



# UNIVERSIDADE DA CORUÑA

## GRAO EN XESTIÓN DIXITAL DE INFORMACIÓN E DOCUMENTACIÓN (2019-2023)

*Análise da produción científica das tres universidades galegas co software estatístico R.*

*Análisis de la producción científica de las tres universidades gallegas con el software estadístico R.*

*Analysis of the scientific production of the three Galician universities with the statistical software R.*

**Autor:** Rubén Fernández Blanco

**Directores:** D. Ángel Manuel González Rueda e D. Javier Tarrío Saavedra.

Ferrol, xuño de 2023



## RESUMO

Nos últimos anos a produción científica adquiriu unha crecente importancia na avaliación de calidade e coa competitividade entre as universidades, sendo así unha referencia importante á hora de obter axudas, prestixios e incluso saber cal é mellor. Dito isto, no presente traballo, temos como obxectivo estudar, analizar e comparar a produción dos grupos de investigación de distintas universidades, empregando técnicas para o tratamento dos datos co paquete Bibliometrix do software estatístico R.

En particular, mostraremos un análise bibliométrico das publicacións do sistema universitario galego, composto pola Universidade da Coruña, Universidade de Santiago de Compostela e a Universidade de Vigo, no período comprendido entre os anos 2012 e 2023, recompilados dende a base datos Web of Science. Entre as técnicas usadas aplicáronse modelos como a lei de Lotka, comparación de revistas, de autores, documentos, redes de colaboración, así como análises de indicadores informétricos.

Os resultados obtidos servirán de axuda para comprender a actividade de cada universidade, así como tamén identificar as fortalezas e as debilidades de cada unha delas. Tendo en conta dito análise debemos situar a cada universidade en un contexto diferente debido a varios factores como pode ser o ano de creación de cada universidade.

Coas ferramentas estatísticas que axudan a realizar este tipo de análises, podemos apreciar a diferenza que existen entre as tres universidades galegas, destacando a Universidade de Santiago de Compostela por ser unha das institucións máis antigas de España con un gran prestixio nacional e internacional, seguidamente das universidades máis novas, UDC e UVigo, que se achegan altamente a un número de produción e de citas comparables á USC.

**Palabras chave:** Bibliometría; WOS; Software R; USC; UDC; UVigo; Biblioshiny.

## RESUMEN

En los últimos años, la producción científica ha adquirido una creciente importancia en la evaluación de la calidad y la competitividad entre las universidades, siendo así una referencia importante a la hora de obtener ayudas, prestigio e incluso saber cuál es la mejor. Dicho esto, en este trabajo tenemos como objetivo estudiar, analizar y comparar la producción de los grupos de investigación de distintas universidades, utilizando técnicas para el tratamiento de los datos con el paquete Bibliometrix del software estadístico R.

En particular, mostraremos un análisis bibliométrico de las publicaciones del sistema universitario gallego, compuesto por la Universidad de A Coruña, la Universidad de Santiago de Compostela y la Universidad de Vigo, en el período comprendido entre los años 2012 y 2023, recopilados desde la base de datos Web of Science. Entre las técnicas utilizadas se aplicaron modelos como la ley de Lotka, comparación de revistas, autores, documentos, redes de colaboración, así como análisis de indicadores informétricos.

Los resultados obtenidos servirán de ayuda para comprender la actividad de cada universidad, así como también identificar las fortalezas y debilidades de cada una de ellas. Teniendo en cuenta dicho análisis, debemos situar a cada universidad en un contexto diferente debido a varios factores, tales como el año de creación de cada una de ellas.

Con las herramientas estadísticas que ayudan a realizar este tipo de análisis, podemos apreciar la gran diferencia que existe entre las tres universidades gallegas, destacando la Universidad de Santiago de Compostela por ser una de las instituciones más antiguas de España con un gran prestigio nacional e internacional, seguidamente de las universidades más nuevas, UDC y UVigo, que se acercan altamente a un número de producción y de citas comparables a la USC.

**Palabras clave:** Bibliometría; WOS; Software R; USC; UDC; UVigo; Biblioshiny.

## ABSTRACT

In recent years, scientific production has gained increasing importance in evaluating the quality and competitiveness among universities, serving as a significant reference when it comes to obtaining funding, prestige, and even determining the best institution. With that said, the objective of this study is to investigate, analyze, and compare the research output of different research groups using data analysis techniques implemented through the Bibliometrix package in the statistical software R.

Specifically, we will present a bibliometric analysis of publications from the Galician university system, which consists of the University of A Coruña, the University of Santiago de Compostela, and the University of Vigo, covering the period from 2012 to 2023. The data was collected from the Web of Science database. Various techniques were employed, including models such as Lotka's law, journal comparisons, author and document analyses, collaboration networks, as well as informetric indicator analysis.

The obtained results will help understand the research activity of each university and identify their respective strengths and weaknesses. Considering this analysis, it is important to contextualize each university differently due to factors such as the year of establishment.

With the statistical tools that help in conducting these types of analyses, we can appreciate the differences that exist among the three Galician universities, with the University of Santiago de Compostela standing out as one of Spain's oldest and most prestigious institutions, renowned both nationally and internationally. Following closely behind are the newer universities, UDC and UVigo, which approach a comparable level of research output and citations to USC.

**Keywords:** Bibliometrics; WOS; Software R; USC; UDC; UVigo; Biblioshiny.

## ÍNDICE

<b>INTRODUCCIÓN</b> .....	<b>7</b>
<b>OBJECTIVOS E HIPÓTESES</b> .....	<b>8</b>
<b>METODOLOXÍA</b> .....	<b>10</b>
<b>RECOMPILACIÓN DOS DATOS</b> .....	<b>10</b>
<b>TRATAMENTO DOS DATOS</b> .....	<b>11</b>
<b>MARCO TEÓRICO</b> .....	<b>13</b>
<b>ESPECIALIDADES MÉTRICAS</b> .....	<b>13</b>
<b>BIBLIOMETRÍA</b> .....	13
<b>CIENCIOMETRÍA</b> .....	14
<b>INFORMETRÍA</b> .....	15
<b>WEBMETRÍA E CIBERMETRÍA</b> .....	15
<b>ALTMETRÍA</b> .....	15
<b>INDICADORES INFORMÉTRICOS</b> .....	<b>17</b>
<b>LEI DE PRICE</b> .....	17
<b>LEI DE LOTKA</b> .....	18
<b>INDICE H</b> .....	19
<b>INDICE G</b> .....	19
<b>OUTROS ÍNDICES</b> .....	20
<b>BASES DE DATOS</b> .....	<b>21</b>
<b>SCOPUS</b> .....	21
<b>WEB OF SCIENCE</b> .....	22
<b>SOFTWARE ESTATÍSTICO RSTUDIO</b> .....	<b>25</b>
<b>SOFTWARE MAPEO CIENTÍFICO (SCIENCE MAPPING)</b> .....	<b>26</b>
<b>BIBLIOSHINY</b> .....	26
<b>VOSVIEWER</b> .....	27
<b>HISTORIA</b> .....	<b>28</b>
<b>UNIVERSIDADE DA CORUÑA</b> .....	28
<b>UNIVERSIDADE DE SANTIAGO DE COMPOSTELA</b> .....	28
<b>UNIVERSIDADE DE VIGO</b> .....	29
<b>SISTEMA UNIVERSITARIO GALEGO (SUG)</b> .....	29

<b>ANÁLISE E INTERPRETACIÓN DOS RESULTADOS .....</b>	<b>30</b>
<b>UNIVERSIDADE DA CORUÑA.....</b>	<b>31</b>
DESCRICIÓN XENERAL .....	31
FONTES .....	33
AUTORES .....	37
DOCUMENTOS .....	39
PALABRAS .....	41
REDE DE COLABORACIÓN .....	42
<b>UNIVERSIDADE DE VIGO.....</b>	<b>43</b>
DESCRICIÓN XENERAL .....	43
FONTES .....	45
AUTORES .....	47
DOCUMENTOS .....	50
PALABRAS .....	52
REDE DE COLABORACIÓN .....	53
<b>UNIVERSIDADE DE SANTIAGO DE COMPOSTELA.....</b>	<b>54</b>
DESCRICIÓN XENERAL .....	54
FONTES .....	56
AUTORES .....	58
DOCUMENTOS .....	60
PALABRAS .....	62
REDE DE COLABORACIÓN .....	63
<b>COMPARACIÓN DAS INSTITUCIÓNIS .....</b>	<b>64</b>
<b>CONCLUSIÓNIS .....</b>	<b>66</b>
<b>BIBLIOGRAFÍA.....</b>	<b>68</b>

## INTRODUCCIÓN

Nos últimos anos a produción científica adquiriu unha crecente importancia na avaliación da calidade e a competitividade entre as universidades, sendo así unha referencia importante para estimar o desempeño das organizacións de educación superior e comparalas. De feito, en España, os indicadores baseados na produción científica son un criterio fundamental no reparto dos fondos por parte dos gobernos autonómicos e estatais.

Esta produción non soamente indica a calidade das universidades, senón que tamén é un indicador clave no desenvolvemento académico e científico dentro das institucións. Xoga un papel fundamental na xeración de coñecementos e na proxección en futuras investigacións nacionais e internacionais. É importante igualmente salientar que, aínda que o último obxectivo dos indicadores de produción científica é a medida do desempeño e da calidade en investigación de institucións, centros, grupos e investigadores, os indicadores de produción científica miden en realidade a magnitude (cantidade), impacto (baseado nas citas) e a visibilidade (nivel de revistas que se publica) desta produción. Merece a pena destacar que para medir a calidade da investigación, sería preciso tamén levar a cabo unha análise cualitativa a través dun comité de expertos.

No presente traballo académico temos como obxectivo principal realizar un análises estatístico da produción científica das universidades do sistema universitario galego, onde se inclúen a Universidade de Santiago de Compostela (USC), a Universidade da Coruña (UDC) e a Universidade de Vigo (UVigo), mediante as técnicas de análises dedicadas as métricas da información.

Estas métricas corresponden a tres tipos de disciplinas (Bibliometría, Cienciometría e Informetría) que teñen por obxectivo determinar as tendencias na produción e comunicación científica, enfocándose na produción de documentos sobre un área determinada do saber a través de identificación de autores, afiliacións, tipos de documentos, palabras chave, etc. Estas posibilidades de medir a información posibilitan a cuantificación de resultados a partir da recollida e do análises dos datos.

Actualmente na sociedade e nas universidades en xeral, podemos detectar que este tipo de análises cada vez xeran máis interese e trazan unha liña importante para saber que



debilidades, ameazas, oportunidades e fortalezas existen en cada universidade. Dito de outro modo, pretendemos examinar e avaliar a cantidade de publicacións científicas xeradas, así como analizar a temática, o impacto das investigacións e outras métricas relevantes que iremos vendo ao longo do traballo.

Para poder levar a cabo este análise, recompilamos datos bibliográficos das tres universidades dende a base de datos Web Of Science (WOS) (Clarivate, 2023) comprendidos entre os anos 2012 e 2023. Estes datos son analizados utilizando técnicas bibliométricas adecuadas que nos permiten extraer conclusións significativas e relevantes sobre a produción científica.

Unha vez temos a base de datos, facemos uso de programas estatísticos como é o caso de R (R Core Team, 2023). A aplicación RStudio co emprego de librerías adecuadas darannos as respostas adecuadas sobre o análise das tres universidades galegas.

Mediante as gráficas sacadas grazas ao programa estatístico R, poderemos obter conclusións precisas de cada universidade ou das tres en xeral. É dicir, poderemos sacar un análise DAFO (Debilidades, Ameazas, Fortalezas e Oportunidades) a partir do cal se podan sacar unhas conclusións de mellora para o futuro investigador das universidades.

## **OBXECTIVOS E HIPÓTESES**

Na realización deste traballo marcamos unha serie de obxectivos xerais e particulares nos que se presente a finalidade de maneira clara. Tamén destacamos as hipóteses que respondan ao plantexamento do problema e á pregunta de investigación.

Os obxectivos son os seguintes:

- Analizar e comparar a cantidade de publicacións científicas xeradas pola USC, UDC e UVigo en un período determinado, neste caso dende 2012 ata a actualidade.
- Avaliar a distribución temática das publicacións científicas das tres universidades galegas e determinar as áreas de investigación máis predominantes en cada unha delas.

- Analizar o impacto das investigacións realizadas ata o momento, tendo en conta os indicadores bibliométricos da nosa base de datos. Observar que factor de impacto posúen as revistas onde se publican ditas investigacións.
- Identificar posibles diferenzas na produtividade científica das tres universidades e determinar si existen grandes diferenzas en termos de cantidade e calidade das investigacións.
- Deixar aberto novas liñas de investigación, é dicir, comparar a comisión interuniversitaria de Galicia coa de outras comunidades autónomas. Dito doutra forma, comparar as universidades galegas con outras nacionais, e incluso internacionais.
- Fomentar o uso de ferramentas estatísticas onde manexar gran cantidade de información de un xeito máis rápido e eficiente como é o caso de R.
- Mellorar na calidade da investigación galega respecto de outras universidades.
- Realizar unha análise DAFO que podería axudar á mellora de investigacións futuras, así como axudar aos investigadores a tomar decisións entre os campos de investigación aínda sen desenvolverse.

Como hipóteses temos as seguintes:

- Hipóteses nula ( $H_0$ ), non existen diferenzas significativas na cantidade de publicacións científicas xeradas polas tres universidades galegas.
- Hipóteses alternativa ( $H_a$ ), existen diferenzas significativas na cantidade de publicacións científicas xeradas polas tres universidades galegas.
- Hipóteses nula ( $H_0$ ), non existen diferenzas significativas en canto a temática das investigacións das tres universidades.
- Hipóteses alternativa ( $H_a$ ), existen diferenzas significativas en canto a temática das investigacións das tres universidades, reflexando áreas de investigación máis destacadas en cada institución.
- Hipóteses nula ( $H_0$ ), non hai diferenzas no impacto das investigacións realizadas, medidos a través de índices bibliométricos.
- Hipóteses alternativa ( $H_a$ ), existen diferenzas no impacto das investigacións realizadas, medidos a través de índices bibliométricos como o factor de impacto das revistas nas que se publican os artigos científicos.

## **METODOLOXÍA**

A necesidade de saber a calidade e a cantidade da produción científica xerada mediante as institucións co fin de observar tendencias xera unha gran controversia na sociedade actual. A gran cantidade de artigos que son expostos á análise para a descrición, comparación e mellora de futuras investigacións, grazas a internet e as novas tecnoloxías, como a constante innovación en intelixencia artificial e a moda do big data, obriga a que esta análise se realice co uso de ferramentas estatísticas.

Neste traballo analizarase a produción científica das tres universidades galegas no período 2012 a 2023, mediante a aplicación de un paquete estatístico do software libre R. Este paquete estatístico Bibliometrix (Aria&Cuccurullo, 2017), creado por Corrado Cuccurullo, serve para a investigación cuantitativa en cienciometría e bibliometría.

Unha vez definida a poboación de interese, que neste caso serían todas as publicacións xeradas da USC, UDC e UVigo, deste traballo e no período que vamos realizar o estudo, procedemos a falar da metodoloxía empregada.

### **RECOMPILACIÓN DOS DATOS**

Para a obtención dos datos é necesario acceder a bases de datos académicas ou repositorios bibliográficos onde atopar toda a produción científica das tres universidades galegas co filtro que comprenda os anos que vamos analizar.

WOS é a base de datos escollida para esta recompilación, xa que aínda que non sexa das bases de datos que máis cantidade de información pode ter, si que é das máis rigorosas á hora de publicar artigos, xa que segue co método de publicación de pares. A colección principal de WOS é core collection, na que se inclúen algún dos índices máis destacados:

- Science Citation Index Expanded (SCIE)
- Social Sciences Citation Index (SSCI)
- Arts & Humanities Citation Index (AHCI)

Para o acceso a WOS debemos ter en conta que se trata de unha base de datos de pago, pero que si temos acceso gratuíto grazas á universidade, de xeito que para poder

entrar debemos autenticarnos co noso correo universitario. Unha vez dentro podemos ver que se poden buscar artigos desde diferentes campos: Título, Autor, Afiliación, Palabras clave, revista, etc. Como na nosa análise vamos a escoller datos das tres universidades galegas, o campo que máis nos interesa buscar é o de afiliacións, de xeito que filtraremos por “Universidade de Coruña”, “Universidade de Santiago de Compostela” e “Universidade de Vigo” marcando como filtro os anos do estudo (2012-2023).

Como ocorre en moitas bases de datos, o manexo de tantos artigos para o seu análise imposibilita a exportación de todas elas en conxunto, de maneira que temos que gardar todos os documentos en un marked list propio de WOS. Unha vez dende aí, debemos exportar os documentos para o posterior análise ca ferramenta estatística, o cal impídese nos que se exporten todas xuntas debendo facer de mil en mil artigos. O ter que facer esta exportación desta forma, dá lugar a que teñamos 120 arquivos en texto plano que se deben xuntar en un único arquivo en formato ZIP.

## TRATAMENTO DOS DATOS

Para o tratamento dos datos vamos a precisar de un software libre chamado RStudio, o cal nos permitirá realizar un análisis exhaustivo a partir da gran cantidade dos datos extraídos no paso anterior.

Para que poidamos realizar o análisis bibliométrico de xeito eficaz, debemos ter cargado o paquete Bibliometrix, o cal nos permite importar e procesar datos dende bases de datos bibliográficas, neste caso WOS, e poder facer análisis de diferentes tipos (análises de coautoría, de citas, de palabras chave, de visualización dos datos).

Unha vez temos instalado RStudio, os pasos para cargar o paquete Bibliometrix e traballar cos datos, son os seguintes:

- Instalar o paquete dende R, mediante a ubicación de Madrid e a función *install.packages(Bibliometrix)*
- Unha vez instalado o paquete, debemos cargar a librería coa función *library(Bibliometrix)*
- Dentro de esta librería executamos *biblioshiny* coa función *biblioshiny()*, o cal nos permitirá abrir a interfaz web para poder traballar cos datos.

- Na interfaz cargamos os datos da base de datos bibliográfica que corresponda, neste caso dende WOS (ver ilustración 1).
- Unha vez se carguen os datos, abrirásenos un widget que nos indicará os tipos de datos que se importaron, así como a súa fiabilidade e o número total de documentos importados.
- Por último, observaremos que datos nos vai interesar analizar e sacaremos as gráficas oportunas que nos permitan comparar as tres universidades.

The screenshot displays the Biblioshiny web interface. On the left is a navigation sidebar with options like 'biblioshiny', 'Data', 'Load Data', 'Gathering Data', 'Filters', 'Overview', 'Sources', 'Authors', 'Documents', 'Clustering', 'Conceptual Structure', 'Intellectual Structure', and 'Social Structure'. The main area shows a table with columns for DOI, AU (Authors), AF (Affiliations), CR (Citations), AB (Abstracts), AR, BA, and BE. Two rows of data are visible, each with a DOI and corresponding author and abstract information. On the right side, there is an 'Import or Load' section with a search bar, a dropdown for 'Please, choose what to do' (set to 'Import raw file(s)'), a 'Database' dropdown (set to 'Web of Science (WoS/WoK)'), a 'Choose a file' section with a 'Browse...' button and a file named 'basedatos.zip', and a 'Start' button. Below this, a 'Conversion results' section shows 'Number of Documents: 55615'. At the bottom right, there is an 'Export collection' section.

DOI	AU	AF	CR	AB	AR	BA	BE
10.1016/j.neurobiolaging.2014.01.017	VILLAR-CHEDA, B.; DOMINGUEZ-MELIJE, A.; VALENZUELA, R.; GRANADO, N.; MORATALLA, R.; LABANDEIRA-GARCIA, J. L.	VILLAR-CHEDA, BEGONA; DOMINGUEZ-MELIJE, ANTONIO; VALENZUELA, RITA; GRANADO, NOELIA; MORATALLA, ROSARIO; LABANDEIRA-GARCIA, JOSE L.	ARES-SANTOS S, 2012, NEUROPSYCHOL DIS, V45, P810, DOI 10.1016/j.nbd.2011.11.005; BABIOR, 2004, CURR OPIN IMMUNOL, V16, P42, DOI 10.1016/j.coi.2003.12.0	IT IS NOT KNOWN WHETHER THE AGING-RELATED DECREASE IN DOPAMINERGIC FUNCTION LEADS TO THE AGING-RELATED HIGHER VULNERABILITY OF DOPAMINERGIC NEURONS AN			
10.1017/S0012217313000875	VERDEJO VM	VERDEJO, VICTOR M.	AKEEL, BILGRAMI, 2006, SELF KNOWLEDGE RESEN; AKEEL, BILGRAMI, 2013, WITGENSTEIN MIND ME, P66; AKEEL, BILGRAMI, 1992, BELIEF MEANING ANONYMOUS, 2008, CANAD	CONCEPT-INDIVIATING POSSESSION CONDITIONS CONSTITUTE ONE OF THE FUNDAMENTAL, ALBERT RATHER UNATTENDED, POLEMICAL TERRITORIES BETWEEN FREEGAN AND ANT			

Ilustración 1. Biblioshiny

## MARCO TEÓRICO

Neste apartado preséntase un marco teórico onde desenvolveremos os enfoques e os fundamentos teóricos que orientan o traballo e o que se coñece ata o momento sobre o tema que se vai estudar. Nel tamén se definen os conceptos e os termos máis relevantes do traballo.

Dividimos o marco teórico en varias partes onde comezaremos falando das especialidades métricas e os seus tipos, indicadores informétricos empregados para o análise, bases de datos de onde sacamos a extracción dos datos, a ferramenta ou entorno de desenvolvemento RStudio co seu software de mapeo científico en análises bibliométricos e a historia das tres universidades galegas.

### ESPECIALIDADES MÉTRICAS

Os estudos métricos da información permítenos avaliar a ciencia e coñecer as particularidades e o desenvolvemento de determinados campos de coñecemento. Cada un destes campos de estudo posúen características particulares e métricas específicas, as cales permiten analizar esta produción investigadora.

Desta forma, vamos a ver as distintas especialidades métricas dende a súa aparición ata a actualidade coa aparición de internet e a transformación da sociedade grazas ás novas tecnoloxías.

### BIBLIOMETRÍA

Cando falamos de bibliometría (Herúbel, 1999) temos que ir ao ano 1934 onde Paul Otlet define o termo como “a parte definida da biblioloxía que se ocupa da medida ou cantidade aplicada aos libros”. A súa proposta fundamentase no feito de que todos os elementos considerados pola biblioloxía son susceptibles de poder ser medidos.

En 1948, Shiyali Ramamrita Ranganathan propón un termo chamado librametrics para designar a “aplicación dos métodos matemáticos e estatísticos á actividade bibliotecaria”. A tradución deste termo ao español defínese como bibliotecometría.

Por outra banda, Alan Pritchard (Pritchard&Wittig, 1981), volve a propoñer o termo bibliometría como “ a aplicación dos métodos matemáticos e estatísticos aos libros

e outros medios de comunicación”. Pritchard é un dos autores máis recoñecidos como creador da bibliometría.

## CIENCIOMETRÍA

En canto a cienciaometría (Sengupta, I. N, 1992) debemos ir ao ano 1963, onde o autor Derek J. De Solla Price, considerado un dos pioneiros deste novo termo, publica o libro “Little Science, Big Science” onde se introduce o concepto de lei de Price para describir a relación entre o tamaño de unha comunidade científica coa súa produtividade.

Un dos exemplos desta lei, foi vista anos antes de que saíse publicado o libro, onde o autor A. Balbi aplicou o análise estatístico a datos do salario de profesores da Universidade de Coimbra, coa cantidade de estudantes por ano e facultades, e a súa analoxía con outros países en 1828. Este estudo realizado por Balbi significou un acontecemento histórico de relevancia e un antecedente claro dos estudos cienciaométricos.

Existen máis autores que foron engadindo e evolucionando este termo. En 1969, Nalimov e Mulcensko definen a cienciaometría como “a aplicación de métodos cuantitativos a investigación sobre o desenvolvemento da ciencia como un proceso informativo”, poñendo de manifesto que o estudo cuantitativo da ciencia pode levarse a cabo mediante métodos estatísticos.

A cienciaometría alcanza un gran desenvolvemento grazas a Eugene Garfield, outro dos autores máis recoñecidos da bibliometría. No ano 1961 funda de Insitute for Scientific Information (ISI) e o creador da base de datos bibliográfica Science Citation Index (SCI), co fin de indexar revistas en diferentes disciplinas científicas. O ISI produciu bases de datos e produtos de información para bibliotecas, universidades, axencias gobernamentais e empresas industriais de todo o mundo. Este termo, agora chamado Web of Science (WOS), revolucionou a forma na que as persoas buscaban información de interese.

Despois de dez anos, no 1971, aparece o Social Science Citation Index, que contén revistas en disciplinas de ciencias sociais. En 1975 crea o Journal Citation Reports (JCR) co obxectivo de medir o impacto das revistas. Eugene Garfiel tamén desenvolve o “factor de impacto” que utiliza para avaliar a relevancia e o prestixio das revistas científicas.

## **INFORMETRÍA**

Unha vez introducidos os termos de bibliometría e cienciometría, debemos ter en conta a informetría. Este termo comeza a empregarse a principios dos anos oitenta, o cal se centra no estudo cuantitativo da produción, difusión e uso da información científica que se ocupa de analizar o impacto da información científica e eficiencia dos sistemas de información. Ademais de ocuparse da medición da información, tamén se ocupa do seu almacenamento e recuperación co apoio das ferramentas informáticas.

## **WEBMETRÍA E CIBERMETRÍA**

Na década dos anos noventa empeza a xurdir un novo fito na nosa sociedade e tamén na historia da bibliometría, a aparición de internet. A partir deste momento, os investigadores déronse conta da alta importancia que tiña poder acceder a información de maneira máis cómoda vía internet. Eugene Garfiel, o creador de ISI, respondeu con rapidez a este auxe informático, expandíndose coa tecnoloxía e lanzando o actual Web of Science en 1997.

Coa aparición das novas tecnoloxías tamén aparecen novos termos relacionados coa bibliometría, pero neste caso orientados á informática. Estes conceptos son os de webmetría e a cibermetría, que se centran no estudo cuantitativo da información na web e nos contornos dixitais.

Estes termos aínda que comparten similitudes, existen algunhas diferenzas, de maneira que, mentres a webmetría (Thelwall, 2009) se ocupa do estudo cuantitativo da web, analizando aspectos como a estrutura da web ou a interconexión entre páxinas, a cibermetría (Torres-Salinas&Moed, 2019), trata dunha disciplina moito máis ampla que abarca a webmetría e o análise da información en contornos máis desenvolto.

## **ALTMETRÍA**

O desenvolvemento que se produce no século XX co nacemento de internet favorece un crecemento na tecnoloxía e unha revolución da maneira de vivir da sociedade. Dita revolución trasládase ás redes sociais na década dos anos 2000 co nacemento das tres redes sociais académicas máis importantes actualmente: Academia.edu, Mendeley e ResearchGate, que fan que haxa un cambio rotundo na maneira da produción académica.



O nacemento destas redes sociais académicas proporciona aos investigadores e aos estudantes novas formas de traballar coa información, é dicir, á hora de encontrar, leer e compartir información relacionada coa súa actividade profesional, ofrecendo así novas ideas e métodos de investigación na comunidade científica.

A raíz desta nova transformación tecnolóxica provocada polas novas ferramentas e maneiras de compartir a información con este tipo de redes sociais, nace en 2010 unha nova métrica, altmetría ou métricas alternativas, que se empregan para medir a actividade, visibilidade e o impacto social e científico das producións académicas nos contornos web.

A primeira definición de este novo termo nace no 2010 en Altmetrics: A Manifesto (Priem et al., 2010) como “creación e estudo de novas métricas baseadas nas redes sociais para o análise e o estudo da comunidade científica”. Ao longo dos anos este termo foi evolucionando, Khodiyar (Khodiyar et al., 2014), pon como estudo da altmetría (Roemer&Borchardt, 2015) de resultados científicos que son tratados de algunha forma na web. En 2015 defínese como disciplina e ao ano seguinte a National Information Standard Organization (NISO) inclúe a definición a calquera métrica que se xere na web e que se refira a actividade académica.

## INDICADORES INFORMÉTRICOS

Nun estudo bibliométrico os indicadores informétricos configúranse como características que explican o comportamento da ciencia e da tecnoloxía. Defínense como uns parámetros que se empregan no proceso avaliativo das investigacións científicas. Permiten analizar o tamaño, crecemento e a distribución da bibliografía, así como identificar os axentes implicados na produción (autores, grupos de investigación, revistas, etc.), pero tamén nos permiten avaliar e obter informacións sobre o crecemento ou a diminución de un campo científico.

Os resultados obtidos grazas a estes indicadores informétricos van a adquirir importancia en decisións futuras por parte das institucións estudadas, neste caso do sistema universitario galego, de maneira que axuden a cambiar políticas de xestión e de decisións, co mesmo fin de obter a maior calidade posible en futuras investigacións.

En este apartado veremos algúns dos indicadores informétricos máis empregados no traballo co fin de obter as análises correspondentes de cada universidade estudada coa axuda da ferramenta de análise.

### LEI DE PRICE

En 1844 Engels propón a “lei do desenvolvemento acelerado da ciencia: a ciencia progresa proporcionalmente a masa de coñecemento acumulada pola xeración precedente”.

En 1956, como comentamos antes, un dos autores que máis atribúe o termo *cienciometría* Derek J. De Solla Price (Price, 1965), propón a coñecida “lei de crecemento exponencial da ciencia”, que para xustificala recorre a múltiples observacións cuantitativas, referidas ao crecemento de revistas científicas, autores, institucións, etc.

Esta lei do crecemento exponencial da ciencia ou tamén chamada lei de Price, divídese en dous principios: lei de crecemento exponencial e lei de crecemento loxístico. Falaremos da primeira lei a cal afecta á investigación:

1ª Lei de Price, lei de crecemento exponencial, pódese enunciar como:

- A ciencia aumenta a interese composto, multiplicándose por unha cantidade determinada en períodos iguais de tempo.
- A taxa de crecemento é proporcional ao tamaño da poboación ou magnitude total.
- Canto máis grande é a ciencia, máis rápido aumenta.

A fórmula desta lei sería a seguinte (ver ilustración 2):

$$N = N_0 e^{bt}$$

*Ilustración 2. Lei de Price*

Onde N é o número de publicacións en un período determinado, t é o tempo e b é un parámetro relacionado coa velocidade de crecemento.

### LEI DE LOTKA

A lei de Lotka (Lotka, 1926), proposta en 1926 por Alfred J. Lotka, é a lei da distribución dos autores segundo a súa produtividade. Esta lei establece que o número de produtores,  $A_n$ , que publican “n” traballos sobre unha materia é inversamente proporcional a “n” ó cadrado. A lei tamén implica que moi poucos autores publican a maioría dos traballos, para unha produtividade alta ( $n >$ ), o número de autores,  $A_n$ , é baixo, xa que ambas variables son inversas.

A fórmula da lei de lotka será a seguinte (ver ilustración 3):

$$A_n = \frac{A_1}{n^2} \equiv A_n = \frac{A_1}{n^m}$$

*Ilustración 3. Lei de Lotka*

Onde “ $A_n$ ” é o número de autores con n firmas. “ $A_1$ ” é o número de autores con 1 firma e “ $n^2$ ” é o número de firmas ao cadrado.

Esta lei de Lotka, xunto coa lei de Price, permítenos obter unha gran información acerca da produción científica das tres universidades galegas. No apartado de análises poderemos observar as comparacións destas institucións.

### **INDICE H**

Este indicador creado por Hirsch (Hirsch, 2005) en 2005 é un indicador bibliométrico empregado para medir a produtividade e o impacto da investigación de un autor. De xeito que se define como número “H” aquel que coincide co maior número de publicacións de ese autor que recibiron polo menos “H” citas.

Un índice “H” alto indícanos que o autor ten unha boa cantidade de publicacións altamente citadas, indicándonos así unha boa contribución científica. En caso de que o índice sexa baixo sería todo o contrario.

Co paso dos anos este índice variará en función do investigador onde primará a súa carreira investigadora ou a colaboración con outros autores. É dicir, pode haber anos que non existan moitas colaboracións co investigador, o cal pode afectar na súa produción científica de maneira que este índice se vexa afectado en menor ou en maior grao.

### **INDICE G**

O índice “G” proposto por Leo Egghe (Egghe, 2006) no ano 2006, igual que o índice “H”, cuantifica a produtividade bibliométrica baseada nas publicacións dos investigadores. Calculase ordenando as publicacións de un investigador polo número de citas recibidas en orde ascendente, numerando a posición e obtendo así novas columnas de datos.

Un autor ten un índice “G” cando, considerando os “G” artigos máis citados do investigador, a cantidade de citas acumulados por estes artigos “G” é superior a “G” ao cadrado.

O índice “G” en comparación co índice “H”, é maior e mais variable que este último, o que fai que se teñan que observar e analizar os dous índices en conxunto.

## OUTROS ÍNDICES

Unha vez vistos os índices máis frecuentes á hora de facer o análise bibliométrico das tres universidades galegas, debemos ter en conta os indicadores máis tradicionais como poden ser os índices de produtividade científica, indicadores baseados en citas, indicadores propios das revistas (JCI, JCR), indicadores de uso, indicadores de reputación, indicadores de colaboración, etc.

Estes outros tipos de indicadores axúdannos tamén a sacar gran parte das gráficas ca súa análise correspondente, os cales non podemos pasar por desapercibido.

## **BASES DE DATOS**

No análise bibliométrico debemos ter en conta a necesidade de sacar a información dalgún repositorio bibliográfico ou de bases de datos académicas que nos proporcionen a información que queremos analizar das tres universidades nos anos comprendidos entre 2012 e 2023.

Actualmente as principais fontes de información bibliométrica son Web of Science e Scopus, de acceso de pago, nos que se encontran os maiores repositorios das principais publicacións científicas. Estas bases de datos tamén teñen os seus indicadores de calidade, e cabe destacar que en España se empregan os artigos indexados no Journal Citation Report de WOS para avaliar a promoción da carreira docente universitaria.

Neste traballo optouse por escoller Web of Science, xa que aínda que por números teña máis publicacións Scopus, as publicacións nesta primeira base de datos requiren de altos requisitos de calidade. Para os investigadores do grao en documentación que empreguen estas bases de datos, necesitarán saber realizar consultas de maneira eficaz para poder elaborar análises estatísticos personalizados que describan, avalíen e comparen institucións, autores, temas de investigación en crecemento, etc. É importante realizar este tipo de análises xa que permitirán cambiar políticas e mellorar en xeral a calidade das publicacións en todos os ámbitos universitarios.

Neste apartado explicaremos as dúas bases de datos, as cales permitirían facer esta análise facendo máis énfases na base de datos de Web of Science, a escollida neste análise bibliométrico.

### **SCOPUS**

Unha das bases de datos máis importantes a nivel científico é Scopus (Elsevier, 2023), fundada pola editorial Elsevier no ano 2004, na que se proporciona acceso a investigacións revisadas por pares. Ofrece información sobre publicacións seriadas, conferencias e patentes, ademais da información bibliográfica.

Scopus conta con ferramentas bibliométricas para o rastrexo, a análise e o comportamento da actividade investigadora, converténdose no principal competidor de WOS. Tamén permite realizar consultas e refinar a búsqueda por resumos, palabras chave,

citas, referencias... así como exportar os datos a xestores bibliográficos como RefWorks. Un dos índices de impacto principal nesta base de datos é SJR (Scimago Journal Rank), seguido de CiteScore, SNIP (Source Normalized Impact per Paper) e Scopus content coverage.

## WEB OF SCIENCE

A base de datos empregada para a realización deste traballo é Web of Science (WOS), a cal é unha das fontes fundamentais na bibliometría, compostas de repositorios multidisciplinares que inclúen información de alta calidade.

A colección principal de WOS é core collection na que se inclúen:

- Science Citation Index Expanded (SCIE)
- Social Science Citation Index (SSCI)
- Arts & Humanities Citation Index (AHCI)

Ao igual que Scopus é de pago. Encontrase dispoñible en todas as universidades, nas que se permiten realizar consultas, incluíndo tamén utilidades como é a visualización de xeito estatístico na propia base de datos. Tamén podemos incluír un dos rankings máis importantes en revistas científicas, o chamado Journal Citation Report (JCR), no que se clasifican as revistas de cada área seguindo unha serie de índices bibliométricos entre os que destaca o índice de factor de impacto.

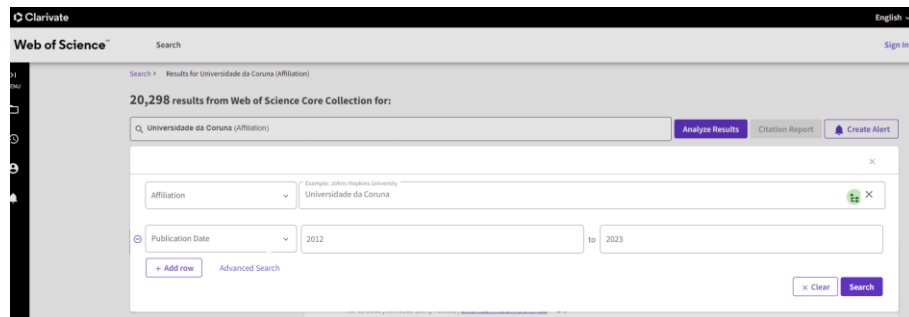
A continuación explicase cal é o proceso para acceder a esta base de datos o cal é de onde se sacan os datos para a súa posterior análise na ferramenta estatística. É importante destacar que para facer uso de WOS debemos ser alumnos ou docentes da universidade na que estudemos, neste caso podemos acceder grazas a Universidade da Coruña.

Primeiramente para acceder a páxina de WOS, debemos acceder á Fundación Española para a Ciencia e a Tecnoloxía (FECYT) quen centraliza e xestiona o acceso as universidades. Destacamos que para acceder existen dúas formas dependendo de onde nos encontremos, é dicir, se nos encontramos na propia universidade debemos estar conectados a rede wifi do campus para que se nos permita o seu uso. Se nos encontramos

fora da universidade debemos acceder aos servizos bibliotecarios da universidade e autenticarse dende fora da UDC senón non poderemos acceder.

O enlace para acceder a WOS dende FECYT é o seguinte: <http://woscc.fecyt.es/>

Unha vez nos encontramos dentro de WOS apareceranos o buscador onde despregaremos e buscaremos por “Affiliations” e “Add date range” onde iremos universidade por universidade engadindo o rango de anos que imos a buscar para o análise bibliométrico (ver ilustración 4).



*Ilustración 4. Web of Science*

Como a nosa queri ou consulta trata das tres universidades galegas e comprendidos entre os anos 2012 e 2023, a cantidade de datos a exportar é moi grande, o cal WOS solo nos permite exportar os datos de 500 a 500.

Para poder exportar e manter os datos, engadímolos a un marked list (lista de favoritos) o cal nos permitía ter gardada a queri cada vez que entramos en WOS, facilitándonos non ter que buscar o mesmo cada día que entrabas a exportar ditos datos.

Unha vez temos cada universidade en un marked list debemos seleccionar o botón de “Export” onde se nos despregaran unha serie de formatos, o cal sabendo que ferramenta e que paquete vamos a empregar para analizar os datos, debemos seleccionar un formato ou outro. No noso caso o paquete empregado vai ser bibliometrix de modo que deberemos seleccionar o formato “Plain Text”.

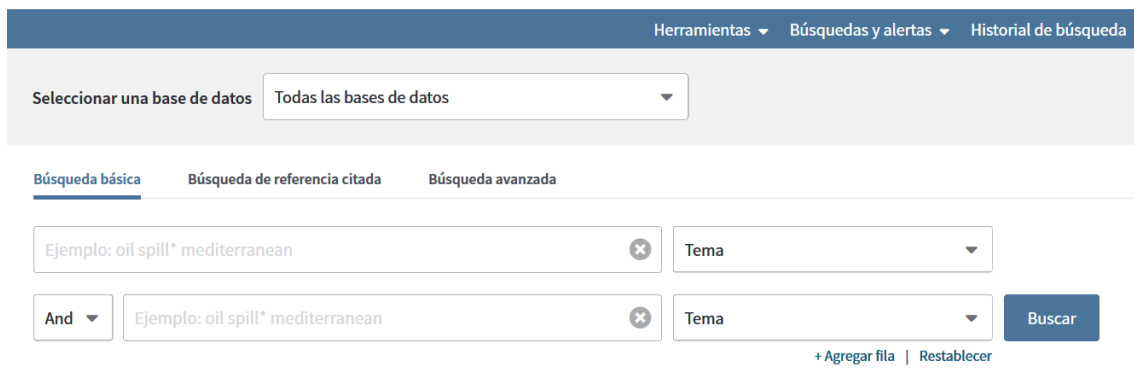
Como dixemos anteriormente, temos que exportar os datos de 500 en 500, e como temos unha gran cantidade de datos: UVigo (19183 publicacións), UDC (16879) e USC (29275), a cantidade de arquivos en formato de texto plano é de uns 150 arquivos, que debemos xuntar un a un para a creación da nosa base de datos definitiva que é a que



vamos a empregar para o análises en RStudio. Este paso é longo, porque debemos ir con cautela seleccionando os datos, evitando repeticións e esperar a velocidade da descarga dos datos, que algúns casos a demora era alta, da mesma forma a unión de todos os arquivos tamén require de un tempo a ter en conta.

Unha vez exportados os datos podemos facer uso das ferramentas analíticas que ofrece WOS e comparalo co análises bibliométrico realizado na nosa ferramenta estatística, se queremos ver que os datos se exportaron ben e ambas coinciden (ver ilustración 5).

## Web of Science



The screenshot displays the Web of Science search interface. At the top, there is a navigation bar with 'Herramientas', 'Búsquedas y alertas', and 'Historial de búsqueda'. Below this, a dropdown menu allows selecting a database, currently set to 'Todas las bases de datos'. Three search modes are available: 'Búsqueda básica' (selected), 'Búsqueda de referencia citada', and 'Búsqueda avanzada'. The search area contains two input fields, both with the example query 'Ejemplo: oil spill\* mediterranean'. The second field is preceded by an 'And' operator. To the right of each field is a 'Tema' dropdown menu. A 'Buscar' button is located to the right of the second field. Below the search area, there are links for '+ Agregar fila' and 'Restablecer'.

*Ilustración 5. Web of Science*

## SOFTWARE ESTATÍSTICO RSTUDIO

No estudo estatístico das tres universidades galegas é imprescindible contar cun programa estatístico se queremos traballar con unha gran cantidade de información. Deste xeito, permítesenos analizar gran cantidade de variables e coñecer as diferentes situacións de cada universidade, así como tamén as súas similitudes.

Este programa estatístico vai ser RStudio, un entorno de desenvolvemento libre, creado no ano 2011, que se converteu en unha das ferramentas máis populares empregadas para data science ou analistas. RStudio posibilita traballar co linguaxe programación R, de unha maneira máis visual sen necesidade de coñecer dito linguaxe. Para poder instalar RStudio debemos ter instalado primeiramente R, o cal vai ser a “antiga” aplicación onde se encontrará o linguaxe de R. Da mesma forma, se queremos instalar algún tipo de paquete debemos instalalo primeiramente na aplicación de R para que posteriormente apareza en RStudio (que é o que nos vai pasar co paquete de biblioshiny).

RStudio vai permitir aos científicos dos datos (data science) e aos analistas crear visualizacións interactivas onde poder ver a información dos datos dunha forma moito máis asumible. Tamén permite a creación de informes con RMarkdown onde poderemos facer páxinas web, xeración de documentos e presentacións en diferentes formatos de unha maneira automática.

No presente traballo faremos uso de un software de mapeo, o cal debemos instalar en R, un paquete chamado bibliometrix. Para executalo en RStudio debemos chamar á librería co comando `library(Bibliometrix)` e despois chamar a biblioshiny co comando `biblioshiny()`. Unha vez executado abriráenos unha interface de forma local no noso ordenador ao que soamente nos teremos acceso na que analizaremos diferentes puntos das investigacións anteriormente exportadas.

## **SOFTWARE MAPEO CIENTÍFICO (SCIENCE MAPPING)**

A técnica de Science Mapping ten como obxectivo mostrar da forma máis intuitiva posible as relacións entre os investigadores, revistas, universidades, liñas de investigación... a través de diferentes gráficos facilitando así o análise bibliométrico. Como no noso traballo empregamos un entorno de desenvolvemento en R, debemos ter seleccionado un software que corra en R. Este software vai a ser biblioshiny procedente da librería bibliometrix.

Bibliometrix é un paquete de R para o análise bibliométrico e cienciaométrico creado por A. Cucurullo onde se incorporan diferentes técnicas de análise descritivo dos datos, ademais de técnicas multivariantes que permiten detectar e describir grupos e relacións entre investigadores, universidades ou revistas. Para non ter que facer uso do código empregaremos biblioshiny.

### **BIBLIOSHINY**

Biblioshiny é unha variante de bibliometrix orientado a usuarios que non están habituados a manexar código en R. De maneira que é unha interface gráfica integrada nunha páxina web local que ten como base biblioshiny (ver ilustración 6). Opera con datos procedentes de diferentes bases de datos como pode ser WOS ou Scopus, no noso caso traballaremos coa nosa base de datos extraída de WOS.

Componse de varios menús onde podemos ver diferentes tipos de gráficos, os cales cada menú despregarase con submenús onde vemos máis gráficos:

- 1) Visión xeneral
- 2) Fontes
- 3) Autores
- 4) Documentos
- 5) Estruturas conceptuais
- 6) Estruturas intelectual
- 7) Estruturas sociais



## **HISTORIA**

O análise bibliométrico onde examinaremos as tres universidades galegas que conforman o Sistema Universitario Galego (SUG), vaines a dar uns resultados que dependendo da historia de cada universidade variarán. Isto debese ao tipo de facultades que posúe cada universidade, na que existirán temas de investigación que se traballará en cada unha delas, así como a cantidade de docentes e de publicacións que varían potencialmente ano tras ano.

De este modo, sendo as tres institucións representativas de cada cidade galega posúen unha historia e circunstancias diferentes. En este apartado trataremos de delimitar cada universidade no seu contexto xeográfico e histórico.

### **UNIVERSIDADE DA CORUÑA**

A Universidade da Coruña (Universidade da Coruña [UDC], 2023) fundada no ano 1989 e ubicada na cidade da Coruña, é a universidade xunto coa de Vigo máis xoven que compoñen o sistema universitario galego. A creación desta universidade debese a alta demanda na sociedade galega de contar con unha institución sen ter que desprazarse a outras cidades.

A UDC comeza a súas actividades académicas no curso 1990-1991 coa apertura das facultades de Dereito e Ciencias Económicas e Empresariais. Ao longo dos anos as facultades foron medrando e engadíndose novos graos como Enxeñaría Informática ou Xestión Dixital de Información e Documentación, entre moitos outros.

### **UNIVERSIDADE DE SANTIAGO DE COMPOSTELA**

A Universidade de Santiago de Compostela (Universidade de Santiago de Compostela [USC], 2023) é unha das universidades máis antigas de España fundada en 1495 polo arcebispo Alonso de Fonseca centrándose nos primeiros anos en disciplinas de teoloxía, filosofía e dereito canónico.

Ao longo dos anos esta universidade foise transformando alcanzando un gran prestixio e converténdose en un centro principal de estudo en Europa. Esta transformación

veu ligada a titulacións implementadas como ciencias da saúde e ciencias sociais e xurídicas.

Esta universidade é das que máis destaca polo prestixio académico, a traxectoria histórica de ser unha das universidades máis antigas e da súa alta actividade investigadora.

### **UNIVERSIDADE DE VIGO**

A Universidade de Vigo (Universidade de Vigo [UVigo], 2023), xunto coa de Coruña, fúndanse no mesmo ano coa lei 11/1989, de 20 xullo, por ordenación do sistema universitario galego. Ambas universidades teñen en común que foron creadas en ambientes industriais e portuarias, non como en Santiago, que xa era creada polo clero séculos atrás.

En canto a área de estudo, é parecida a UDC onde destacan enxeñarías, tecnoloxía e ciencias sociais e xurídicas.

### **SISTEMA UNIVERSITARIO GALEGO (SUG)**

O sistema universitario galego (Xunta de Galicia, 2023) está composto polas tres universidades: Universidade de A Coruña (UDC), Universidade de Santiago de Compostela (USC) e a Universidade de Vigo (UVigo). Aquí as tres universidades traballan en conxunto para ofrecer unha ampla oferta en graos e mestres, e fomentar a investigación e a transferencia de información científica na comunidade autónoma de Galicia.

O SUG crease no ano 1989 descentralizando así o sistema universitario desa época xa que soamente existía a Universidade de Santiago de Compostela. No ano 1990 aprobouse a creación da UDC e da Uvigo. Unha vez creado ordénanse os campus que formarían parte das tres universidades, Ourense e Pontevedra quedan adscritos á Uvigo, Ferrol queda adscrito á UDC e por último, o campus de Lugo queda adscrito á USC.

## ANÁLISE E INTERPRETACIÓN DOS RESULTADOS

Unha vez temos a nosa base de datos, conxunto de todos os textos planos extraídos de WOS dende o ano 2012 ata a actualidade, procedemos a realizar o estudo estatístico descritivo de cada universidade, mediante a axuda de RStudio e da librería biblioshiny a través da súa interface web, a cal nos permite visualizar cada tipo de gráfica e descargala en imaxes.

Como tiñamos unha base de datos en bruto coas tres universidades xuntas, os resultados obtidos non eran moi específicos á hora de ver que predominaban os autores, publicacións e revistas da universidade da que máis datos temos, que era a USC.

De xeito que unha vez cargados os datos en biblioshiny procedemos a descargalos en formato csv, para posteriormente en Excel realizar o filtro e dividir a base de datos xeral en tres bases de datos distintas, onde en cada unha se atopa cada universidade. Desta forma, volvemos a realizar o análise por cada universidade podendo ver resultados desglosados por cada unha delas.

Os datos resultantes nas tres bases de datos son os seguintes:

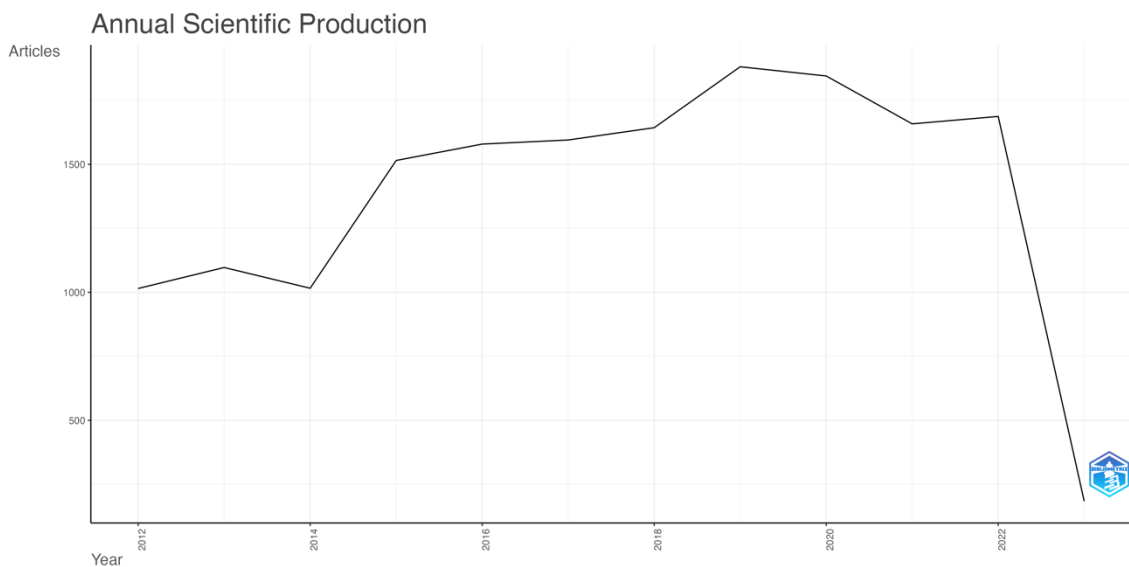
- Universidade da Coruña: 16879 documentos.
- Universidade de Santiago de Compostela: 29275 documentos.
- Universidade de Vigo: 19183 documentos.

O resultado total de documentos nas tres universidades galegas fan un total de 65337 documentos xerados dende o ano 2012 ata 2023. Dito isto procedemos a analizar cada universidade e vendo as distintas gráficas resultantes do devandito análise bibliométrico. En cada universidade engadiremos algunhas que outras gráficas diferentes.

## UNIVERSIDADE DA CORUÑA

### DESCRIPCIÓN XENERAL

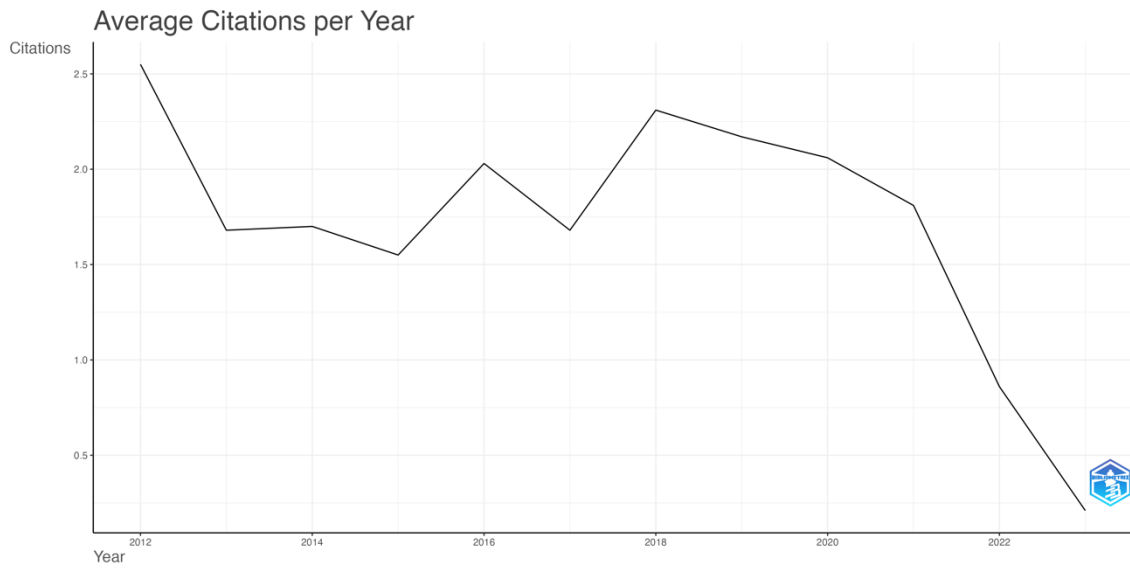
Na seguinte gráfica (ver ilustración 8) podemos ver a **producción científica anual** da UDC, na cal podemos apreciar que o número de documentos varía no ano 2014 pero a partires dese ano en diante sube considerablemente esta produción chegando case a 2000 artigos no ano 2019. Tamén se aprecia que despois do ano 2019 as producións baixan podendo ser a causa a un factor sanitario a nivel mundial como é o COVID-19, aínda que tampouco se observa unha baixada tan forte. Por último, no ano 2023 as producións caen porque nos encontramos actualmente en este período, o que fai que en WOS aínda non se teñan moitas publicacións actuais.



*Ilustración 8. Producción científica anual*

A seguinte gráfica (ver ilustración 9) indícanos as **citas promedias por ano**, onde se observa que no ano que máis citas temos é no ano onde comezamos a realizar este análises bibliométrico, no 2012. Vemos que un dos peores anos en citas foi o 2015 con 1.6 de citas, e a partires de aí temos unhas subidas e baixadas considerables





*Ilustración 9. Citas promedias por ano*

## FONTES

Na seguinte gráfica (ver ilustración 10) podemos observar as **fontes máis relevantes** onde aparecen as 10 revistas que máis número de documentos posúen, encabezando a revista “Annals of the rheumatic diseases”. De feito, obsérvase que as revistas onde máis publican os investigadores da UDC son as relacionadas con medicina, especificamente con reumatoloxía e cardioloxía. Isto está relacionado con dous grupos moi fortes dentro da universidade, o GRINCAR, de cardioloxía, dirixido por Javier Muñiz e María Generosa Crespo Leiro, e o de reumatoloxía, dirixido por Francisco Blanco.

Tamén é importante salientar que nesta listaxe están revistas full open access (só publican en open access e é o investigador quen se ocupa de pagar o custo de publicación, polo que están máis orientadas ao negocio, publican máis artigos), mega revistas (de moi ampla temática, diferentes ciencias, gran número de artigos publicados), que publican un grande número de artigos, como son Plos One e Scientific Reports.

Tamén hai que destacar a incidencia doutro grupo de revistas full open access, pertencentes á editorial china (aínda que radicada en Basilea) MDPI. Son moito máis rápidas na publicación que as demais, sacan un número grandísimo de números especiais, publican moitísimos artigos a través dunha política agresiva de captación. Estas revistas son Sustainability en Sensors, sétima e décima no listaxe respectivamente. Que estean estas e non outras ten relación a que están enfocadas as ciencias ambientais e sostenibilidade, a primeira, orientada á enxeñaría, tecnoloxía e monitorización, a segunda, ámbitos en auge na actual Industria 5.0 (dixitalización e sostenibilidade).

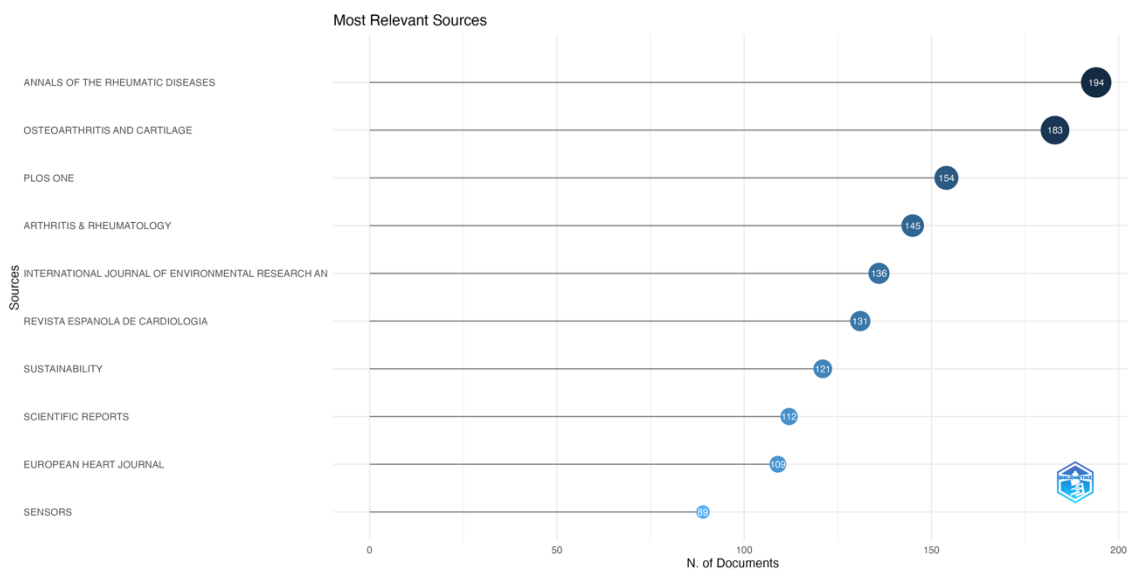


Ilustración 10. Fontes máis relevantes

En canto as **fontes locais máis citadas** (ver ilustración 11) destaca a revista “Plos One” con un total de 4077 citas locais. Hai unha relación entre citas e prestixio das revistas. Pódese ver en que aquí figuran “Nature” e “Science” (que non son as revistas onde máis publicamos pero que publicar un só artigo nelas garante moitas citas). Tamén existen moitas citas en revistas médicas.

As revistas de química tamén soen ter moitas citas, máis que moitas outras ciencias. Aquí aparece “Journal of American Chemical Society” e “Inorganic Chemistry”. Isto tamén está relacionado coa forza do centro de investigación CICA (Centro Interdisciplinar de Química e Bioloxía [CICA], 2023), da UDC. Tamén se pode ver a actividade do CITIC (Centro de Investigación en Tecnoloxías de la Información y Tecnoloxías [CITIC], 2023) a través da revista “Lecture Notes in Computational Science”.

Outras que teñen moitas citas son as revistas que publican moitos artigos como “Plos One” e “Science of the Total Environment”, aparte de focalizarse en ámbitos de profusa citación (bioloxía, química e medicina a primeira, ciencias ambientais a segunda).

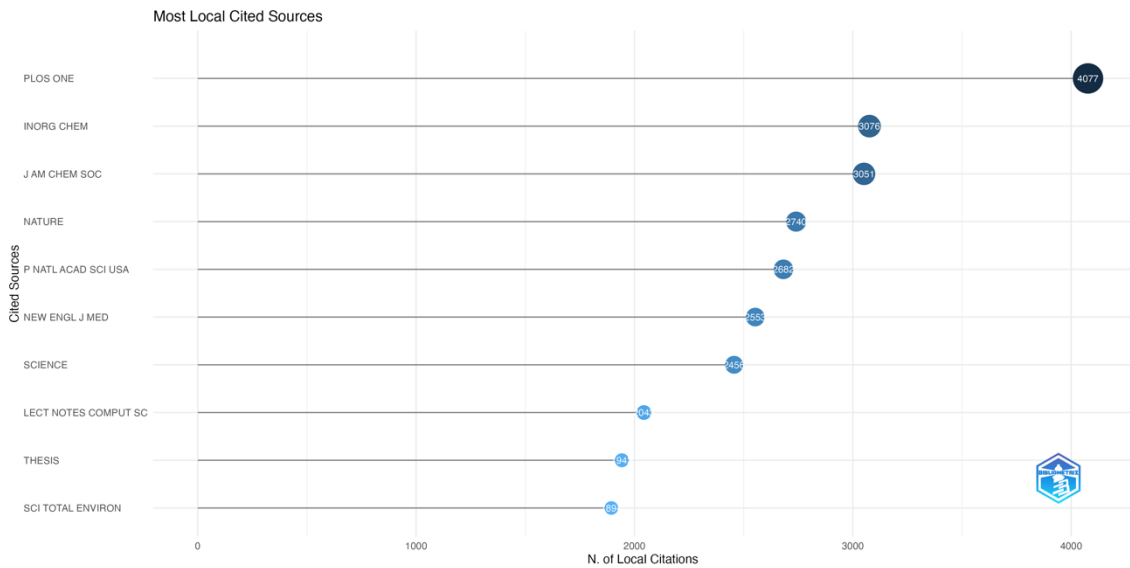


Ilustración 11. Fontes locais máis citadas

A seguinte gráfica (ver ilustración 12) móstranos o **impacto local das fontes mediante o índice h**, onde destaca a revista “Inorganic Chemistry” con un total de 29 h, seguida da revista de “Plos One” con 28h.

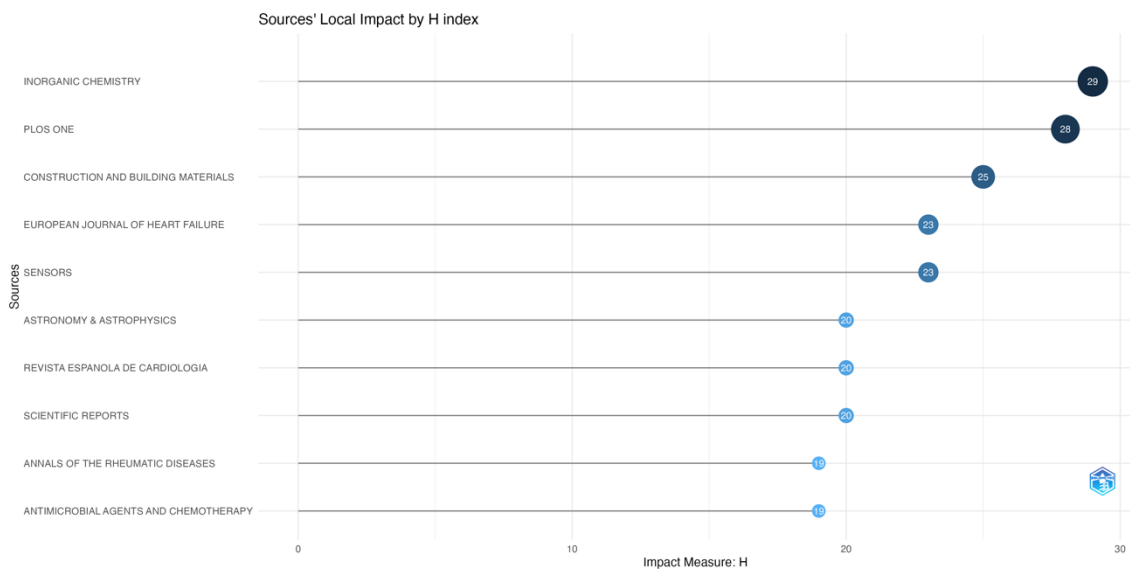


Ilustración 12. Impacto local das fontes mediante o índice h

Esta gráfica (ver ilustración 13) móstranos as **fontes de produción ao longo do tempo** dende o ano 2012 ata o ano 2023, onde podemos observar que a primeira revista “Annals of the rheumatic diseases” é a que máis continuidade tivo ao longo do tempo.

Tamén destaca a revista “International Journal of Environmental Research And Public Health” que aínda que foi das máis baixas, no ano 2019 subiu considerablemente sendo case unha das mellores revistas e facendo unha gran competencia en moi pouco tempo. Tamén se pode observar cando empezou a destacar o grupo de reumatoloxía na UDC a través das novas liñas de investigación, ou alternativamente, pódese ver que a revista “Osteoarthritis and Cartilage” entra no WOS.

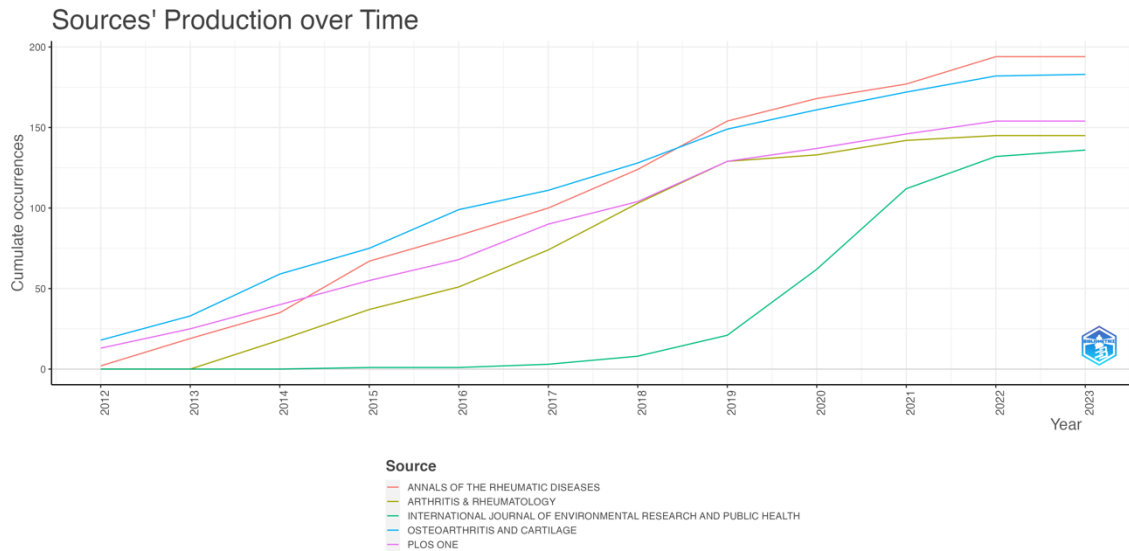
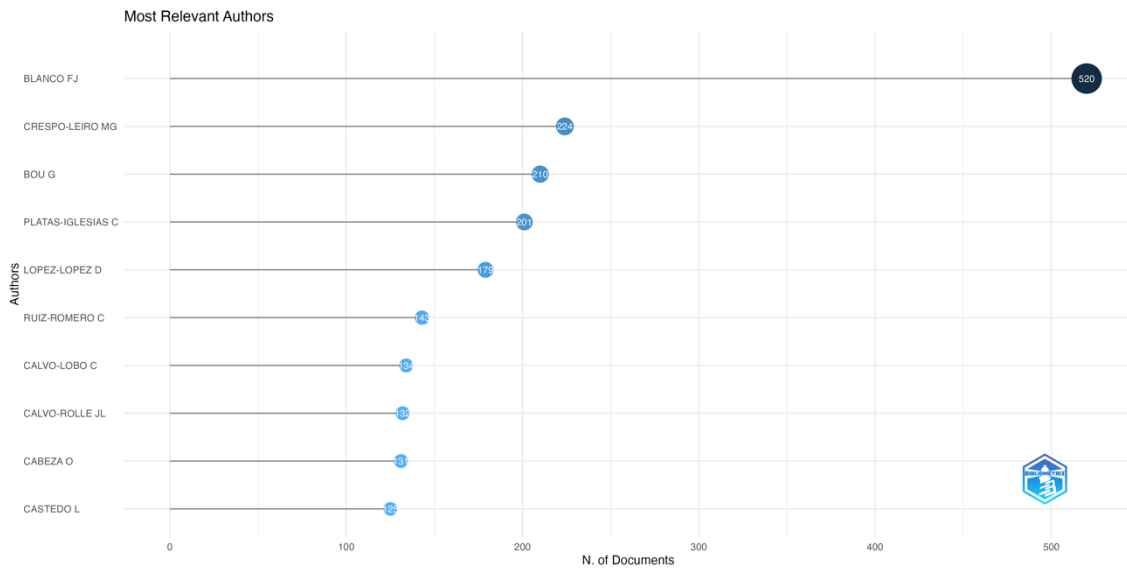


Ilustración 13. Fontes de produción ao longo do tempo

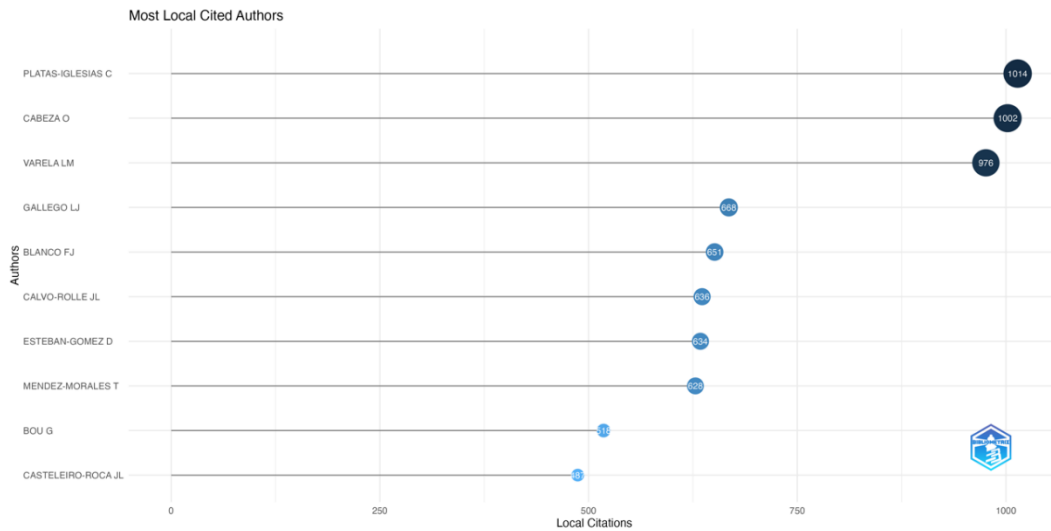
## AUTORES

Podemos apreciar os **autores máis relevantes da Universidade da Coruña** (ver ilustración 14), o que máis número de documentos publicou é Blanco FJ con un total de 520 documentos.



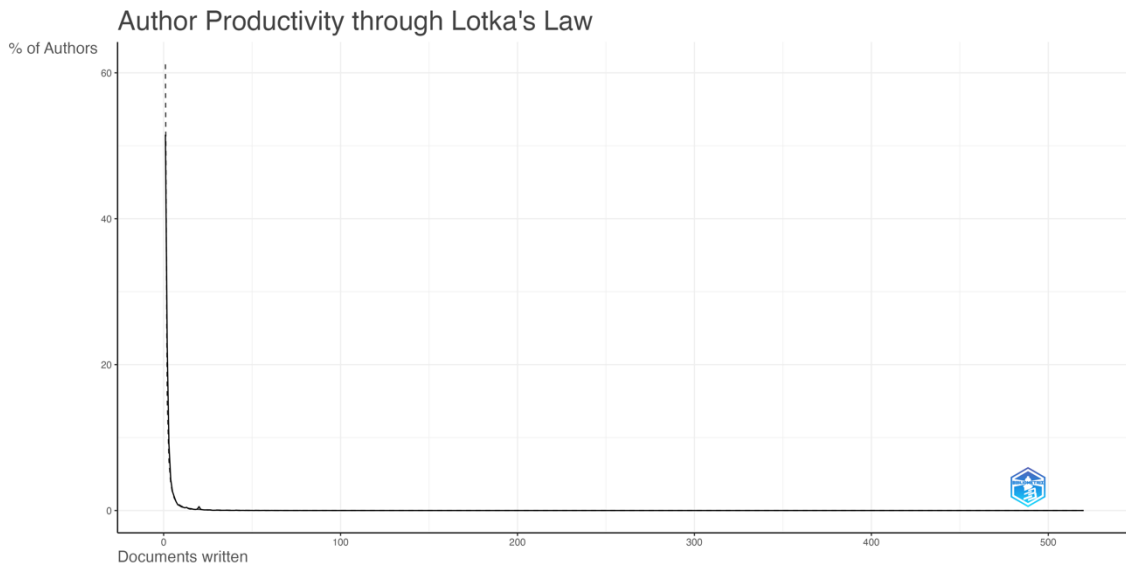
*Ilustración 14. Autores máis relevantes*

En canto aos **autores locais máis citados** (ver ilustración 15) encabézase por tres autores destacados con un número de citas comprendido entre 950 e 1000. Estes autores son Platas-Iglesias C como primeiro, Cabeza O como segundo e Varela LM como terceiro.



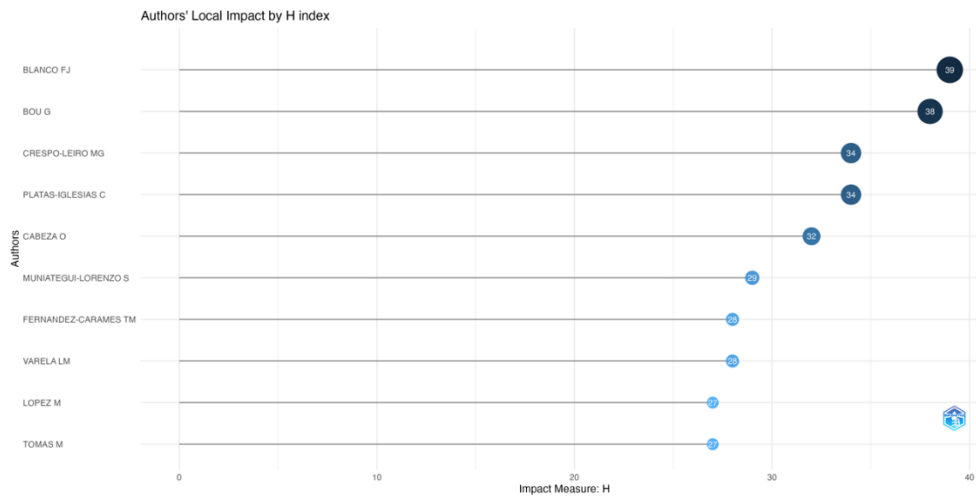
*Ilustración 15. Autores locais máis citados*

Na seguinte gráfica (ver ilustración 16) vemos a **lei de Lotka** onde se obtén o axuste dos datos. O porcentaxe de autores decrece lentamente respecto do número de documentos. Como se ve, existe un número reducido de investigadores, a elite, que son os altamente produtivos e que a maioría dos investigadores publican un número relativamente pequeno de traballos. Tamén se podería chamar, lei de reparto asimétrico da produción científica.



*Ilustración 16. Lei de Lotka*

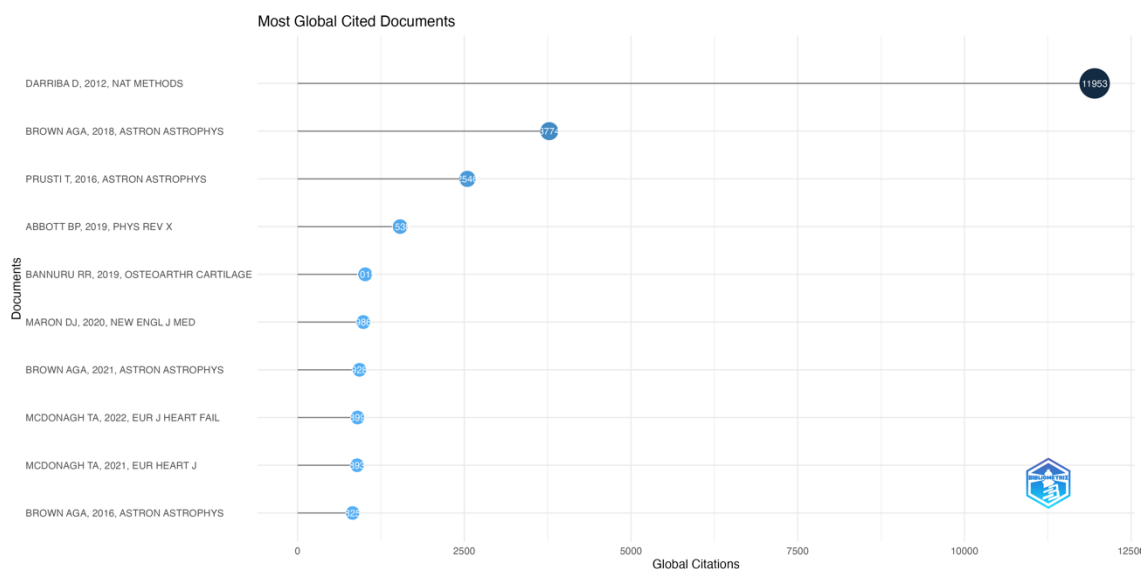
No **impacto local dos autores por índice H** (ver ilustración 17) vemos que o autor Blanco FJ volve a repetir como o autor que máis índice H ten con 39h. Este autor é un dos investigadores máis relevantes no eido da reumatoloxía no mundo.



*Ilustración 17. impacto local dos autores por índice H*

## DOCUMENTOS

Como **documentos globais máis citados** (ver ilustración 18) temos con 11953 citacións a publicación “Darriba D, Taboada GL, Doallo R, Posada D.. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772”.



*Ilustración 18. Documentos globais máis citados*



Como **documentos locais máis citados** (ver ilustración 19) temos, con 49 citas, a publicación “Esteban-Gomez D, 2012, Chemphyschem”.

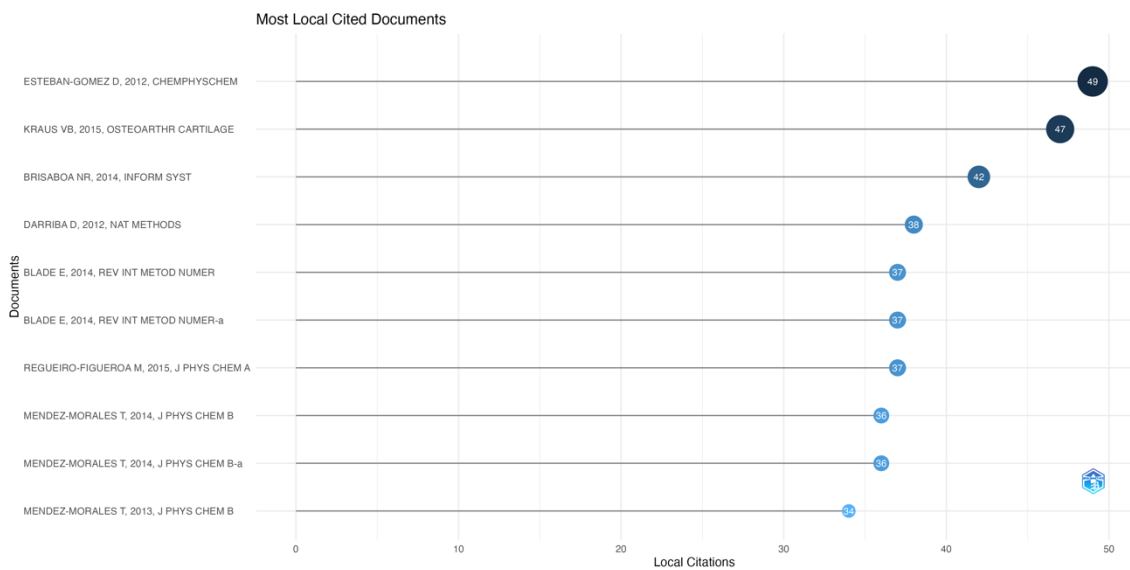


Ilustración 19. Documentos locais máis citados

As **referencias locais con máis citas** (ver ilustración 20) encontrase con 144 citas a referencia “Cohen-J., 1988, Stat Power Anal Beha”.

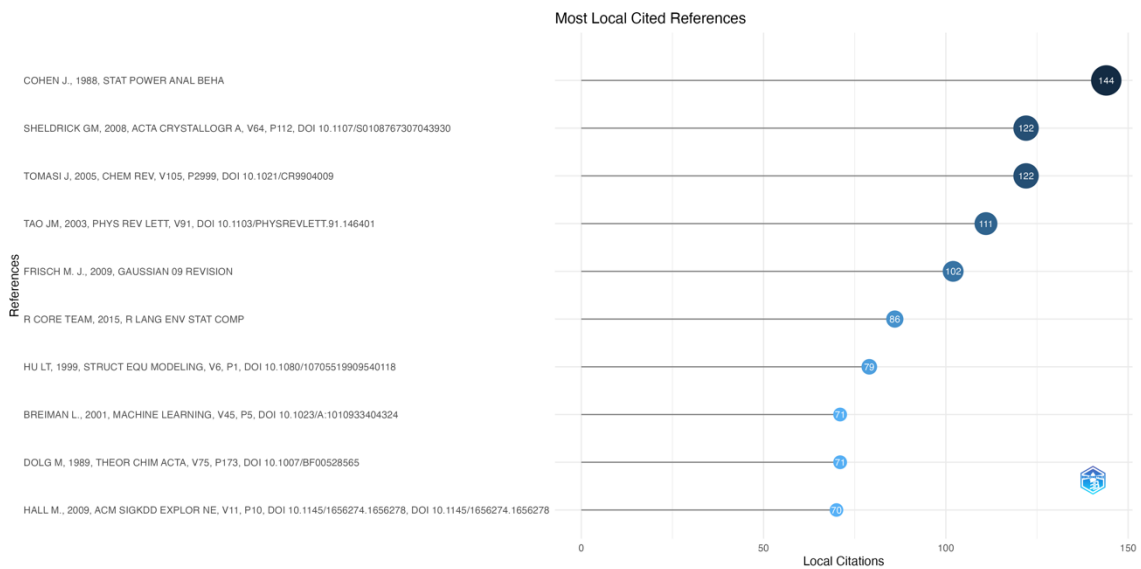


Ilustración 20. Referencias locais con máis citas

## PALABRAS

Como **palabras máis relevantes** (ver ilustración 21) temos “spain”, “covid-19” e “machine learning”. Estas palabras son sacadas das palabras chave dos documentos, o cal pódenos axudar a saber que temáticas son as máis predominantes na UDC.

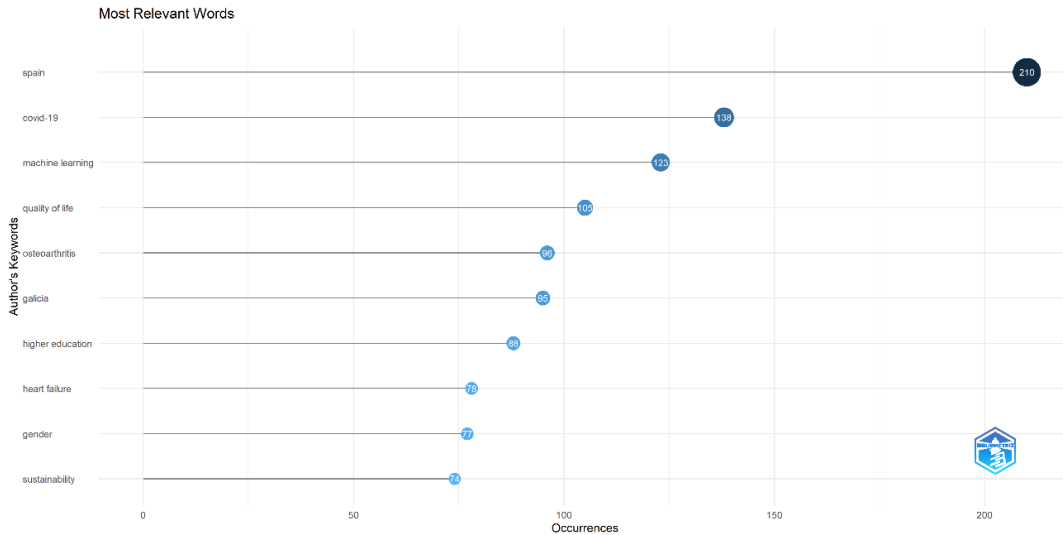


Ilustración 21. Palabras máis relevantes

Outra maneira de representar as palabras máis destacadas de maneira máis visual é este tipo de gráfico (ver ilustración 22) onde mostran as palabras máis usadas de un tamaño máis grande que as demais. Como observamos unha das palabras que podemos observar é a de “osteoarthritis”.



Ilustración 22. Gráfico de palabras



## UNIVERSIDADE DE VIGO

### DESCRIPCIÓN XENERAL

Na seguinte gráfica (ver ilustración 24) podemos ver a **producción científica anual** de UVigo, o cal podemos apreciar que o número de documentos baixa un pouco no ano 2014 pero a partires dese ano en adiante sube considerablemente esta produción sobrepasando 2000 artigos no ano 2021. Tamén se aprecia que despois do ano 2019 as producións baixan podendo ser a causa a un factor sanitario a nivel mundial como é o COVID-19. Por último, no ano 2023 as producións caen porque nos encontramos actualmente en este período, o que fai que en WOS aínda non se teñan moitas publicacións actuais.

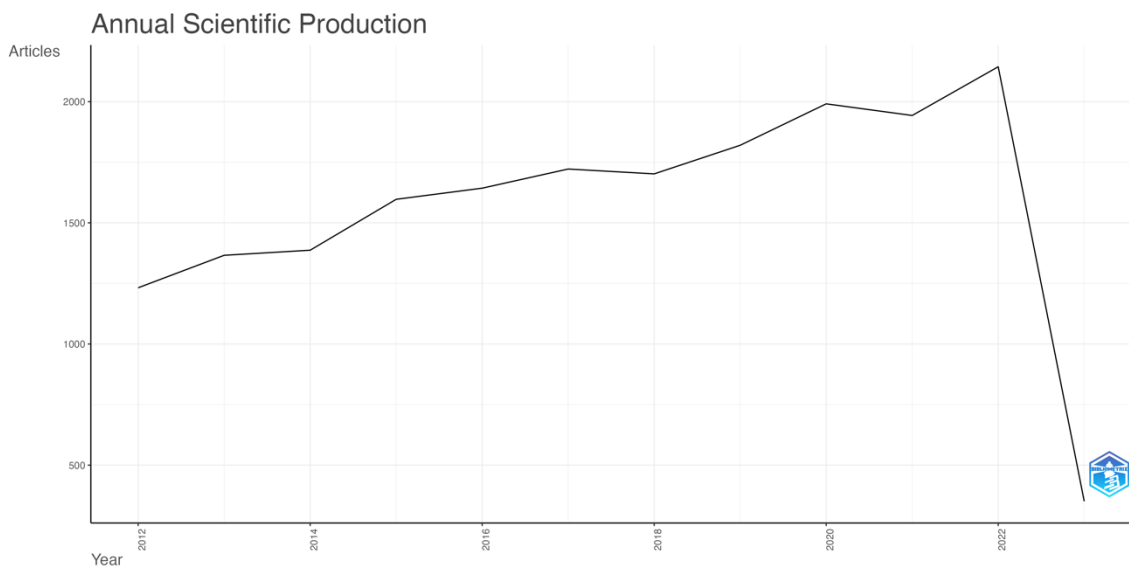
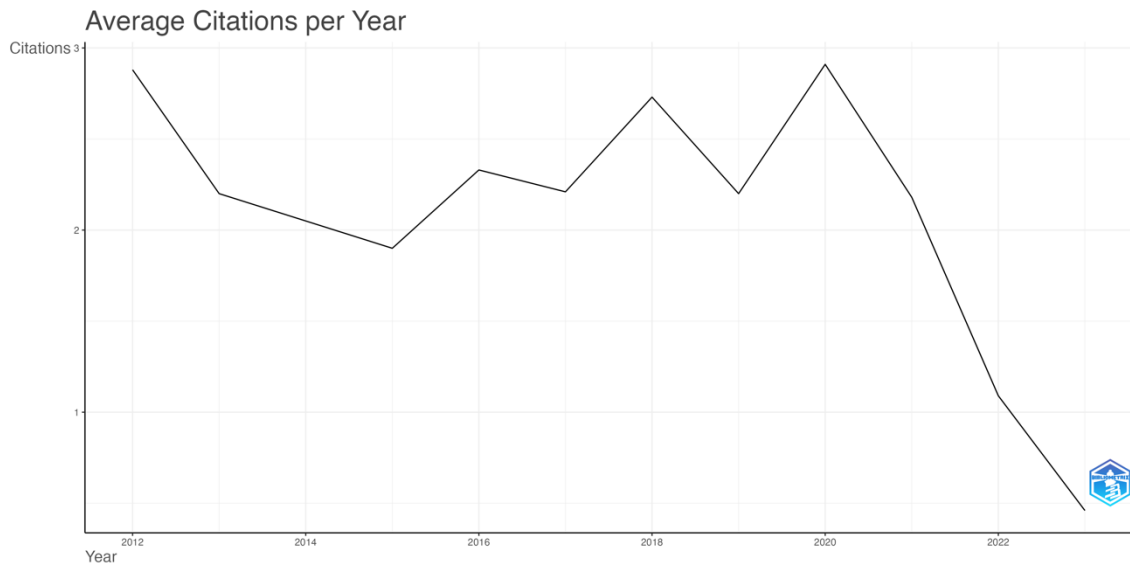


Ilustración 24. Producción científica anual

A seguinte gráfica (ver ilustración 25) indícanos as **citas promedias por ano** onde se observa que no ano que máis citas aparecen é o 2020. Vemos que un dos peores anos en citas foi o 2012 con 1.5 de citas, e a partires de aí temos unhas subidas e baixadas considerables.



*Ilustración 25. Citas promedias por ano*

## FONTES

Na seguinte gráfica (ver ilustración 26) podemos observar as **fontes máis relevantes** onde aparecen as 10 revistas que máis números de documentos posúen, encabezando a revista “Science of the total environment”.

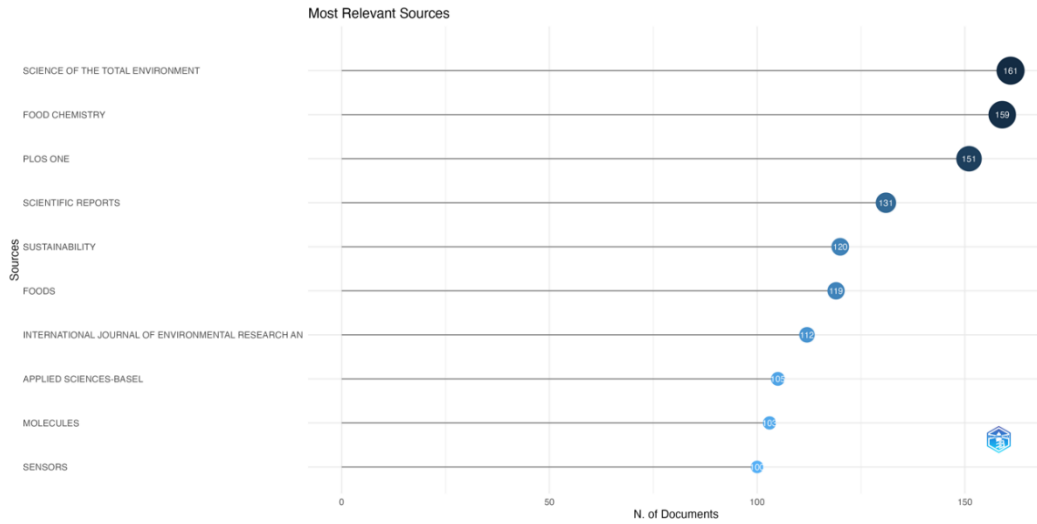


Ilustración 26. Fontes máis relevantes

En canto as **fontes locais máis citadas** (ver ilustración 27) destaca con a revista “Food Chem” con un total de 7410 citas locais.

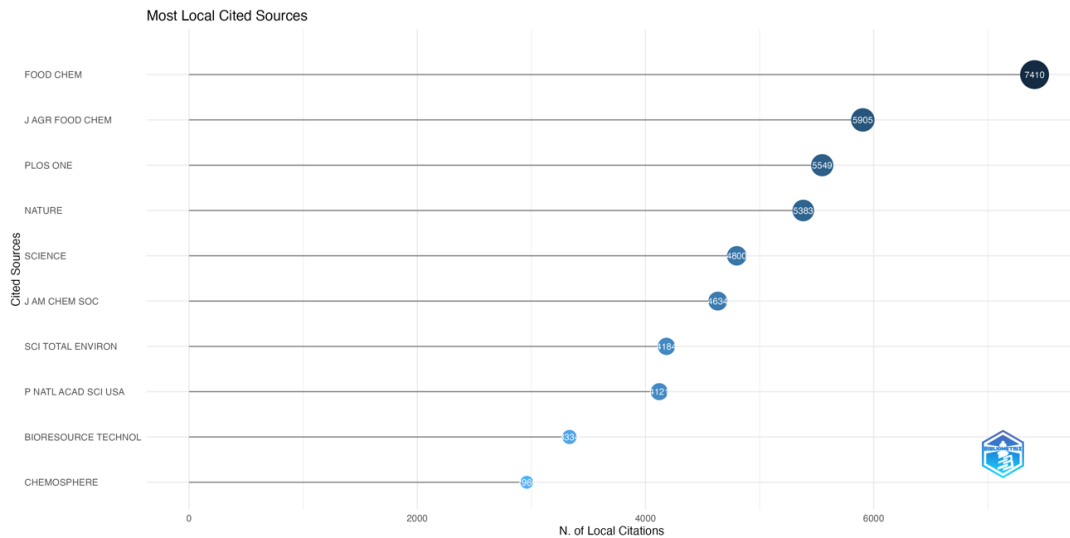
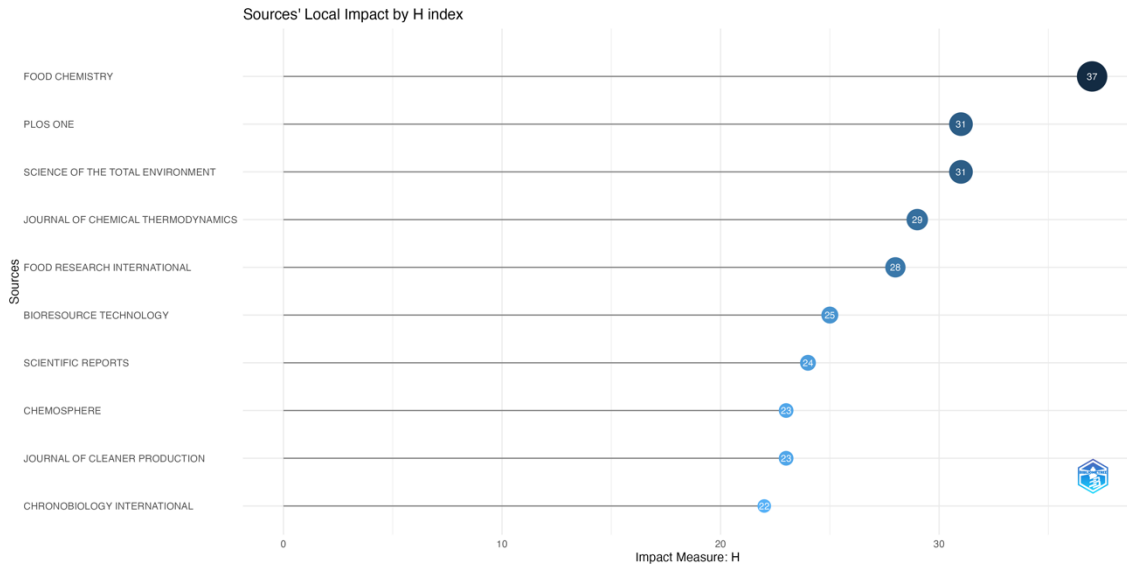


Ilustración 27. Fontes locais máis citadas

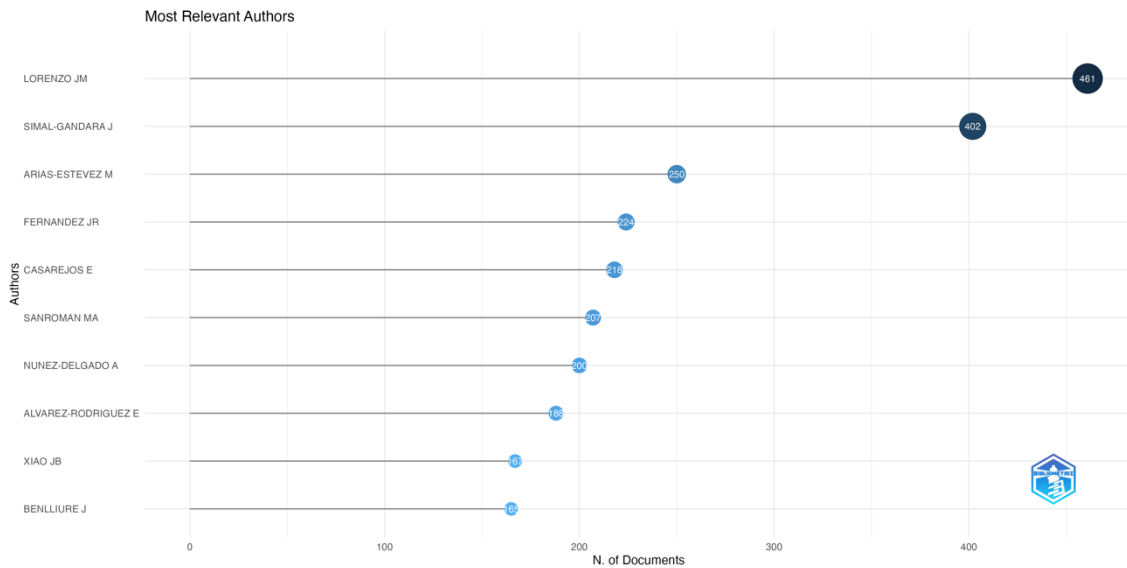
A seguinte gráfica (ver ilustración 28) móstranos o **impacto local das fontes mediante o índice h**, onde destaca a revista “Food Chemistry” con un total de 37 h, seguida da revista de “Plos One” con 31h.



*Ilustración 28. Impacto local das fontes mediante o índice h*

## AUTORES

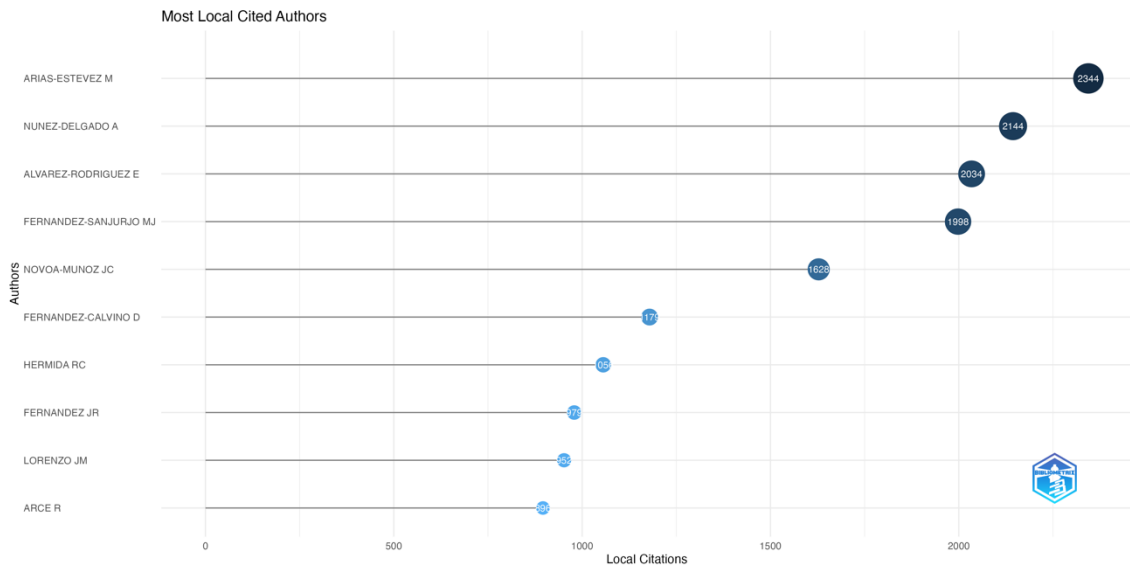
Podemos apreciar os **autores máis relevantes da Universidade de Vigo** (ver ilustración 29), o que máis número de documentos publicou é Lorenzo JM con un total de 461 documentos.



*Ilustración 29. Autores máis relevantes*

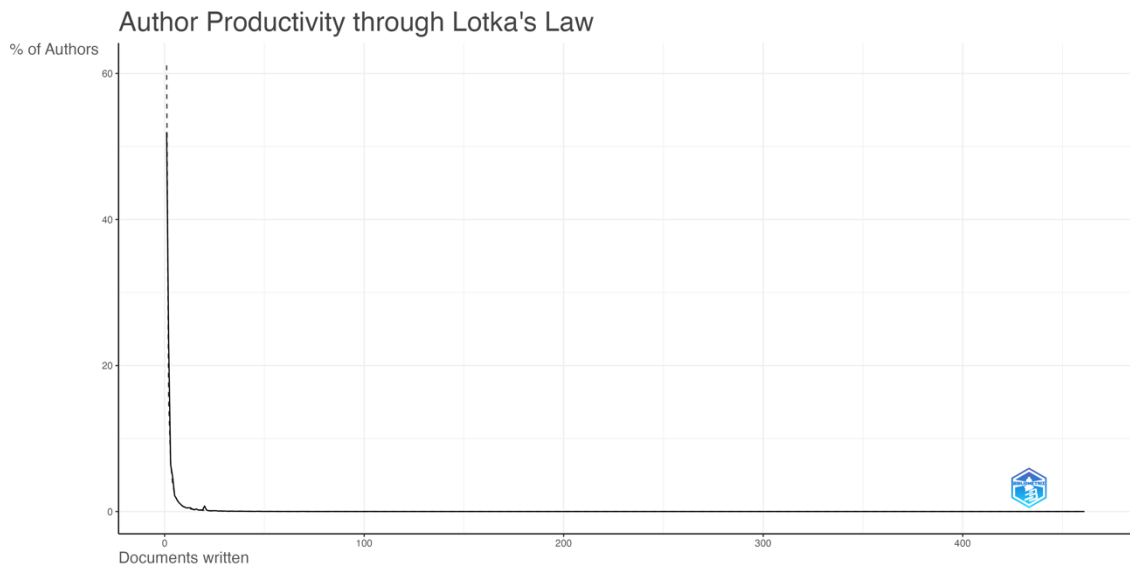
En canto aos **autores locais máis citados** (ver ilustración 30) encabezase por tres autores destacados con un número de citas comprendido entre 2000 e 2350. Estes autores son Arias Estévez M, Núñez Delgado A e Álvarez Rodríguez E.





*Ilustración 30. Autores locais máis citados*

Na seguinte gráfica (ver ilustración 31) vemos a **lei de lotka** onde se obtén o axuste dos datos. O porcentaxe de autores decrece lentamente respecto do número de documentos.



*Ilustración 31. Lei de Lotka*

No **impacto local dos autores por índice H** (ver ilustración 32) vemos que o autor Liz Marzan LM é o autor que máis índice H ten con 55h.

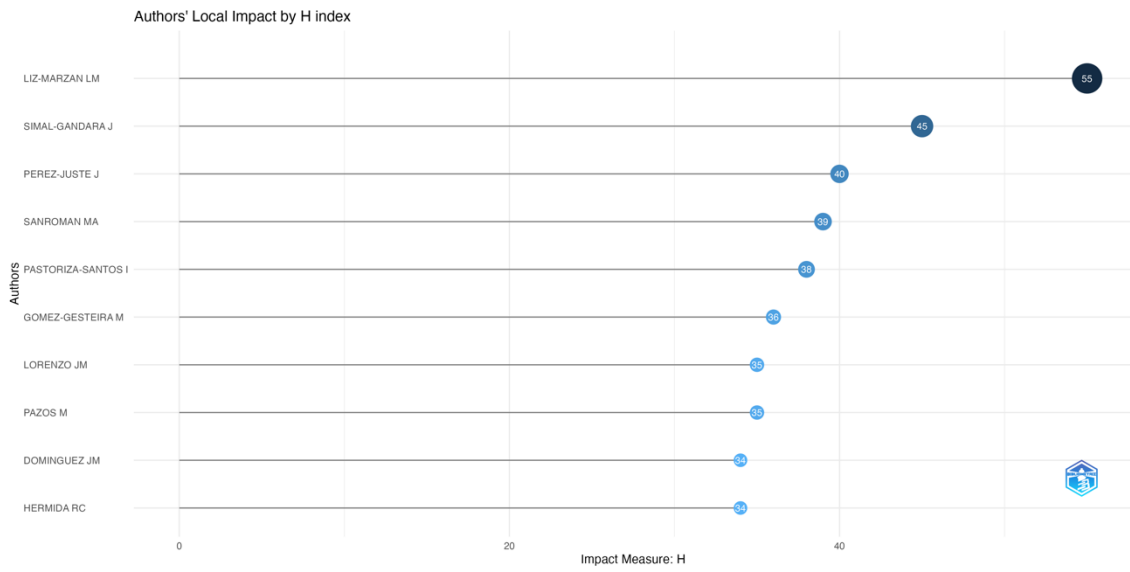


Ilustración 32. Impacto local dos autores por índice H

## DOCUMENTOS

Como **documentos globais máis citados** (ver ilustración 33) temos con 11953 citacións a publicación “Darriba D, 2012, Nat Methods”.

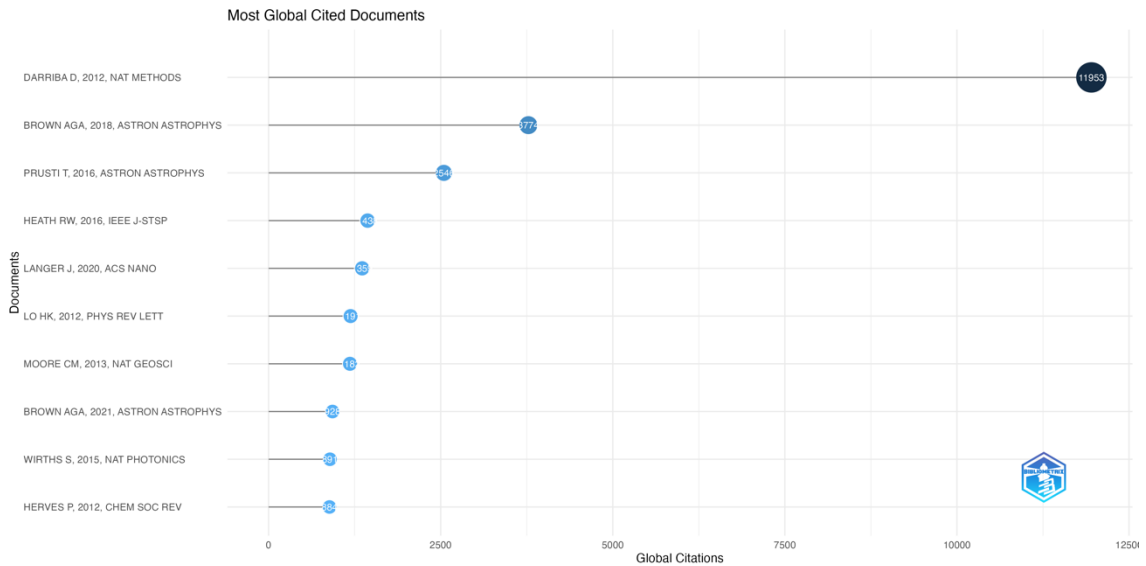


Ilustración 33. Documentos globais máis citados

Como **documentos locais máis citados** (ver ilustración 34) temos con 72 citas a publicación “Gimeno L, 2012, Rev Geophys”.

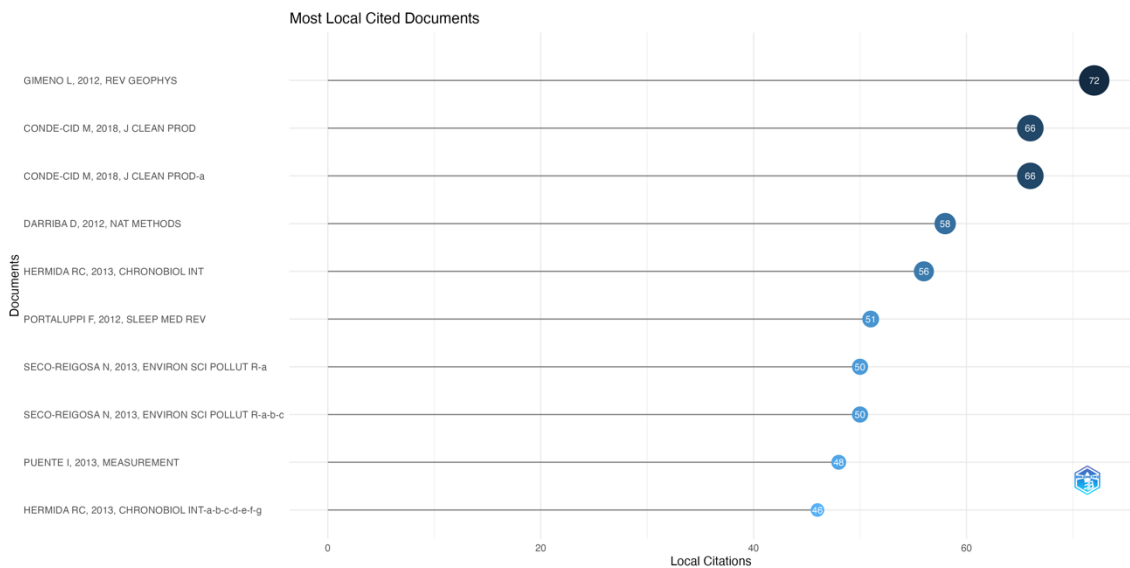
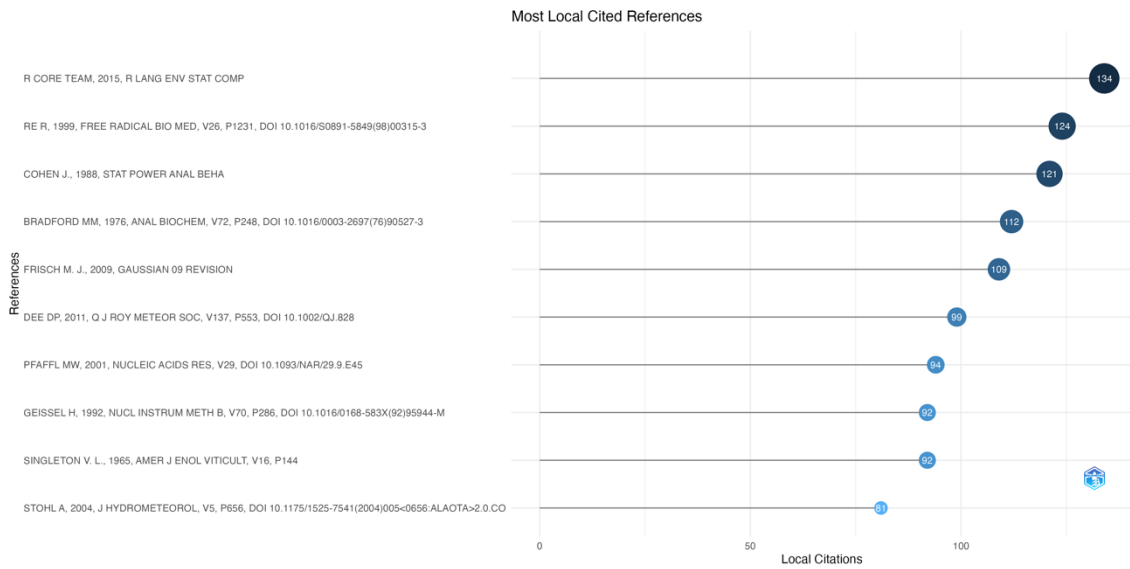


Ilustración 34. Documentos locais máis citados

As **referencias locais con máis citas** (ver ilustración 35) encontráse con 134 citas a referencia “R Core Team, 2015, R Lang Env Stat Comp”.

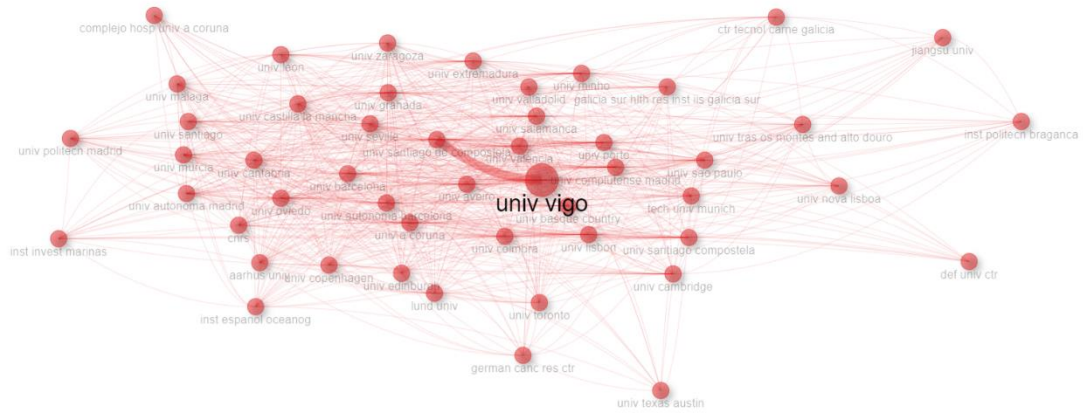


*Ilustración 35. Referencias locais con máis citas*



## REDE DE COLABORACIÓN

A gráfica de rede de colaboración (ver ilustración 38) da UVigo móstranos con que universidades e institucións públicas participa máis en canto a produción.

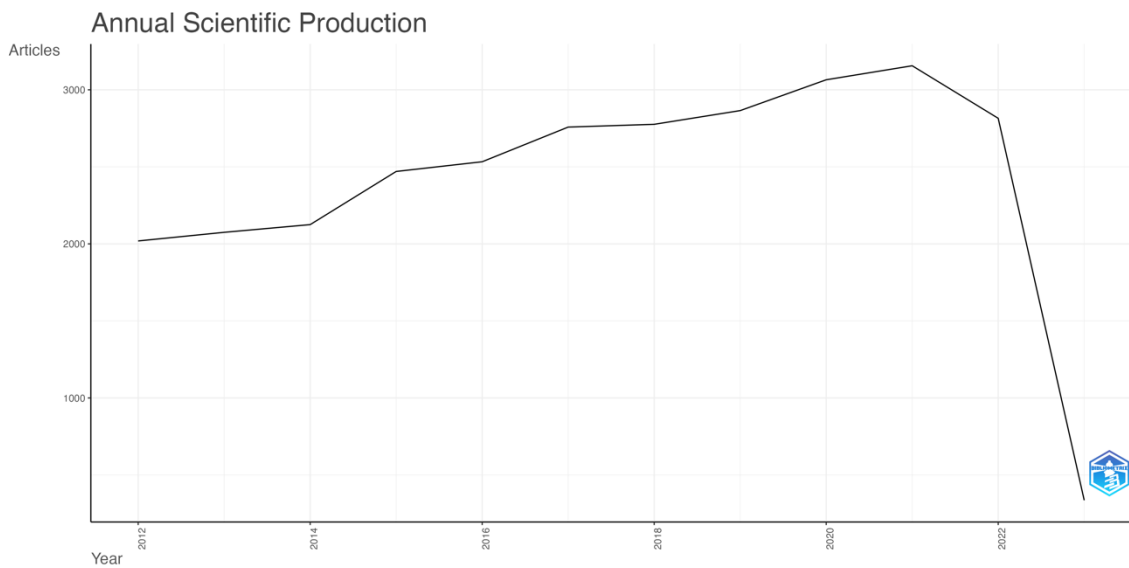


*Ilustración 38. Rede de colaboración*

## UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

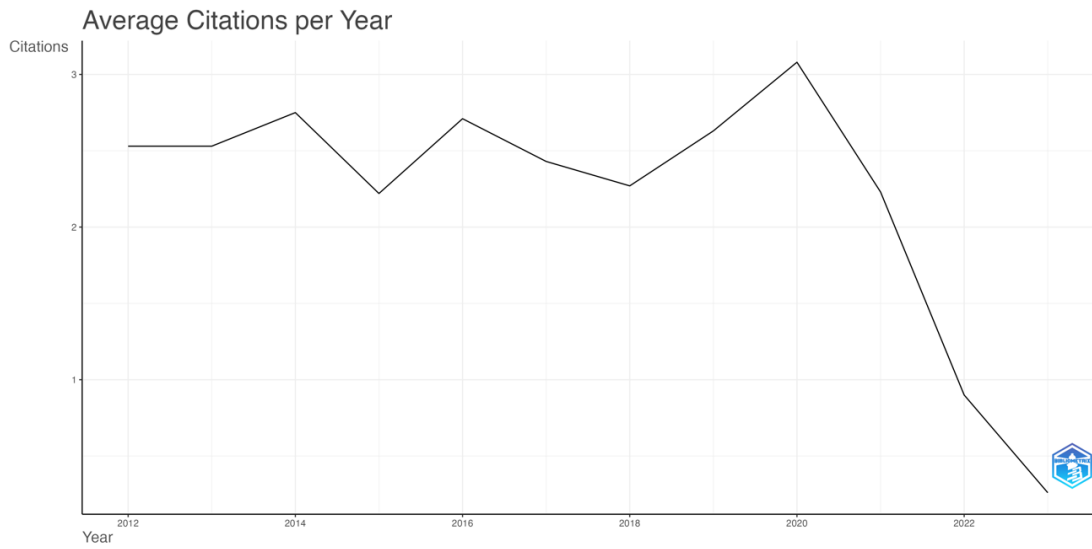
### DESCRINCIÓN XENERAL

Na seguinte gráfica (ver ilustración 39) podemos ver a **producción científica anual** da USC, o cal podemos apreciar que o número de documentos é crecente chegando case a 3000 artigos no ano 2021. Por último no ano 2023 as producións caen porque nos encontramos actualmente en este período, o que fai que en WOS aínda non se teñan moitas publicacións actuais.



*Ilustración 39. Producción científica anual*

A seguinte gráfica (ver ilustración 40) indicánnos as **citas promedias por ano** onde se observa que no ano que máis citas temos é no ano 2020. Vemos que un dos peores anos en citas foi o 2015 con 2.5 de citas.



*Ilustración 40. Citas promedias por ano*



## FONTES

Na seguinte gráfica (ver ilustración 41) podemos observar as **fontes máis relevantes** onde aparecen as 10 revistas que máis número de documentos posúen, encabezando a revista “Journal of High Energy Physics”.

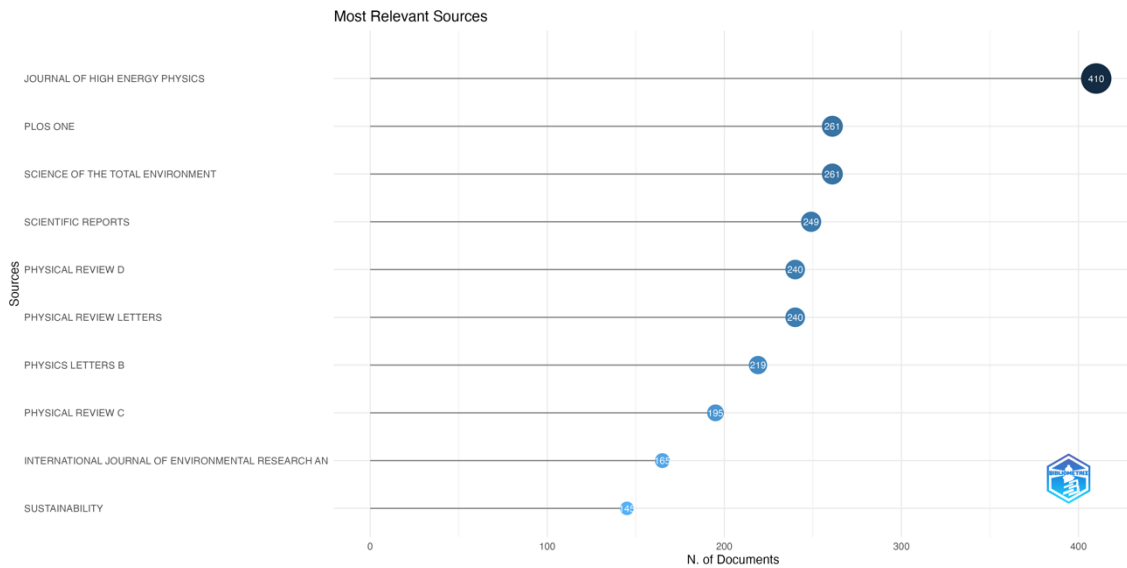


Ilustración 41. Fontes máis relevantes

En canto as **fontes locais máis citadas** (ver ilustración 42) destaca con a revista “Phys Rev D” con un total de 14919 citas locais.

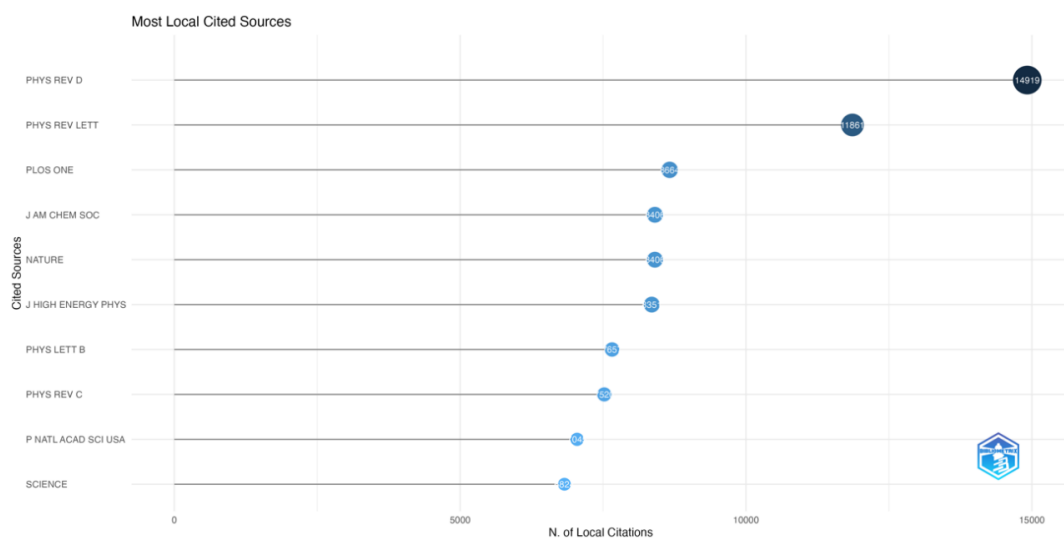
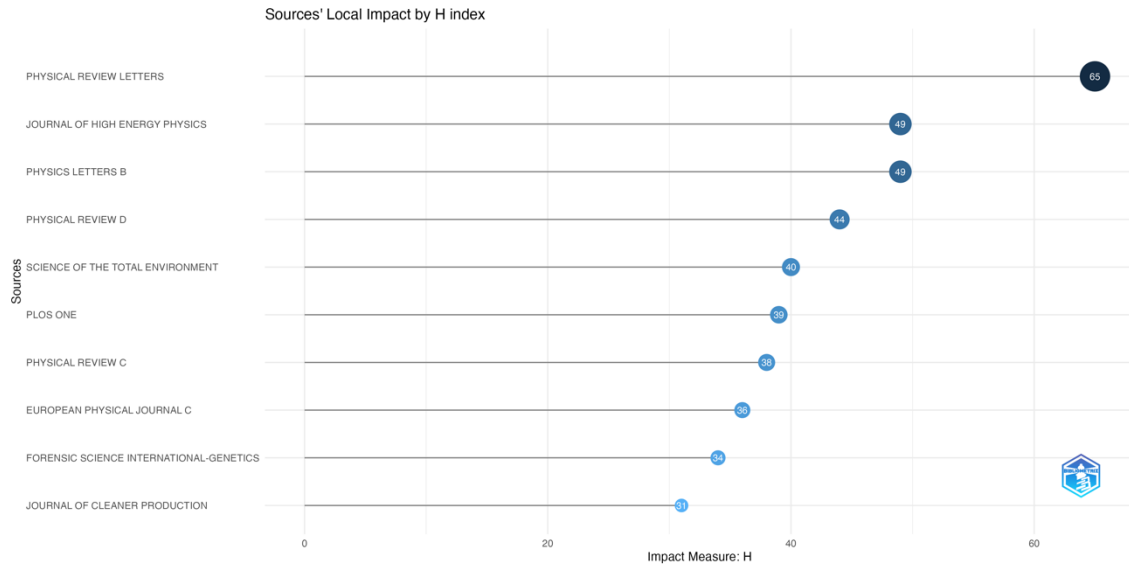


Ilustración 42. Fontes locais máis citadas

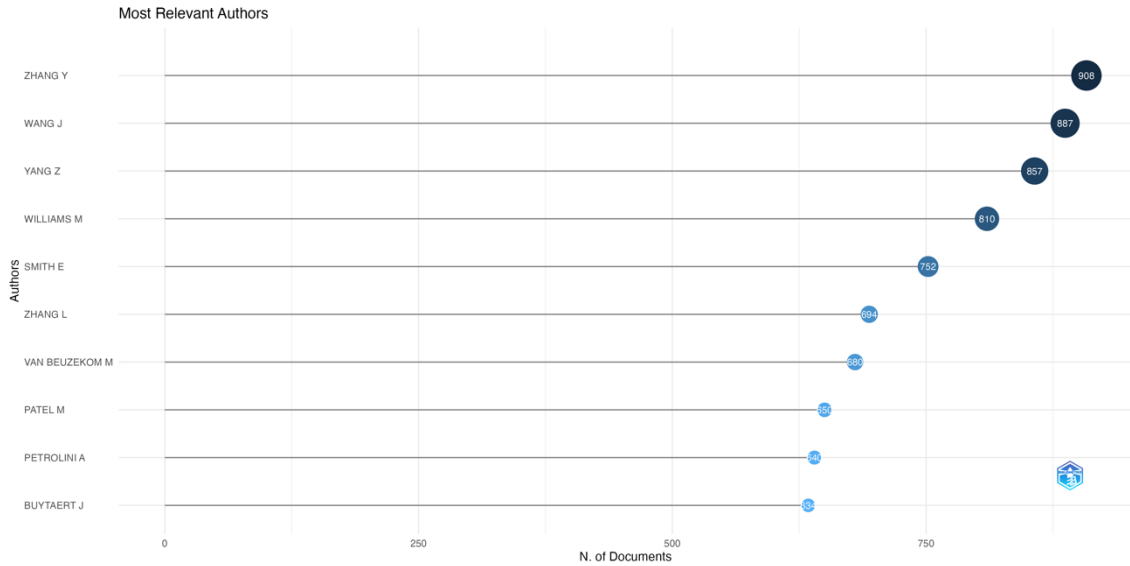
A seguinte gráfica (ver ilustración 43) mostranos o **impacto local das fontes mediante o índice h**, onde destaca a revista “Physical Review Letters” con un total de 65 h, seguida da revista de “Journal of High Energy Physics” con 49 h.



*Ilustración 43. Impacto local das fontes mediante o índice h*

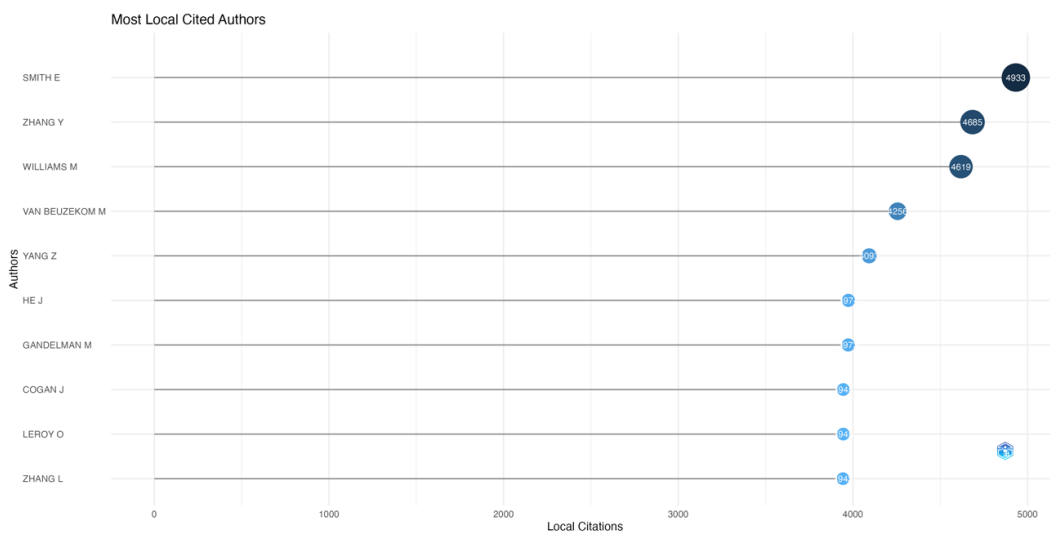
## AUTORES

Podemos apreciar os **autores máis relevantes da Universidade de Santiago de Compostela** (ver ilustración 44), o que máis número de documentos publicou é Zhang Y con un total de 908 documentos.



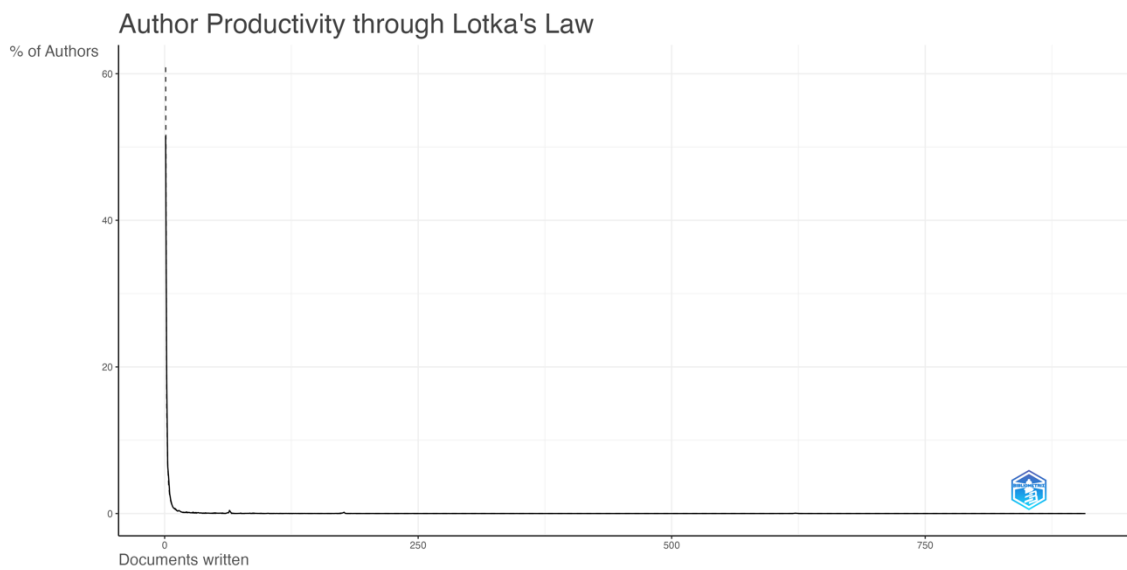
*Ilustración 44. Autores máis relevantes*

En canto aos **autores locais máis citados** (ver ilustración 45) encabezase por tres autores destacados con un número de citas comprendido entre 4500 e 5000. Estes autores son Smith E, Zhang Y e Williams M.



*Ilustración 45. Autores locais máis citados*

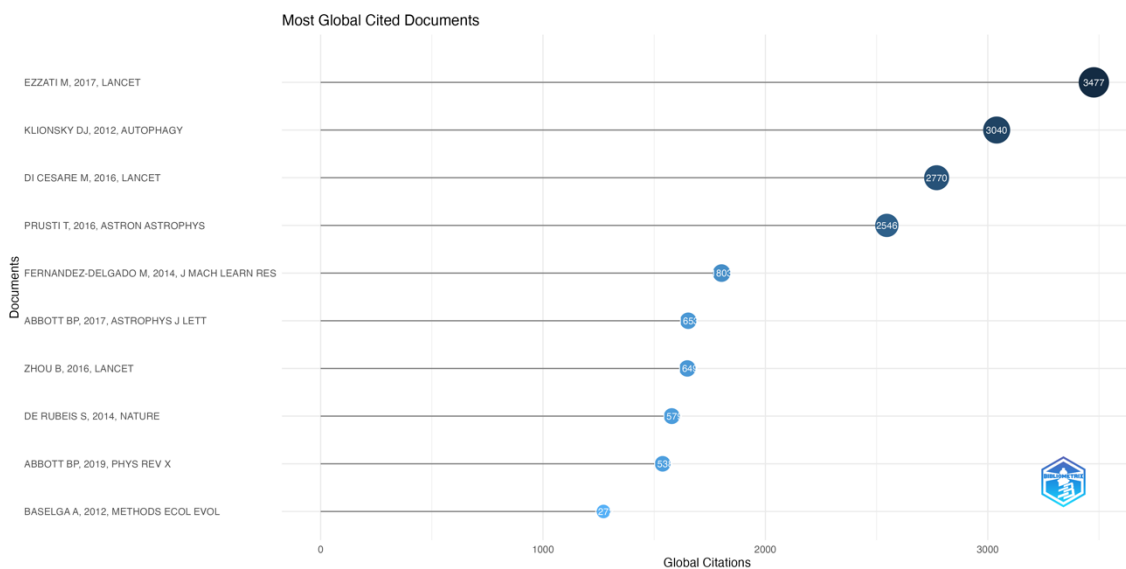
Na seguinte gráfica (ver ilustración 46) vemos a **lei de lotka** onde se obtén o axuste dos datos. O porcentaxe de autores decrece lentamente respecto do número de documentos.



*Ilustración 46. Lei de Lotka*

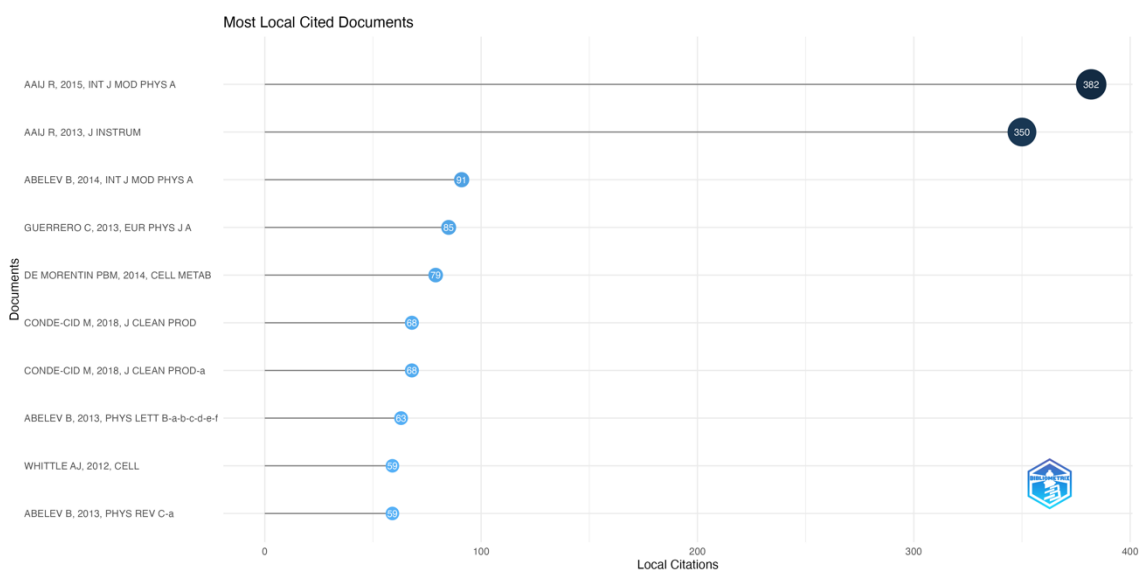
## DOCUMENTOS

Como **documentos globais máis citados** (ver ilustración 47) temos con 3477 citacións a publicación “Ezzati M, 2017, Lancet”.



*Ilustración 47. Documentos globais máis citados*

Como **documentos locais máis citados** (ver ilustración 48) temos con 382 citas a publicación “AAIJ R, 2015, Int J Mod Phys A”.



*Ilustración 48. Documentos locais máis citados*

As referencias locais con máis citas (ver ilustración 49) encontrase con 632 citas a referencia “Agostinelli S, 2003, Nucl Instrum Meth A”.

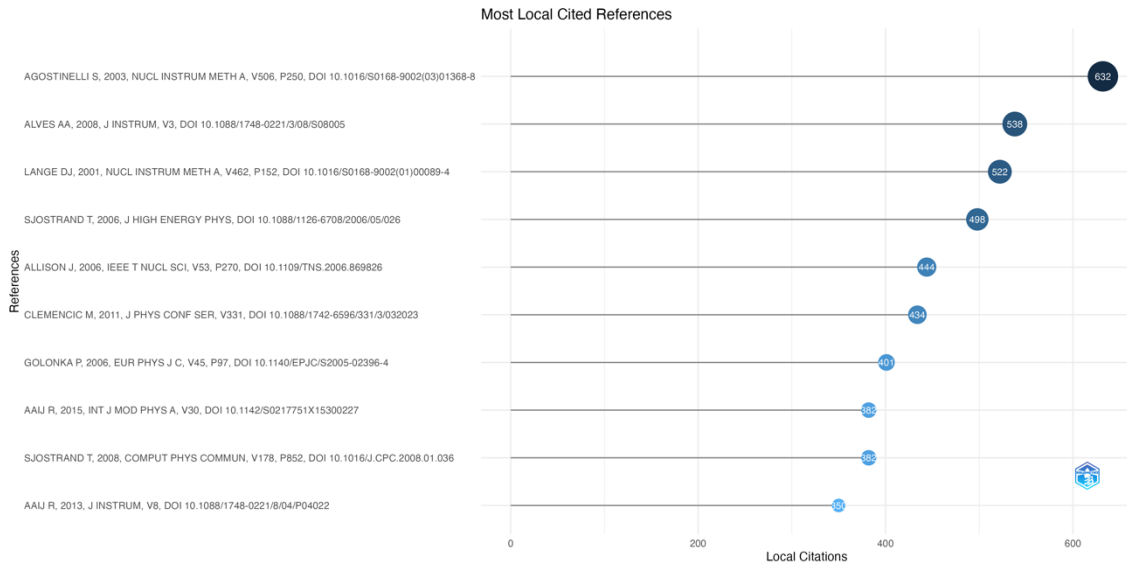


Ilustración 49. Referencias locais con máis citas

## PALABRAS

Como **palabras máis relevantes** (ver ilustración 50) temos “spain”, “galicia” e “covid-19”. Estas palabras son sacadas das palabras chave dos documentos, o cal pódennos axudar a saber que temáticas son as máis predominantes na USC.

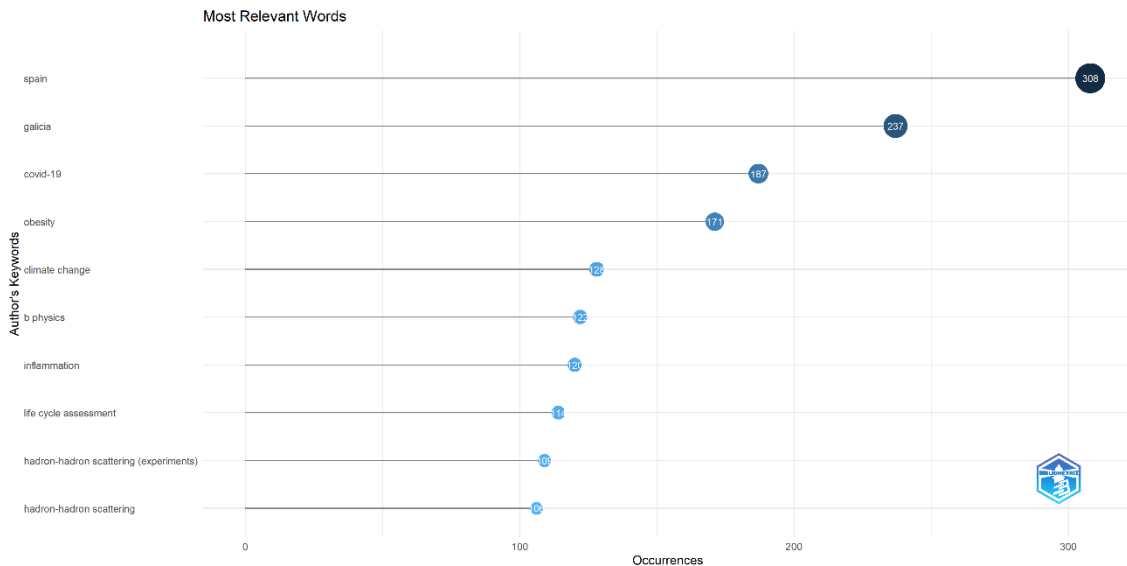


Ilustración 50. Palabras máis relevantes

Outra maneira de representar as palabras máis destacadas de maneira máis visual é este tipo de gráfico (ver ilustración 51) onde mostran as palabras máis usadas de un tamaño máis grande que as demais.

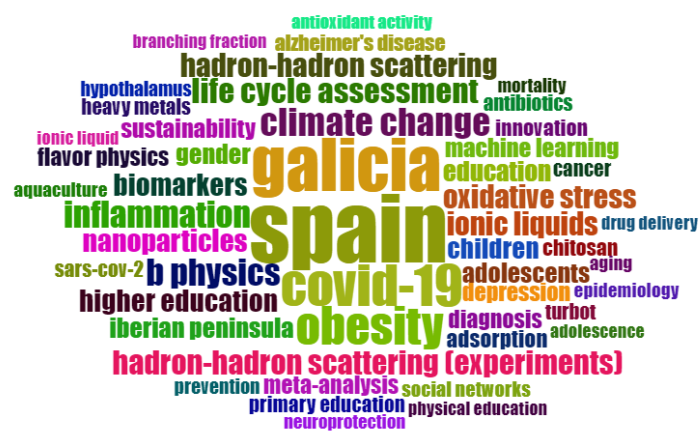


Ilustración 51. Gráfica de palabras





## COMPARACIÓN DAS INSTITUCIÓNS

Nestas gráficas de información (ver ilustracións 53, 54 e 55) aparecen os datos das tres universidades, UDC, USC e UVigo, nas cales podemos apreciar a diferenza na cantidade de documentos que existen entre as tres universidades, o número de autores, o número de coautores, as revistas, o índice de citacións, o número de coautores por documento, etc.

Vemos como a universidade máis destacada por maioría en case todos os campos é a USC, isto pode vir dado a que sexa unha das institucións máis antigas e con máis recoñecemento nacional e internacional.

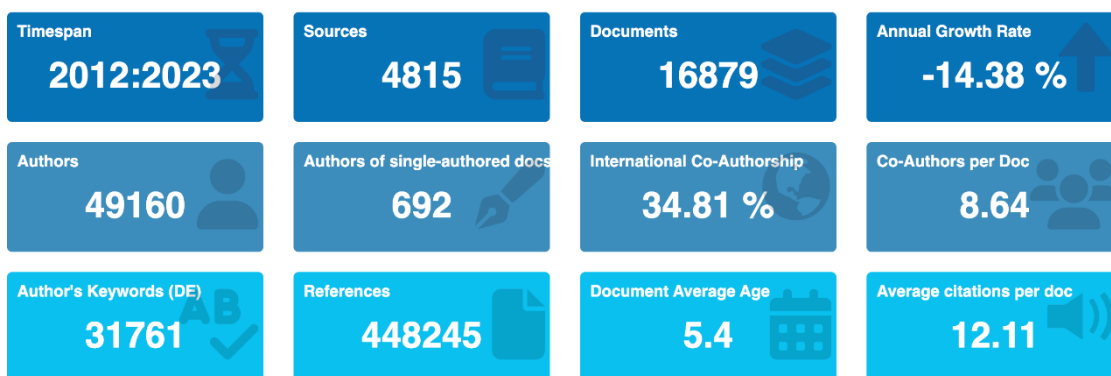


Ilustración 53. Universidade de Coruña

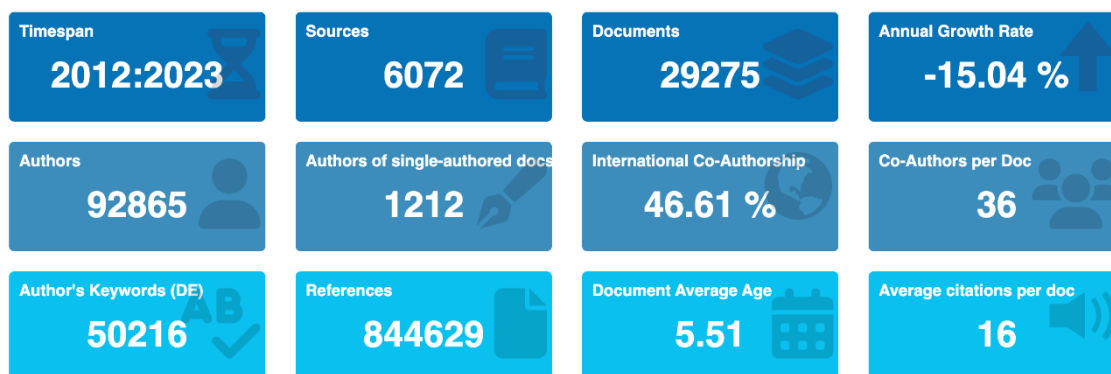


Ilustración 54. Universidade de Santiago de Compostela



Ilustración 55. Universidade de Vigo

En esta gráfica (ver ilustración 56) podemos observar o ritmo de produción crecente en cada unha das universidades ao longo dos anos, dende o 2012 ata a actualidade. Observase que a USC encabeza como primeira universidade, seguida de UVigo e UDC.

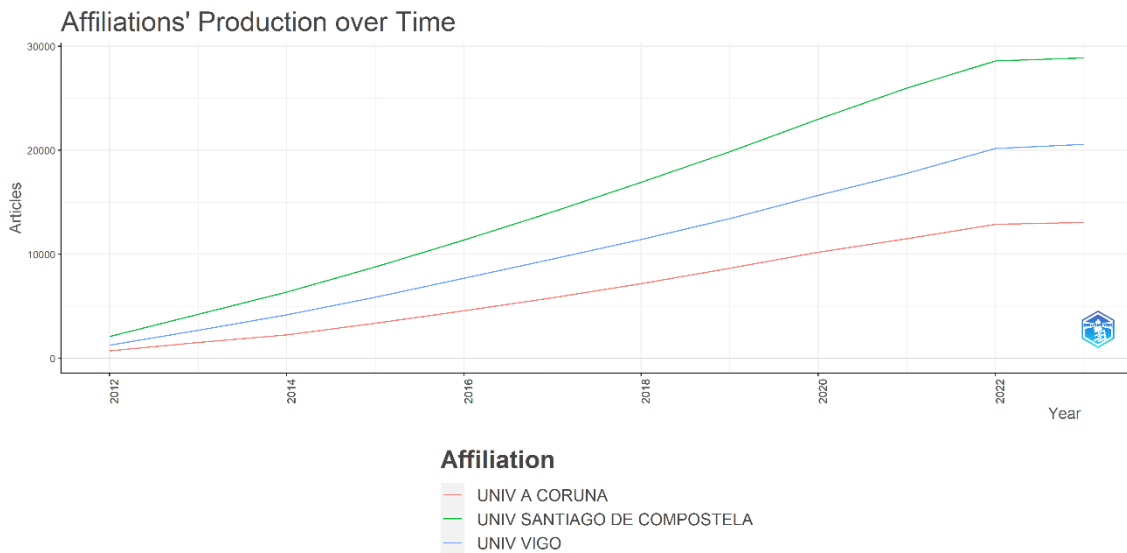


Ilustración 56. Producción das institucións

## CONCLUSIÓNS

No traballo aplícase unha metodoloxía de análise da información bibliométrica, desde un punto de vista estatístico coa axuda do software libre R e coa axuda do paquete bibliometrix. Este procedemento permite levar a cabo un tratamento dos datos estatístico máis específico a partir de bases de datos moito máis grandes, permitindo traballar con gran cantidade de datos, sen limitación de publicacións científicas. A base de datos escollida para o análise é WOS, este repositorio contén información relevante das tres universidades galegas, podendo tamén refinar a búsqueda nos anos comprendidos entre 2012 e o 2023.

Dito isto, a análise correspóndese cun estudo estatístico de tipo descritivo da produción científica do sistema universitario galego no cal se observa como rede de investigación principal a USC, tendo índices máis altos que a UDC e a UVigo. Isto acontece, revisando a historia de cada universidade, porque a USC é das universidades máis antigas, con gran prestixio a nivel nacional e internacional, cun grupo moito maior de investigación, que a súa vez tamén está subestimado polo gran número de autores, que aparte de xerar gran cantidade de información, tamén posúen un gran número de citas. Tanto a UDC como a UVigo, dependendo do ano que se creou, moi posterior a USC, tamén son universidades que se achegan aos datos da USC, con altos números de citas e con unha gran demanda de produción na última década.

En canto á UVigo respecto da USC, sabendo que é unha das universidades máis novas xunto coa UDC, os índices son moi superiores e en moitos casos, aproximándose á USC. Isto é debido a que a universidade ten un alto porcentaxe de produtividade con un alto impacto de citas para a comunidade científica. Como podemos observar ao longo do traballo, a UVigo é a universidade que segue á USC en canto a produción en estes anos estudados.

Tamén observar que tanto a UDC como a UVigo encabezan o cumprimento da lei de crecemento exponencial na ciencia, observando que son as institucións que máis se acercan á USC nos últimos anos en canto a velocidade de crecemento da produtividade.

Coa lei de Lotka aplicada ás tres universidades galegas, vemos que existe un número reducido de investigadores, a elite, altamente produtivos, mentres que os demais

investigadores publican un número pequeno de publicacións científicas. Estas casuísticas veñen acompañadas de varios factores como pode ser o caso do tipo de contratacións que fai cada universidade en cada seu respectivo ano, podendo traer investigadores altamente produtivos en un período corto de tempo.

Unha vez sacamos que universidade é máis produtiva, así como as que máis medraron esta última década, tamén se avalían os índices bibliométricos de cada fonte ou revista, autor e publicacións, analizando aqueles que máis destacan dentro de cada universidade.

É de especial interese saber en que revistas publica cada universidade e investigador, de feito, observamos que existen tres tipos de revistas donde máis se publican artigos con moitas citas. Estas revistas son as de full open access, mega revistas (Plos One) e MDPI, estas últimas moito máis rápidas que as demáis. Con isto podemos sacar como conclusións en que revistas publican máis os investigadores de cada universidade, afectando tamén ao crecemento e a produtividade de cada unha, xa que non é o mesmo publicar nunha revista que tarda un ano en facer oficial a investigación que unha que tarda dous meses.

Respecto ao análise bibliométrico das palabras clave dos autores de cada universidade, podemos observar que temas se tratan en cada unha delas. Observamos por exemplo que na UDC destaca a osteoartritis debido a alta produtividade e a calidade que temos sobre este tema, grazas a un dos autores que o estuda e o publica. De xeito que cada universidade vai a ter diferentes puntos de vista sobre os temas a tratar, excepto un que si que aparece en todas con alta incidencia, o covid-19, que abarcou un dos temas máis analizados destas tres universidades.

Por último, podemos observar nunha das gráficas a rede de colaboración que ten cada universidade, destacando o grao de internacionalización, moi por encima, que ten a USC respecto de UDC e UVigo. É das universidades que máis colabora con outras institucións públicas doutros países. Observando a UDC e a UVigo colaboran moito con universidades españolas doutras comunidades autónomas.

## BIBLIOGRAFÍA

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975. Elsevier.

CICA. (2023). Centro Interdisciplinar de Química e Bioloxía. Recuperado de <https://cica.udc.gal/>

CITIC. (2023). Centro de Investigación en Tecnologías de la Información y las Comunicaciones. Recuperado de <https://citic.udc.es/>

Clarivate. (s.f.). Web of Science. Recuperado de <https://clarivate.com/webofsciencelgroup/solutions/web-of-science/>

Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics*, 69(1), 131-152.

Elsevier. (s.f.). Scopus. Recuperado de <https://www.scopus.com/>

Hérubel, J. P. (1999). Historical Bibliometrics: Its Purpose and Significance to the History of Disciplines. *Libraries and Culture*, 34(4), 380-388.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572.

Khodiyar, V. K., Rowlett, K., & Lawrence, R. N. (2014). Altmetrics as a Means of Assessing Scholarly Output. *Learned Publishing*, 27(5), S25-S32.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317-323.

Price, D. J. de S. (1965). Networks of scientific papers. *Science*, 149(3683), 510-515.

Pritchard, A., & Wittig, G. G. (1981). *Bibliometrics: A Bibliography and Index*, Vol. 1. Watford: ALLM Books.

Priem, J., Tarabonelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A Manifesto. Recuperado de <http://altmetrics.org/manifesto/>

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Recuperado de <https://www.R-project.org/>

Roemer, R. C., & Borchardt, R. (2015). Meaningful Metrics: A 21st-Century Librarian's Guide to Bibliometrics, Altmetrics, and Research Impact. Chicago: Association of College and Research Libraries.

Sengupta, I. N. (1992). Bibliometrics, Informetrics, Scientometrics and Librametrics: An Overview. *Libri*, 42(2), 75-98.

Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1-116.

Torres-Salinas, D., & Moed, H. F. (Eds.). (2019). *Proceedings of the 17th International Conference of the International Society for Scientometrics and Informetrics (ISSI)*. Rome, Italy.

Universidade da Coruña. (s.f.). Universidad de A Coruña. Recuperado de <https://www.udc.es/>

Universidade de Santiago de Compostela. (s.f.). Universidad de Santiago de Compostela. Recuperado de <https://www.usc.es/>

Universidade de Vigo. (s.f.). Universidad de Vigo. Recuperado de <https://www.uvigo.gal/>

Van Eck, N. J., & Waltman, L. (s.f.). VOSviewer. Recuperado de <https://www.vosviewer.com/>

Xunta de Galicia. (s.f.). Educación Universitaria. Recuperado de <https://www.edu.xunta.gal/portal/enseanzas-universitarias>