cap al català —una de les coses que la gent més fa servir— és precisament on els models fallen més. No perquè la llengua sigui pobra: perquè la dieta del model ho és.

Desplaçament per fricció

Imaginem una estudiant de batxillerat. Vol fer un treball sobre la història de la postguerra al seu poble. Obre ChatGPT. Pregunta en català. La resposta arriba. És correcta, però lleugerament menys precisa. La pròxima vegada, sense pensar-hi, pregunta en castellà o en anglès. Perquè vol el millor resultat. Perquè té pressa. Perquè és més còmode.

Això és el desplaçament per fricció. I és, en bona mesura, la manera com es perden les llengües avui. Sense decrets, sense prohibicions: per acumulació de friccions petites que fan que la llengua de menor resistència sempre acabi guanyant. En l'educació, a la feina, a l'administració, en la cultura.

Quan un metge li demana a un assistent clínic com tractar una pacient, i l'assistent respon en un català ple de calcs, que assumeix marcs sanitaris aliens als nostres, les conseqüències deixen de ser abstractes. Una llengua pot sobreviure com a codi de comunicació i morir com a infraestructura cultural. Una llengua mal representada per la IA és una llengua que es buida des de dins.

Sis milions ho fan millor que deu

Arribats aquí, algú podria pensar: és la mida. Som una llengua petita. Què hi farem. Som el 0,47 % de l'internet textual del món; un 0,47 % de la dieta dels grans models.

Però el benchmark inclou una quarta llengua comparable: el danès. Sis milions de parlants. Quatre milions menys que el català.

Doncs bé. El danès, amb menys parlants que nosaltres, obté resultats significativament millors a la IA. La bretxa del danès amb l'anglès és la meitat que la del català.

| *Sis milions de parlants ho fan millor que deu. Per què?*

Perquè Dinamarca és un estat. Perquè ha invertit, durant anys, a construir corpus digitals de qualitat en danès. Perquè ha negociat amb les grans empreses tecnològiques. Perquè ha finançat la participació dels seus investigadors en els grans consorcis de datasets multilingües. No és el nombre de parlants. És la prioritat política. Són les institucions. És la cadena de decisions petites que algú, en algun lloc, ha pres.

Aquí ho haurem de fer a la catalana: recreant un model que coneixem bé, l'aliança entre societat i institucions. La bretxa del català no és un destí: és el resultat de decisions. I per tant es pot revertir prenent decisions diferents.

La finestra és aquesta dècada

La IA generativa no és un truc per escriure textos més de pressa. És una nova capa d'infraestructura. Com en el seu moment ho van ser les carreteres, la xarxa elèctrica o Internet. Una infraestructura que determina qui pot participar plenament en la societat i en quines condicions.

La finestra per intervenir existeix. Però és estreta. Els estàndards d'ús de la IA es fixen ara. Les empreses tecnològiques decideixen ara quines llengües

prioritzen. Els hàbits dels nostres estudiants es formen ara. D'aquí a deu anys, capgirar tot això serà molt més car i molt més difícil.

El català no arriba tard. Arriba just a temps. Però just a temps vol dir ara. I ara vol dir que ens hi hem de posar ahir.

Els ajuntaments són infraestructura de llengua

Probablement, llegint això, més d'un alcalde o alcaldessa pensa: «Tot això està molt bé, però jo soc alcaldessa d'un municipi de vuit mil habitants. Tinc la recollida d'escombraries, el pavelló esportiu, la guarderia, els jubilats i mil coses més. Què hi pinto, jo, en una conversa amb Silicon Valley?».

Hi pinteu molt. Més del que sembla.

Recordem la xifra: un de cada cinc dispositius està configurat en català. Cada ajuntament que configura els seus ordinadors, tauletes i telèfons en català és un senyal estadístic que viatja directament a les decisions que es prenen a l'altre costat de l'Atlàntic o a la Xina. Un sol no es notarà. Però si ho fan dos-cents ajuntaments, o tres-cents, o tots els de Catalunya, el senyal és visible, mesurable, comptabilitzable.

| *L'ajuntament és infraestructura de llengua.*

Tres palanques municipals

1. El senyal digital del vostre ajuntament

Cap endins: configurar els ordinadors, les tauletes i els telèfons municipals en català. Demanar-ho als proveïdors de tecnologia. Que sigui la configuració per defecte de l'organització. És la idea darrere de la campanya Configura, que impulsem amb Òmnium Cultural, Plataforma per la Llengua i Softcatalà. Un acte tècnic petit amb un doble efecte: el senyal estadístic que arriba a les empreses tecnològiques globals, i la normalització interna. Quan un treballador

municipal veu que l'ordinador de l'oficina va en català, i que ningú no ha hagut de demanar permís per fer-ho, la llengua deixa de ser una opció militant i es converteix en l'opció natural. Que és exactament el que ha de ser.

Cap enfora: els llocs web i serveis digitals. Una part molt important de les dades amb què s'entrenen els models d'IA prové del web. Tenir la pàgina en català no n'hi ha prou: cal que estigui ben etiquetada, ben estructurada, tècnicament neta, perquè els sistemes automàtics de rastreig la reconeguin com a contingut en català d'alta qualitat. Des d'Accent Obert estem preparant una guia de bones pràctiques específica per a administracions locals.

I, una cosa importantíssima: caldrà que us convertiu en investigadors d'arxius —els del vostre ajuntament i els del vostre municipi— per ajudar-nos a construir l'immens corpus pendent: el de la cultura catalana. Després hi aprofundim.

2. La contractació pública

És, potser, la palanca més poderosa. Una de les coses que mesura el benchmark és quins models tracten millor el català i quins el tracten pitjor. I la diferència entre uns i altres és abismal. Els millors models —entre el 70 i el 75 % de rendiment en català— són els de tres o quatre empreses concretes. A l'altre extrem, alguns dels models més econòmics del mercat —els que es venen com a solució barata per a institucions amb pressupost ajustat— ofereixen rendiments en català per sota del 40 %. El Phi 4 de Microsoft baixa al 38 % en català, gairebé la meitat que el seu rendiment en anglès. El Llama de Meta es queda al 51 %.

Per què importa, això? Perquè són precisament els models que les administracions —incloses moltes europees— despleguen cada cop més per estalviar costos. Sense saber-ho. Sense que ningú no els ho hagi dit. I el resultat és una administració que ofereix als ciutadans, en català, un servei objectivament inferior al que ofereix en castellà o en anglès.

Cal una auditoria de quines eines d'IA estem fent servir a l'administració, i quin és el seu rendiment real en català. La farem. No és un retret a ningú: és una crida a saber què tenim contractat.

I, a partir d'ara, fer servir el plec de condicions com el que és: un instrument de política lingüística. Quan un ajuntament contracta una plataforma de gestió, un sistema d'atenció ciutadana, una aplicació mòbil o una eina d'IA, hi pot incloure requisits lingüístics. Que el servei funcioni en català. Que l'assistència tècnica s'ofereixi en català. Que les eines d'IA hagin estat avaluades en català i que el proveïdor pugui demostrar que la qualitat hi és equivalent a la d'altres llengües. Sembla un tràmit menor. No ho és. La suma de milers de plecs de condicions és un senyal de mercat poderosíssim. Les empreses tecnològiques s'adapten al que els clients exigeixen. I els ajuntaments són clients institucionals amb poder de compra real. Us facilitarem les eines objectives per fer-ho amb criteri.

3. Els creadors digitals del territori

A molts municipis hi ha joves que fan vídeos, podcasts, contingut per a xarxes socials. Alguns en català, molts no. Per què? Perquè ningú no els ha dit que el català és una opció viable en el món digital. Han vist que en castellà o en anglès les plataformes els donen més visibilitat. Però les dades diuen una altra cosa. Segons l'Informe Llista, els projectes de menys de 5.000 seguidors —que són més del 62 % de tot l'ecosistema digital en català— van doblar les seves visualitzacions en un any. Els creadors emergents en català no van contra corrent: estan creixent.

La Casa de la Creació Digital en Català, que estem impulsant a Barcelona, és un model que pot tenir antenes al territori. Però fins que existeixi, els ajuntaments podeu fer una cosa molt senzilla: visibilitzar i reconèixer els creadors locals que ja ho fan en català. Donar-los suport des dels canals de comunicació municipals. Convidar-los a parlar a l'institut del poble. Cada creador que opta pel català és un node d'una xarxa que creix i que, de retruc, enriqueix el corpus digital del país.

Aquí entrem nosaltres

Tres palanques municipals, tres àmbits d'acció. Però l'ajuntament no pot fer aquesta jugada sol. Algú s'ha d'asseure davant de Google, d'Anthropic, de Mistral, de Common Crawl —el gran repositori global de dades web que alimenta la majoria de models d'IA del món—. Algú ha de fer els benchmarks i mantenir-los actualitzats. Algú ha de coordinar l'ecosistema perquè els esforços se sumin enlloc de duplicar-se. Aquí entrem nosaltres.

Accent Obert és l'hereva de la Fundació .cat, que aquest any compleix vint anys. Vint anys d'ençà que un grup de pioners del català digital es van presentar a l'ICANN —les Nacions Unides d'Internet— a proposar una cosa inèdita: el reconeixement de les cultures en un nou món on, fins llavors, només existien estats. El .cat va ser el primer domini del món vinculat a una cultura, no a un estat.

Ara fa poc més d'un any vam canviar de nom. No per una operació de màrqueting, sinó per ampliar la missió dels webs i els dominis a tot l'àmbit digital. Avui la partida es juga a la IA, als navegadors dels cotxes, als menús de les televisions, a les apps de serveis socials que assisteixen la nostra gent gran. Volem que puguem viure plenament en català en aquest àmbit. En català i des del nostre context cultural. No n'hi ha prou que la IA parli en català: volem que respongui qui era Mercè Rodoreda, o com va ser la Guerra Civil, des del nostre context. Que ens parli de fiscalitat i de salut des del nostre marc.

L'estratègia que hem desplegat per a aquest moment s'articula en un projecte que anomenem IAxCAT (títol provisional), amb tres eixos. **Dades i qualitat lingüística:** assegurar que el català que aprèn la IA sigui ric, divers i culturalment representatiu. **Infraestructura oberta:** coordinar els actors de l'ecosistema per evitar la fragmentació. **Observatori:** el benchmark que us he descrit, que volem actualitzar regularment perquè en cada moment puguem saber en quin estat es troba el català a cada sistema. I, sostenint-ho tot, una capacitat transversal: **la diplomàcia tecnològica.** Parlar amb les grans empreses globals i exigir, amb dades i amb evidència, que el català hi sigui ben representat. Amb la legitimitat de vint anys d'experiència.

Tot això ho farem. Però hi ha una cosa que una organització com la nostra no pot fer com la fan els ajuntaments: arribar a les persones. A les famílies, als comerciants, als centres educatius. Al treballador municipal que cada matí obre l'ordinador i decideix —encara que no ho sàpiga— en quina llengua viu la seva jornada laboral.

La batalla de la llengua en l'era de la IA es lluita en dos fronts alhora. El front digital, el dels models i les infraestructures, és el nostre. El front social, el de la vida quotidiana, és el dels ajuntaments. I un sense l'altre no es guanya.

Primer va ser el punt. Ara toca l'accent

Sovint, quan parlem de la llengua catalana, ho fem des d'un lloc de defensa. De resistència. Com si la nostra feina fos aguantar el cop i preservar allò que tenim. Entenc aquell lloc: vinc del món de la gestió cultural i sé què vol dir treballar per preservar i transmetre allò que som.

Però preservar no és prou. Mai no ho ha estat. I avui menys que mai.

Una llengua es manté pel que permet. Pel que activa. Pel que fa possible. Si el català és útil —si permet treballar millor, aprendre més de pressa, comunicar-se amb precisió, accedir a coneixement de qualitat—, la gent l'usarà. Si no ho és, per molt que ens en sentim orgullosos, quedarà relegat als espais on l'orgull pot pagar el cost de la fricció. I no tothom pot pagar aquell cost.

Per això la batalla digital importa tant. Perquè és la batalla per fer que el català sigui útil. Pràctic. Eficient. Una llengua amb la qual viure plenament en el món que ja és aquí.

I per això la dada del 27 % és tan important. No perquè sigui inevitable: al contrari, perquè sis milions de parlants de danès ja han demostrat que es pot fer millor. La bretxa del català no és un destí. És una decisió. I les decisions es prenen. A Califòrnia, sí. Però també als plens municipals, als plec de condicions, a les configuracions per defecte dels ordinadors d'un ajuntament.

Penseu en els pròxims deu anys. Els infants que ara tenen vuit o deu anys entraran al mercat laboral en un món on la IA serà tan present com avui ho és Internet. Si en aquell món el català és una llengua de primera, podran viure plenament en català sense haver de canviar de llengua per eficiència. Sense la fricció que avui fa que més de la meitat dels catalanoparlants cedeixin. Sense el cost silenciós que paga cada dia qui tria la seva llengua en un entorn que no l'afavoreix.

Aquell és el món que volem construir. Una llengua viva, operativa, capaç d'habitar els espais on avui es decideix el futur.

Fa vint anys, un grup de persones va tenir la gosadia de creure que una cultura sense estat podia tenir un domini propi a Internet. Que l'ICANN reconeixeria que existim. Que el .cat no era una extravagància, sinó una necessitat. Ho van aconseguir contra el que semblava possible. Amb feina, amb aliances, amb visió de llarg recorregut.

Primer va ser el punt. Ara toca posar-hi l'accent. I posar-hi l'accent vol dir entendre que la llengua és política d'infraestructures. Que els drets digitals són drets lingüístics. Que la configuració d'un dispositiu és un acte polític. Que un plec de condicions ben redactat pot ser més poderós que un discurs sencer.

Els ajuntaments teniu el territori. Teniu la confiança de la gent. Teniu la capacitat d'arribar allà on cap algoritme no arriba: a les persones concretes, en els seus contextos reals.

Fem-ho plegats.