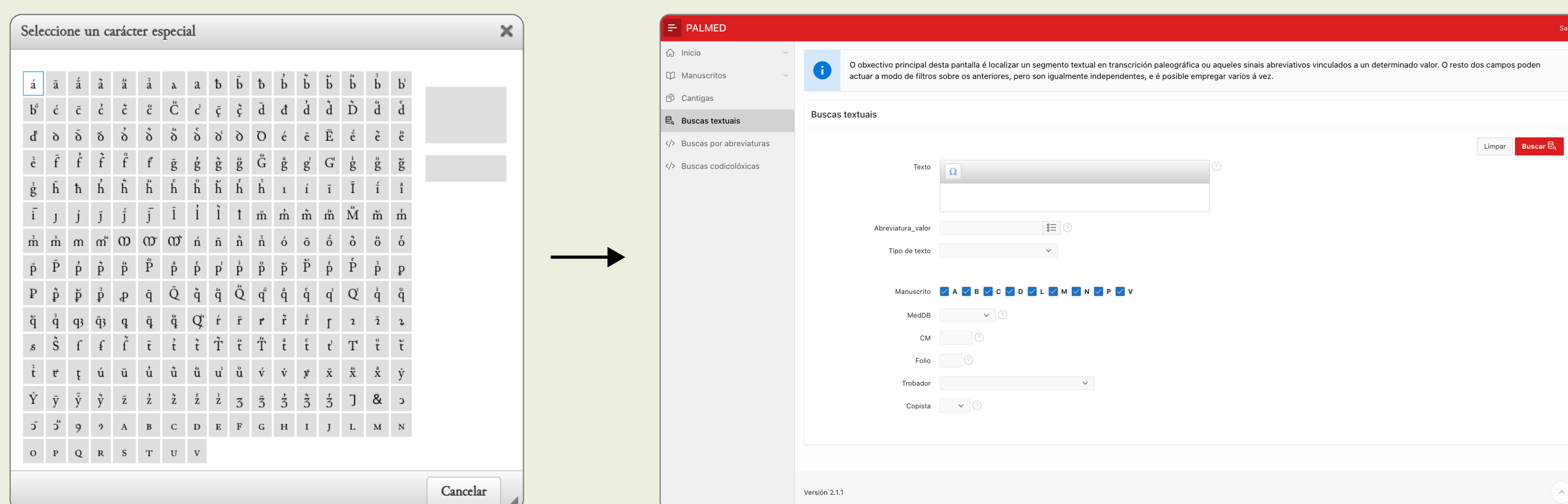


# EDICIÓN PALEOGRÁFICA e VARIACIÓN LINGÜÍSTICA NOS CANCEIONEIROS GALEGO-PORTUGUESES

Pilar Lorenzo Gradín - Carmen de Santiago Gómez (USC - CRPIH)

A Base de datos paleográfica da lírica galego-portuguesa (*PalMed*) ofrece, desde 2020, a posibilidade de visualizar as transcricións conservadoras de todos os testemuños que conteñen as cantigas dos trobadores, á vez que permite realizar buscas combinadas de todos aqueles elementos de interese para o estudo codicolóxico, gráfico, ecdótico e lingüístico de dito corpus ([www.cirp.gal/palmed](http://www.cirp.gal/palmed)).

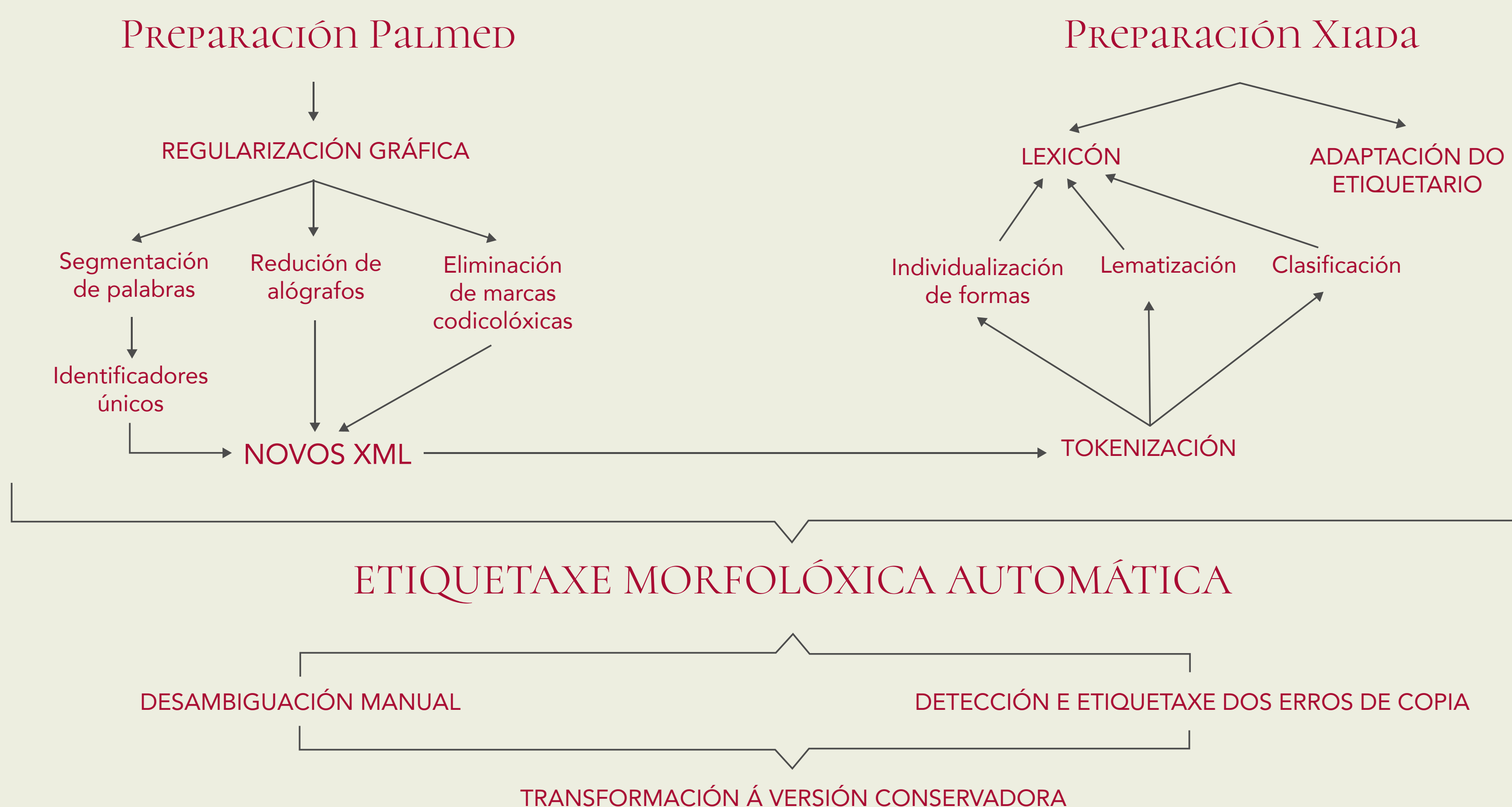
A precisión na etiquetaxe en XML de cada un dos elementos gráficos presentes nos manuscritos serviu para construír unha pantalla de buscas textuais desde a que se pode atopar calquera carácter en todos os contextos a través dun teclado electrónico.



Estas consultas textuais poden afinarse grazas a diversos campos habilitados na mesma pantalla (valor dunha abreviatura, tipo de texto, manuscrito, cantiga, trobador, copista...).

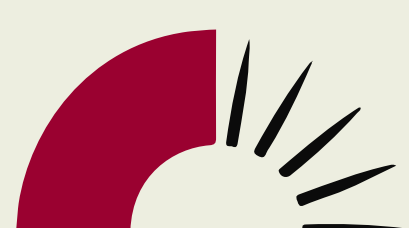
Como complemento de *PalMed*, concibiuse o proxecto *Cancioneiros galego-portugueses. Da paleografía dixital á gramática histórica* (financiado polo Ministerio de Ciencia e Innovación), que ten por obxectivo principal a lematización e a marcaxe morfolóxica de cada termo transcrito, de modo que os usuarios poidan realizar buscas textuais tamén por lemas, formas e categorías gramaticais.

Para levar a cabo este labor adaptouse o sistema de etiquetaxe automática deseñado polo equipo de XIADA (proxecto CORGA, <http://corpus.cirp.gal/xiada/>) ás esixencias do galego medieval.



No etiquetario eliminouse a subclasificación por tipo nas conxuncións, así como as categorías de locución (que se etiquetarán a man mediante marcas que permitirán conservar a información morfolóxica de cada elemento que as integra), de signo de puntuación e de categoría periférica. Así mesmo, engadíronse etiquetas para clasificar os adverbios pronominais, as partículas de reforzo verbal, as palabras noutras linguas e creouse un código numérico para relacionar entre si os lexemas e morfemas das formas verbais analíticas.

Para explotar os datos acadados, deseñárase unha interface de consulta que permita realizar buscas morfolóxicas filtradas por calquera dos criterios que xa existen para *PalMed*, o que facilitará o estudo exhaustivo da variación lingüística nos diversos niveis que configuran a tradición manuscrita galego-portuguesa.



PID2020- 113491GB-I00