

# Navigating the Research Landscape: Effective Information Management Strategies

1st Future Research Symposium in South Africa, Johannesburg, 13 March 2024

Research and Innovation Division | Dale Tower | [dalet@uj.ac.za](mailto:dalet@uj.ac.za)



UNIVERSITY  
OF  
JOHANNESBURG

# How UJ uses big data in support of evidence-based research decision making: a practitioner's perspective ...

- Research Information Management Systems (RIMS) have leveraged “big data” and have made unprecedented volumes of research information available to us.
- In this context, “big data” refers to the publication and citation databases available from Elsevier and others, institutional RIMS, and the wealth of **productivity, performance, collaboration, impact** and **visibility** metrics that we can derive from the research that is stored and/or indexed in these systems.
- Today I'll be sharing, from a very hands on, *practitioners' perspective*:
  - How the UJ Research Office is using some of this big research data to help **support informed**, and **evidence-based** research decision making.
  - What we have in the pipeline to help **inform the world** about UJ's research and researchers.



# Agenda

- **Bibliometric profiles**
  - To support decisions relating to hiring, promoting and retaining scholars.
- **Bibliometric analyses**
  - To provide deep insight into the research activities and quantitative performance of our academic departments, research centres and institutes.
- **Pure**
  - Share the platform we are putting in place to expose UJ's research to the world.
  - A hands-on example of we are already using Pure to identify publications for our national assessment that might have otherwise been overlooked, allowing us to offset Pure's cost and improve/justify our return on investment.



# Disclaimer!



- UJ does not exist in isolation.
- Like others, we are aware of and are affected by the advancing “**metric tide**”, the pressure to “**publish or perish**” and the obvious and unanticipated consequences of research metrics.
- We know about the **methodological flaws in the ranking systems**, the **paper mills**, **salami slicing**, **plagiarism**, **predatory journals**, **coercive authorship** and many other perverse practices.
- We understand that the **concepts of “quality”, “impact” and “performance”** cannot be reduced to a few numbers.

- However: we believe there is a role for **responsible** quantitative research evaluation, **as long as:**
  - 1. There is an understanding that publication metrics are just another form of evidence.**
  - 2. Metrics are not used reductively but instead supplement & enhance, rather than replace qualitative evaluation.**





# Bibliometric Profiles – what are they?

- Proprietary document, approved by UJ's SenEx, comprising a focused, citation-based overview of a scholar's recent research activities.
- Data is sourced from Google Scholar, Scopus, SciVal and the current UJ RIMS.
- It comprises a basket of **well documented & internationally accepted**\* research metrics.
- The profile is compiled by the Research Office whenever UJ considers scholars for:
  - **Recruitment**
  - **Promotion**
  - **Retention**
  - **Funding and certain internal awards**

\* See the Snowball Metric Initiative (<https://snowballmetrics.com>) for more on the responsible use of metrics in research evaluation.



# Bibliometric Profiles – what are they NOT?

The profile is **NOT USED TO MAKE A FINAL DECISION**, rather it comprises **evidence** that feeds into a **decision-making process** that **gives UJ the best chance** of recruiting, promoting, retaining , funding and awarding scholars who:

- Are active in research areas that align with UJ's the strategic objectives (and/or those of the unit)
- Are consistently productive.
- Have proven track records of establishing & maintaining diverse international collaboration networks.
- Produce research that is highly visible.
- Have proven track records of publishing in leading journals in their fields & producing highly cited work.





# BIBLIOMETRIC PROFILE



# ELSEVIER SCOPUS

Scholar	Frederick J. Raal
Scopus ID	7003901975
	0000-0002-9170-7938

~20 Year Review: Key SciVal Metrics Since ~1996			
Publications	Citations	Citations/Pub	h-index
270	24,272	89.9	59

Total Scopus Publications	Overall h-index
287	60

% Publications by Institution	% Publications by Subject Area	% Publications by Scopus Source	Scopus Source	Title	#Publications	#Citations	SNIP Impact
			Journal of Clinical Lipidology		9	53	1.05
			South African Medical Journal		7	51	0.62
			European Heart Journal		7	1,159	6.75
			Atherosclerosis		6	358	1.37
			New England Journal of Medicine		5	1,546	17.19
			The Lancet		5	200	25.79
			Journal of the American College of Cardiology		5	330	6.42
			Current Opinion in Lipidology		5	49	0.89
			Cardiovascular Journal of Africa		4	17	0.43
			PLoS ONE		3	10	1.25

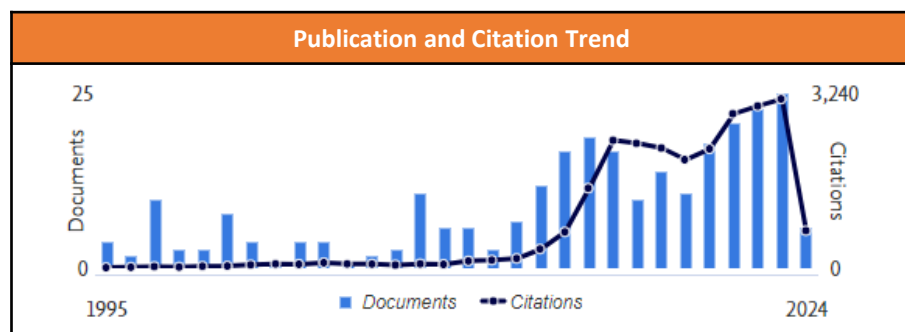
% International co-authorship over the past 5 complete years	% Publications in the top 10% most cited over the past 5 complete years	% Pubs in the top 10% highest impact journals (SNIP) over past 5 complete years
Raal, Frederick Johan: <b>66.7%</b>	Raal, Frederick Johan: <b>26.5%</b>	Raal, Frederick Johan: <b>34.8%</b>

**Field-Weighted Citation Impact (FwCI).** What is the field-specific academic impact of the scholar's publications?

5 Year FwCI	<b>4.18</b>	318% more than average	The <b>FwCI</b> is the ratio of <u>citations</u> received by the scholar relative to the <u>Scopus world average</u> for publications in the same subject field/s, publication type/s and publication year/s. FwCI is an indicator of <b>citation</b> or <b>academic impact</b> .
Since 1996	<b>5.75</b>	475% more than average	

**Field-weighted View Impact (FwVI).** How visible are the scholar's publications to the academic community?

5 Year FwVI	<b>1.79</b>	79% more than average	The <b>FwVI</b> is the ratio of <u>publication views</u> received by the scholar relative to the <u>Scopus world average</u> for publications in the same subject field/s, publication type/s and publication year/s. Can also be seen as an indicator of <b>future citation potential</b> .
Since 1996	<b>2.54</b>	154% more than average	





# Bibliometric Profiles

**NB:** a bibliometric profile **DOES NOT TELL YOU ANYTHING** about scholar's:

- Peer's opinions of their work.
- Ability to raise research funding.
- Teaching and learning activities.
- Community engagement.
- Student supervision.
- Contributions to intellectual property & commercialization.
- Ratings, prestigious awards ... among others.

- These activities remain the **primary considerations** when evaluating a scholar's academic achievements.
- The bibliometric profile provides **additional, objective, supplemental evidence-based metrics** to **support** a well-rounded and comprehensive evaluation of a scholar's research activities.



# Bibliometric Profiles (message to scholars)

- Research metrics may be accessed, and research activities may be evaluated without your knowledge!
  - You may already have public research profiles (Google Scholar, Scopus, WoS) depending on where you publish.
  - As your profiles are public, organisations do not need your permission to access them and may do so without knowledge.
- Scholars
  - It is **CRITICAL** that you take ownership of sources of metrics (Google Scholar, Scopus, Clarivate, etc.)
  - Make sure that at all times your profiles are:
    - ❖ Accurate
    - ❖ Up-to-date
    - ❖ Complete





---

# Bibliometric Analyses

# Bibliometric Analyses – what are they?

- Build on and extend the concept of bibliometric profiles.
- Comprise the **aggregated metrics** of the individual scholars in an entity (i.e. a department, research centre, etc.) as a **proxy for the research activities** or “performance” of the entity.
- Research Office performs bibliometric analyses whenever a Dean/Vice Dean, Head of Department, Director of Centre/Institute, etc., engage in:
  - Strategic planning sessions for their entities.
  - Compiling annual reports (including selected insights/key outcomes).
  - Annual or periodic reviews (sustainability, continued funding, etc.).
  - Performance contracting (especially to establish baseline performance).
  - Benchmarking with similar internal or external entities.



- Intention is to provide **executives with insights** (which otherwise may not be visible or easily attainable) into the **publication & citation-based activities and performance** of the entities, such as:
  - The key **research fields, strengths** and **weaknesses** of the entity.
  - Trends in key metrics (**productivity, scientific impact, visibility, collaboration**, etc.).
  - The extent of publication in **leading journals** resulting in **highly visible and cited** research.
  - Contributions of the entity's publications to **subsidy income**.
  - Performance in terms of some of the **ranking metrics and indicators**.
  - **Benchmarking**.
  - **Modelling** the potential impact of recruiting (or losing) key scholars.

## Bibliometric Analyses – what are they NOT?

- Are **NOT USED TO MAKE FINAL DECISIONS** pertaining to entities and their performance, sustainability, funding, renewability, staff recruitment, retention, promotion, etc.



# Chemical Sciences - Overview

## Summary metrics

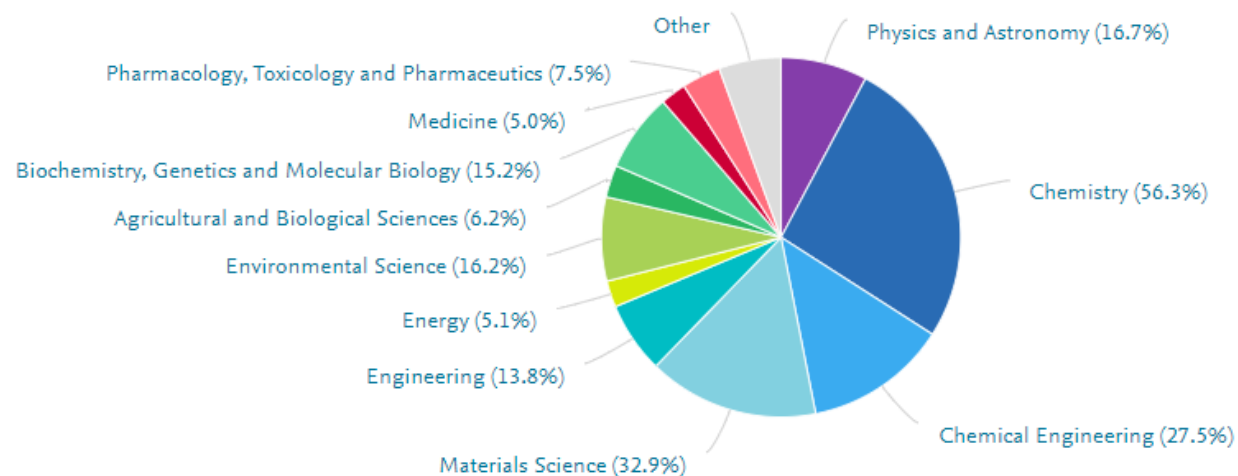


→ **Field-Weighted Citation Impact** is a field-normalised indicator of scientific (citation-based) impact, measured by the number of times that an entity's publications are cited more (or less) than the Scopus world average for other publications **of the same type and in the same field.**

Chemical Sciences publications were cited 1.45 times (or 45% more than) the global Scopus average when compared to similar publications in similar sources in the Scopus database over the review period.

## Publication share by Subject Area

Segment size represents relative publication share per Subject Area. Note that a publication can be mapped to multiple Subject Areas.



# Chemical Sciences – Key performance indicators

## Performance indicators

### Outputs in Top Citation Percentiles

Publications in top 10% most cited worldwide



Chemical Sciences:  
**14.0%**

### Publications in Top Journal Percentiles

Publications in top 10% journals by SNIP



Chemical Sciences:  
**4.3%**

### International Collaboration

Publications co-authored with researchers in other countries/regions



Chemical Sciences:  
**38.9%**

### Academic-Corporate Collaboration

Publications with both academic and corporate affiliations

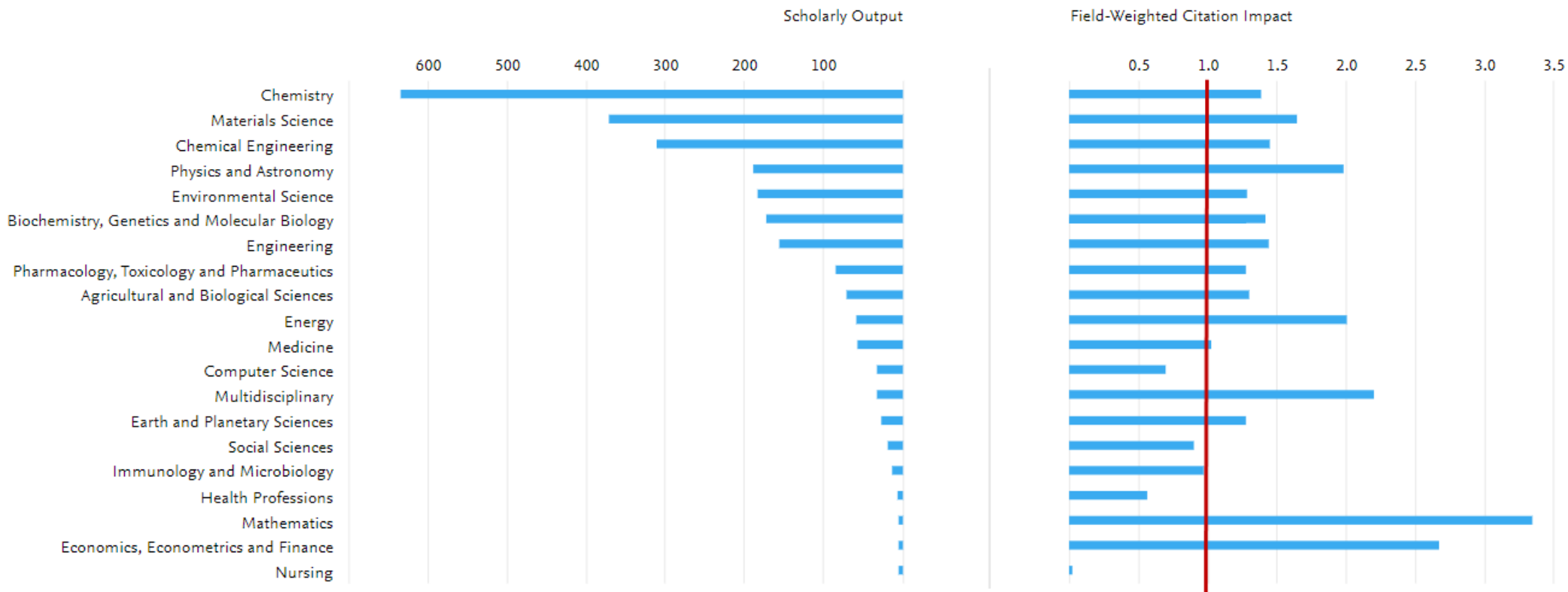


Chemical Sciences:  
**1.1%**



# Chemical Sciences – Research strengths & weaknesses

Publications by Elsevier's “All Science Journal Classification (ASJC) system”



The chart above indicates the fields (as determined by the relevant journal classifications) in which Chemical Sciences is most productive (number of publications) and the Field-weighted Citation Impact (FwCI) of the publications in each field. A FwCI value greater than 1.0 indicates greater scientific (citation-based) impact compared to other publications in the same field in the global Scopus database. A FwCI of less than 1.0 indicates lower scientific impact compared to other similar publications.

Publications classified as “**Mathematics**” have the greatest impact of 3.35 times (or 235% more than average) while publications classified as “**Nursing**” have the lowest impact (0.00). Note that publications are classified according to the journals in which they are published. This can be confusing when publications appear to have been classified in unrelated or even incorrect fields. See slide 7 for more details on the classification of publications.

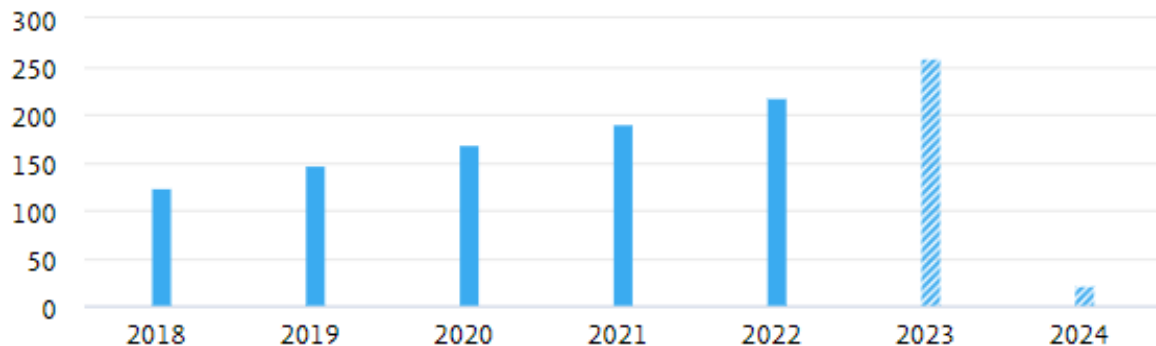
Chemical Sciences achieved above-average scientific impact (in terms of the FwCI indicator) in 15 out of the 20 (75%) ASJC subject fields in which it published during the review period.





# Department – Quantity vs Impact

## Scholarly Output

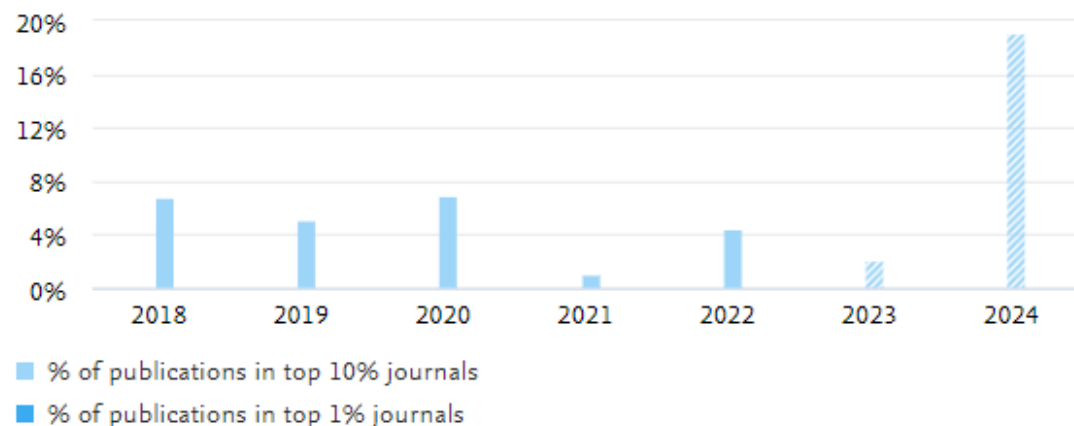


1,127

number of publications by researchers of Chemical Sciences

## Publications in Top Journal Percentiles

Share of publications of Chemical Sciences that are in the top journals by SNIP



46 (4.3%)

number of publications in the top 10% journals by SNIP

**Source-normalised Impact per Paper (SNIP)** is a **source-normalised impact factor** which accounts for the average number of times publications in a Scopus source are cited relative to the average number of times publications in other sources **in the same fields** are cited.

For example, 46 (or 4.3%) of Chemical Science's publications were published in the top 10% of sources in the same publication fields. A source is a publication type (e.g. journal article, conference proceeding, book, etc.).



# Chemical Sciences – Opportunities to improve scientific impact

Top Scopus sources by volume of publications per source, and the Source-normalised impact per paper (impact factor of the source). A Scopus source can be a book, chapter, conference proceeding or journal in which outputs are published.

- ▲ Increasing publications
- ▼ Decreasing publications

Scopus Source	Scholarly Output	Citations	Researchers	SNIP 2022
ACS Omega	24 ▲	424	12	0.94
RSC Advances	22 ▲	412	16	0.87
Journal of Environmental Chemical Engineering	17 ▲	442	11	1.29
Scientific Reports	16 ▲	232	10	1.31
Catalysts	16 ▲	144	11	0.86
Nano-Structures and Nano-Objects	15	276	4	1.02
New Journal of Chemistry	14	367	13	0.68
Molecules	14 ▲	376	10	1.17
ChemistrySelect	14 ▼	80	8	0.49
Journal of Molecular Structure	13 ▲	201	7	0.95

**Source-normalised Impact per Paper (SNIP)** is an impact factor that allows the citation characteristics of publications in different fields to be compared.

A SNIP value greater than 1.00 indicates a **“high(er) impact”** publication source where outputs published in the source are on average cited more than the Scopus average when compared to other outputs in similar fields/disciplines over the same period.

A SNIP value of 1.31, for example, means outputs in that source were cited 1.31 times (or 31% more) on average than similar outputs. A value of 0.49 means that outputs in that source were cited 0.49 times (or 51% less) on average than similar outputs over the period.

Chemical Sciences published most frequently in **“ACS Omega”** which has a SNIP of 0.94, indicating that outputs in this source are cited on average **0.94 times (or 6% less) than other sources in the same subject field.**

Of the top 10 sources, 4 (40%) had SNIP impact factors higher than the Scopus average for sources in similar fields/disciplines. Chemical Sciences may want to consider the possibility (if applicable or possible) of identifying/publishing in higher-impact sources as there is some evidence that this may increase Chemical Sciences’ (and UJ’s) citation rate and other citation-based metrics which play a role in institutional rankings.



# Chemical Sciences – Collaboration analysis

## Geographical Collaboration

International, national and institutional collaboration by Chemical Sciences in the selected year range.



Metric		Scholarly Output	Citations	Citations per Publication	Field-Weighted Citation Impact
International collaboration	38.9%	435	5,592	12.9	2.20
Only national collaboration	27.7%	310	3,969	12.8	1.11
Only institutional collaboration	32.7%	366	5,051	13.8	0.87
Single authorship (no collaboration)	0.7%	8	8	1.0	0.18

Internationally co-authored publications were cited 2.20 times (or 120% **more** than) the Scopus average when compared to similar publications in the same field, of the same type and over the same period.

Publications with **single authorship (no collaboration)** were cited 0.18 times (or 82% **less** than) the Scopus average when compared to similar publications in the same field, of the same type and over the same period.

## Academic-Corporate Collaboration

Academic-corporate collaboration by Chemical Sciences in the selected year range.



Metric		Scholarly Output	Citations	Citations per Publication	Field-Weighted Citation Impact
Academic-corporate collaboration	1.1%	13	40	3.1	0.52
No academic-corporate collaboration	98.8%	1,114	14,802	13.3	1.46

Publications with **corporate (industry) collaboration** were cited 0.52 times (or 48% **less** than) the Scopus average when compared to similar publications in the same field, of the same type and over the same period.














# Chemical Sciences – Insight into the effectiveness and impact of partnerships

▲ Increasing collaboration

▼ Decreasing collaboration

## Top collaborating Institutions

by number of publications co-authored with Chemical Sciences

	Institution	Co-authored publications	Citations received for co-authored publications	Co-authors	Field-Weighted Citation Impact
1.	  University of Johannesburg	1,119 ▲	14,620	534 ▲	1.45
2.	 University of Calabar	90 ▲	1,320	114 ▲	5.97
3.	 MINTEK	79 ▲	1,361	25 ▲	1.22
4.	 University of South Africa	62 ▲	618	59 ▲	1.05
5.	 University of the Witwatersrand	59 ▲	473	45 ▲	1.04
6.	 National Research Foundation	57 ▲	469	27 ▲	1.42
7.	 Council for Scientific and Industrial Research	56 ▼	934	27 ▼	1.07
8.	 University of Pretoria	39 ▲	191	22 ▲	0.65
9.	 Université de Yaoundé I	35 ▲	177	55 ▲	0.58
10.	 Mahatma Gandhi University, Kottayam	30 ▼	351	22 ▼	1.22

Apart from collaboration within UJ, Chemical Sciences' collaboration with University of Calabar was the most beneficial in terms of the **volume of co-authored papers**.

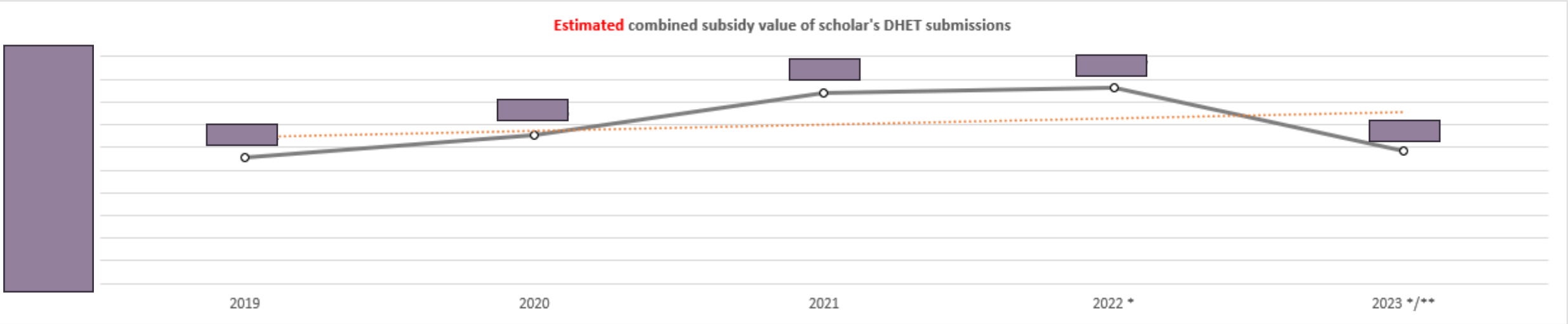
Collaboration with University of Calabar yielded the **highest scientific impact (5.97)**, with 497% more citations than similar publications.

Collaboration with the Université de Yaoundé yielded the **lowest scientific impact (0.58)**, with 42% less citations than similar publications.



# Chemical Sciences –Contribution to publication subsidy income

- Publication subsidy income is just once source of income which may or may not be significant depending on the entity.
- But understanding how an entity has contributed to the institution’s publication subsidy income in the past may be useful metric to include in more comprehensive financial models for future planning.



Average **estimated\*** combined subsidy value of scholar's DHET submissions per year. Only publications submitted to the Research Office are included.

# Bibliometric Analyses – Scenario Modeling – Recruitment Impact

## Before

### Summary metrics

366  
Scholarly Output

16 ▼  
Researchers

4.72  
Field-Weighted Citation Impact

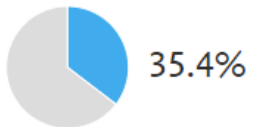
70.5% All Open Access

19,772  
Citation Count

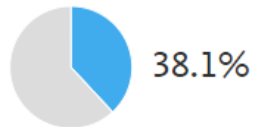
54.0  
Citations per Publication

### Performance indicators

Outputs in Top Citation Percentiles  
Publications in top 10% most cited worldwide

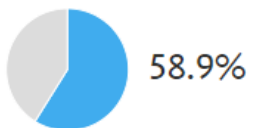


Publications in Top Journal Percentiles  
Publications in top 10% journals



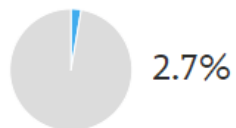
International Collaboration

Publications co-authored with researchers in other countries/regions



Academic-Corporate Collaboration

Publications with both academic and corpo



## After

### Summary metrics

443  
Scholarly Output

17 ▼  
Researchers

4.90  
Field-Weighted Citation Impact

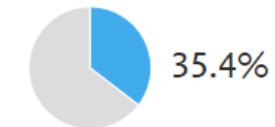
67.3% All Open Access

21,328  
Citation Count

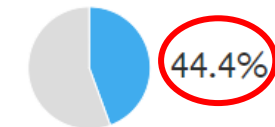
48.1  
Citations per Publication

### Performance indicators

Outputs in Top Citation Percentiles  
Publications in top 10% most cited worldwide

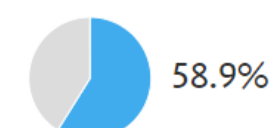


Publications in Top Journal Percentiles  
Publications in top 10% journals



International Collaboration

Publications co-authored with researchers in other countries/regions



Academic-Corporate Collaboration

Publications with both academic and corporate affiliations



# Bibliometric Analyses

**NB:** bibliometric analyses **DO NOT TELL YOU ANYTHING** about entities’:

- Peer’s opinions/review.
  - Fund raising abilities.
  - Teaching and learning activities.
  - Community engagement.
  - Student supervision.
  - Contributions to intellectual property & commercialization.
  - Prestigious awards ... among others.
- These activities remain the **primary considerations** when evaluating entities achievements.
  - The bibliometric analysis provides **additional, objective, supplemental evidence-based metrics** to **support** a well-rounded and comprehensive evaluation of an entity’s research activities.





---

# The Pure Research Portal





# Many stakeholders are interested in UJ's research

*UJ staff | Members of the public | Members of the press | Government agencies | Funding agencies  
Scholars from other institutions | Prospective students | Parents | Potential commercial partners*

- And they want to know:
  - What is UJ's primary research focus and what are UJ's research strengths?
  - Who are UJ's experts in specific research areas?
  - How has UJ responded to various local, regional and international events such as the Covid 19 pandemic, the 4<sup>th</sup> industrial revolution, the rise of AI, etc.
  - What collaboration opportunities exist for external scholars and experts?
  - How do we highlight prestigious awards and recognitions?
  - What trends are evident in terms of the quantity and impact of UJ's research?
  - What resources does UJ have to support research (Centres, Institutes, equipment, facilities, etc.)
- How to respond?



# The problem

UJ's research information is currently stored in **different systems** which **limited** or **no integration**, such as:

- The Research Office's InfoEd **RIMS** - which only store subsets of UJ's research for specific purposes.
- The Library's **Institutional Repository** – which focuses on theses and dissertations and some publications.
- Proprietary databases such as **Scopus** and **Web of Science** - which only index certain publications.
- Answers may not always be possible as these systems are **access controlled**, requires **subscriptions**, requires **internal authorisation** to access, and often **specialized skills** to extract and interpret the results.



# The solution: Pure Portal



- Aggregates multiple sources of research information in a controlled manner into a single curated searchable, reportable resource.
- By exposing UJ's research in a public portal, the **quality**, **quantity** and **impact** of UJ's research, as well as UJ's **research capacity**, **capabilities** and **facilities** will be evident to stakeholders.
- Pure addresses a requirement that cannot be met by the other disparate data sources:
  - ❖ Allows UJ's **research activities to be “mapped” to UJ's internal organizational hierarchy** of Faculties, Departments, and Research Centres.
    - This allows insight into the productivity, quality and impact of UJ's research at a granular level.
    - *Answer questions at an institution, faculty, department, centre, scholar level.*
    - Enables **targeted strategic initiatives & interventions, customized tracking & monitoring of key performance areas and benchmarking** at a level beyond just comparing UJ to other institutions.



# Early wins: using Pure and Scopus data to enhance ROI

297 DHET units for Scopus-indexed articles worth approx. R38 million identified that otherwise may have gone unclaimed.



## Units per capture source

**Report Year**

2022 | 2023 | 2024

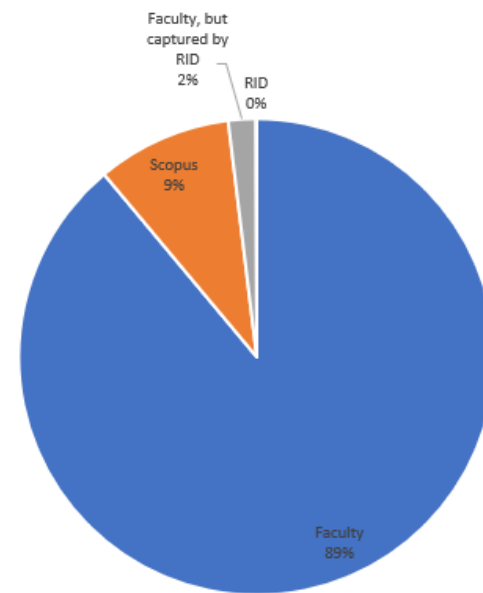
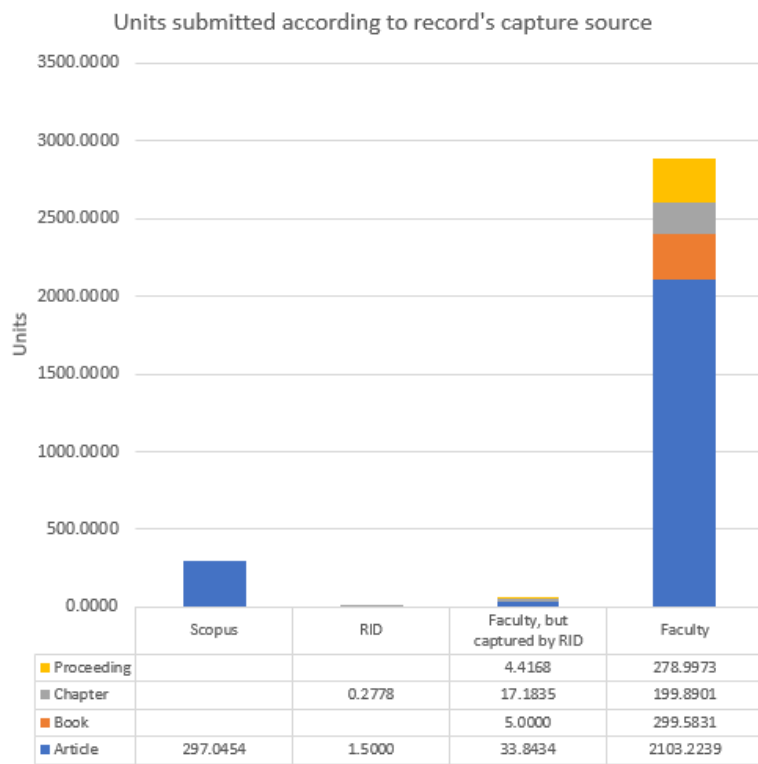
**College/Faculty**

- College of Business and Economics
- Faculty of Art, Design and Architecture
- Faculty of Education
- Faculty of Engineering and the Built Environment
- Faculty of Health Sciences
- Faculty of Humanities
- Faculty of Law
- Faculty of Science
- Johannesburg Business School
- Non-Academic

**Status**

- Approved
- Disapproved
- Duplicate

Units	Article	Book	Chapter	Proceeding	Total
Scopus	297.0454				297.0454
RID	1.5000		0.2778		1.7778
Faculty, but captured by RID	33.8434	5.0000	17.1835	4.4168	60.4437
Faculty	2103.2239	299.5831	199.8901	278.9973	2881.6944
<b>Total</b>	<b>2435.6127</b>	<b>304.5831</b>	<b>217.3514</b>	<b>283.4141</b>	<b>3240.9613</b>



Source: InfoEd RIMS / RID Dashboard

### Key to Capture Sources

#### Faculty:

Publications captured and submitted to the Research Office by the college/faculty.

#### Faculty, but captured by RID:

Publications identified by the college/faculty, but captured by the Research Office. These are usually publications submitted by the college/faculty with insufficient supporting documentation (that had to be sourced by the Research Office in order to ensure compliance with the DHET policy), or publications submitted after the annual January 31st submission deadline.

#### RID:

Publications identified and submitted by the Research office on behalf of the college/faculty. **These publications would otherwise have gone unclaimed.**

#### Scopus:

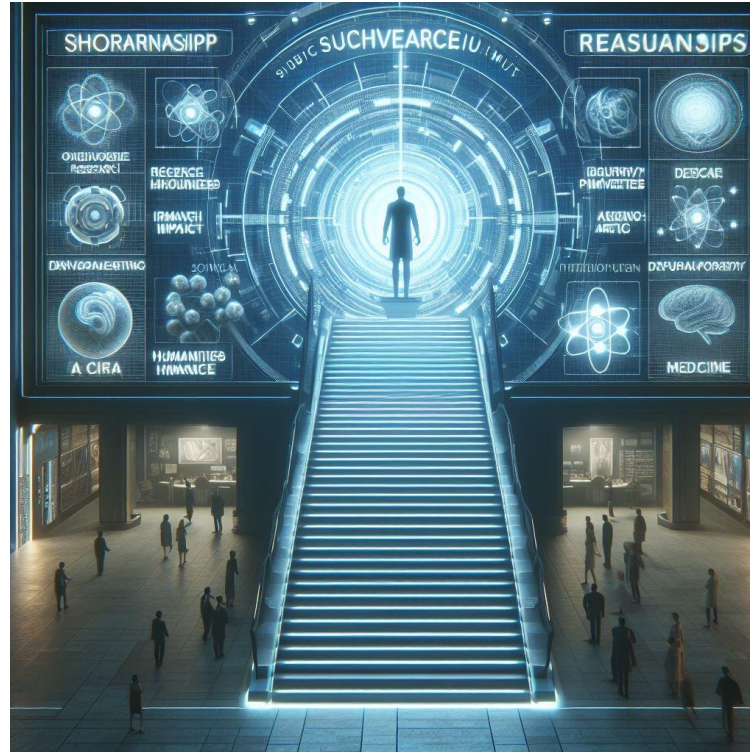
Publications identified and submitted by the Research Office as a result of comparing UJ's publications in the Scopus database to the publications submitted by the college/faculty. **These publications may otherwise have gone unclaimed.**

#### (Blank):

Still to be classified during the RID review process.

# Pure demonstration

(if there's time, or on the sidelines)



UJ Pure Research Portal

[www.pure.uj.ac.za](http://www.pure.uj.ac.za)



# 1st Future Research Symposium in South Africa 13 March 2024

## Navigating the Research Landscape: Effective Information Management Strategies

“Big data in support of evidence-based  
research decision making:  
a practitioner’s perspective”

Dale Towert | [dalet@uj.ac.za](mailto:dalet@uj.ac.za)  
Director: Research Intelligence

University of Johannesburg  
Research & Innovation Division

# Thank you | Questions



*“AI, generate a steampunk-style photo of me working in my office on big data and quantitative research evaluation techniques to examine and understand research productivity and impact.”*

