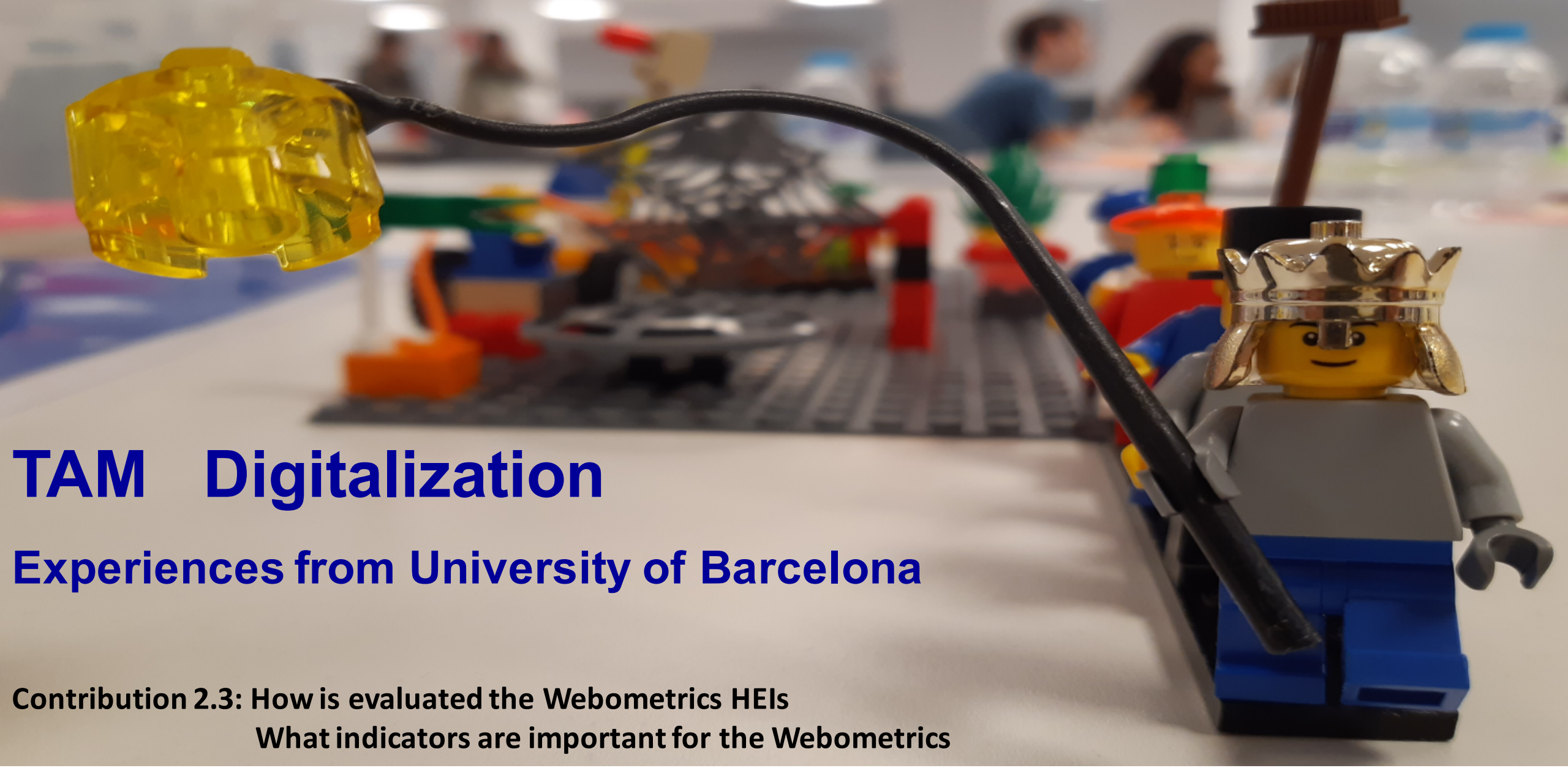


TAM Digitalization

Experiences from University of Barcelona

Contribution 2.3: How is evaluated the Webometrics HEIs
What indicators are important for the Webometrics

UNIVERSITAT DE BARCELONA

SPHERE
Support and Promotion for Higher Education Reform Experts

Erasmus+

About Webometrics Project

The WEBOMETRICS ranking of world universities is produced by Cybermetrics Lab CCHS - a unit of the Spanish National Research Council CSIC -, the main public research body in Spain.

It offers ranking according to their **web presence** (an assessment of the **scholarly contents**, **visibility,** and **impact** of universities on the web)

The term Webometrics

According to the definition by Björneborn and Ingwersen webometrics represents a **joint** (synergy) activity, i.e., application of other two approaches in one element known as **bibliometrics** and **informetrics**, **for researching of the web, its information resources, structure, and technologies.**

The name "webometrics" was defined in 1997, and it was created and established by Tomas Almind and Peter Ingwersen with an intention to show that informetric analysis **can be applied to the web as an important source for measuring values (weight/sense) of documents and information**.

Webometrics Main Hypothesis

The central hypothesis of this approach is that web presence is a reliable indicator of the global performance and prestige of the unversities and as such, it is an indirect way to measure all the universities missions: teaching, research and transfer.

Webometrics Project Intention

It intend to motivate both institutions and students to have a web presence that reflect accurately their activities.

If the web performance of an institution is below the expected position according to their academic excellence, academic authorities should reconsider their web policy, promoting substantial increases of the volume and quality of their electronic publications.

Webometrics is…

- Based on a composite indicator that includes both the volume of the web contents and the visibility and impact of web publications according to the number of external links that they recevce.
- Published since 2004
- Published twice a year (data is collected during the first weeks of January and July) for being public at the end of both months.
- Built from a DB of over 24000 HEI - Higher Education Institutions – are shown in the main rankin and more are covered in regional lists.
- Skewed towards websites which maintain web publications in English on their websites. So North American Univ are relatively common in the top 200, while small and medium size biomedical institutions and German, French, Italian and Japanese universities were less common in the top ranks.

  Possible reasons include publishing via independent research councils (n.b.: CNRS, Max Plank, CNR, …) or the large amount of non-englishweb contents, which are less likely to be linked.

Webometrics is…

- Most small universities do not public all their content on their websites, but only publishing on magazines thereby losing out on the ranking.

- Many organizations try to beat the algorithm by falsifying the underlying data but this might constitute a small anomaly.

CSIC Building - Spain -

Webometrics tools for collection of data from the internet

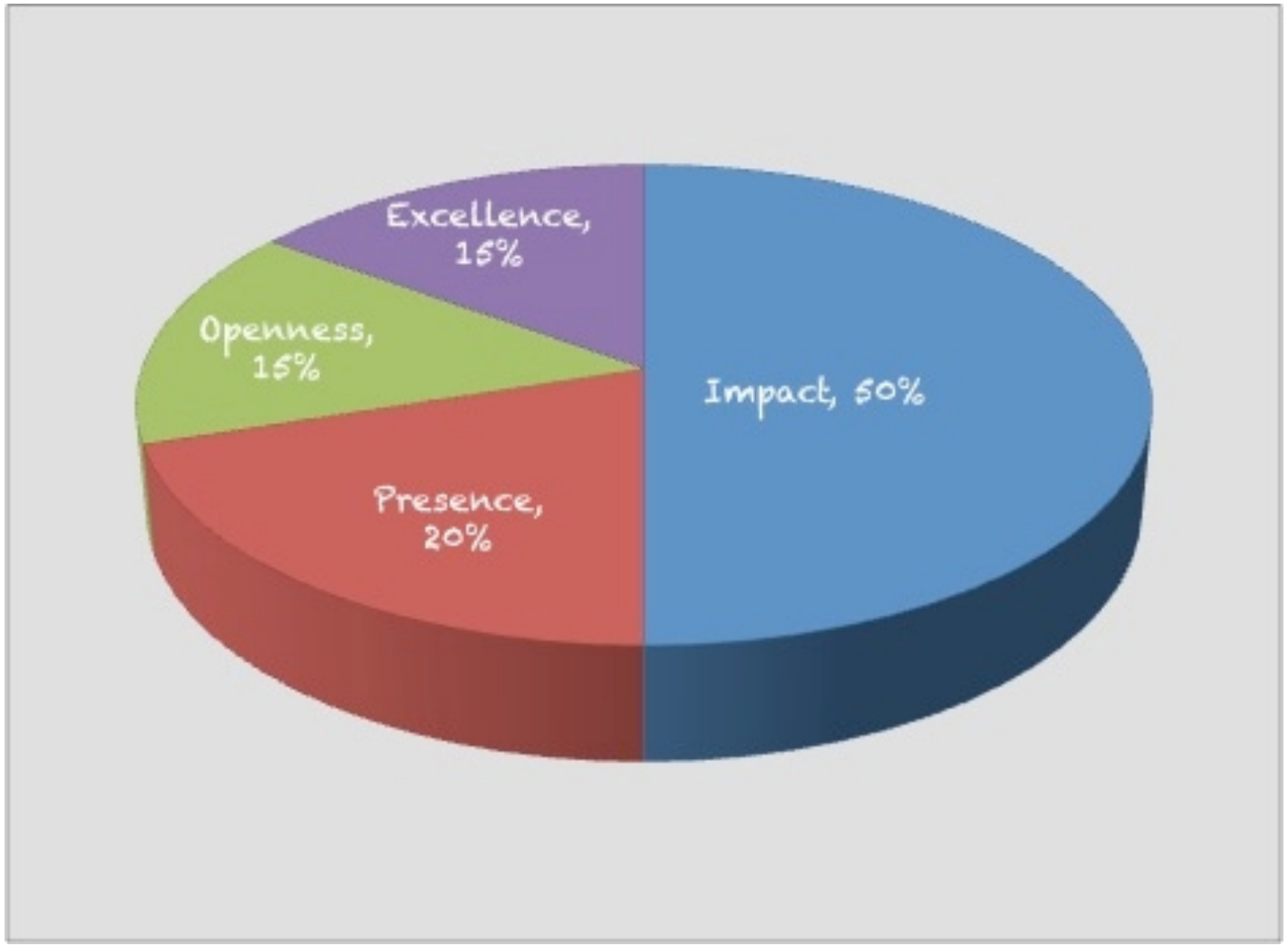
Web tools such as search engines, web crawlers, and webometrics software which are used for collecting data from the web are called Webometrics tools .
The area of research in the field of webometrics can, in a wider sense, be divided into the following segments:

- Analysis of contents of websites
- Analysis of web technologies
- Analysis of application of web contents
- Analysis of structure of web links

Basically, there are two categories of sources of information which can be used in the research of webometrics:

- Commercial web search engines
- Personal web crawlers

Criteria that Webometrics follow



Criteria that Webometrics follow

Indicator	Meaning	Coverage	Source	Weight
Impact	Number of backlinks Number of back domains	Current (historical)	Majestic SEO ahrefs	50%
Presence	Number of webpages (all)	Current	Google	20%
Openness	Number of papers (pdf, doc, docx, ppt)	2007 - 2012	Google Scholar	15%
Excellence	Numbers of papers in the 10% top cited	2003 - 2010	Scimago	15%

Criteria that Webometrics follow

Impact

Refers to the quality of the contents

It is evaluated through a "virtual referendum", counting all the external unlinks that the University web domain receives from third parties. Those links help recognize the institutional prestige, the academic performance, the value of the information and the usefulness of the services as introduced in the webpages according to the criteria of millions of web editors from all over the world.

.The link popularity and diversity are important for a strong impact.

Criteria that Webometrics follow

Presence

Represents the total number of webpages hosted in the main web domain of the university as indexed by Google.

It takes into consideration all formats of webpages recognized by Google (static and dynamic pages and other rich files).

The contribution of everybody in the organization is important for a strong presence

Having additional domains or alternative central ones for foreign languages and/or marketing purposes is penalized.

Criteria that Webometrics follow

Openess

Stands for the global effort to set up institutional research repositories.

Takes into account the number of rich files (pdf, doc, docx, ppt) published in dedicated websites according to the academic search engine Google Scholar.

The objective is to consider recent publications.

Criteria that Webometrics follow

Excellence

This indicator shows how many academic papers are published in high impact international journals.

Since they consider that using the total number of papers could be misleading, they are restricting the indicator to only excellent publications.

Excellent publications refers to the university scientific output being part of the 10% most cited papers in their respective scientific fields.

Criteria and weights used in the WR indicator

Criteria	Indicator	Sources	Weight
Size	Number of pages (*S*)	Google, Yahoo, Live, Exalead	25%
	Number of rich files (PDF, PPT, DOC, and PS) (*R*)	Google	12.5%
	Number of papers (*Sc*)	Google Scholar	12.5%
Visibility	Number of external links (*V*)	Yahoo, Exalead, Live	50%
Luminosity	Number of external outlinks		
Subdomains	Number of subdomains		
Popularity	Number of visits		

Table 1.
Criteria and weights used in the calculation of the WR indicator [51].

Equation WR - Web Ranking -

$$WR = 2 \times Rank(S) + 1 \times Rank(R) + 1 \times Rank(Sc) + 4 \times Rank(V)$$

S: Number of Pages

R: Number of Rich Files

Sc: Number of Papers

V: Number of External Links

Spain

ranking	World Rank	University	Det.	Presence Rank*	Impact Rank*	Openness Rank*	Excellence Rank*
1	116	Universitat de Barcelona	»	164	185	154	103
2	130	Universitat Politècnica de Catalunya BarcelonaTech	»	22	144	58	278
3	141	(1) Universitat de València	»	163	134	176	227
4	142	Universidad Complutense de Madrid	»	134	140	86	243
5	156	Universitat Autònoma de Barcelona	»	86	276	59	151
6	180	Universidad de Granada	»	144	212	132	252
7	197	Universidad de Sevilla	»	69	218	71	352
8	226	Universidad Politécnica de Madrid	»	517	152	134	409
9	232	Universidad Politécnica de Valencia	»	235	282	159	316
10	255	Universidad Autónoma de Madrid	»	699	327	300	222
11	308	Universidad de Zaragoza	»	590	391	123	343
12	315	Universitat Pompeu Fabra	»	262	277	460	446
13	350	Universidade de Santiago de Compostela	»	51	599	277	398
14	369	Universitat d'Alacant / Universidad de Alicante	»	60	346	192	688

Strategies for sustainable Webometrics Ranking Performance

UNIVERSITAT DE BARCELONA

SPHERE
Support and Promotion for Higher Education Reform Experts

Erasmus+

Uniform resource locator naming: HEI should choose a unique domain name for used by all websites to avoid confusion.

Creation of contents: Effective website is made possible only with the effort of large group of authors and potential users.

.

Conversion of contents: Important resources available in non-electronic formats need to be easily converted to web pages

Interlinking: Providde the ability to hyper connect the information and contents of the websites

Language: Web audience is truly global, and language versions specially in English are mandatory, not only for the main pages but also for selected sections and specially for scientific documents.

Rich and media files: hypertext markup language is the standard format for creating websites although sometime it is advisable to use rich file formats like adobe acrobat pdf or microsoft worf document

Interactive search engine: Institutions web designers should avoid cumbersome navigation menus based on flash, java or JavaScript that block robot access

Popularity and statistics: number of visits is important although it is necessary to monitor origin, distribution and reason for reaching the websites.

Archiving and persistence: Maintaining the old copy of outdated material in the site is mandatory as sometimes relevant information is lost when the webpage is redesigned or updated.

Standards for enriching sites: The use of meaningful titles and descriptive metatags can increase the visibility of the pages. Standards like the Dublin Core can be used to add auhoring info, keywords and other data about the web sites

Open access initiatives: Electronic access to scientific publications and other academic materials are regarded as crucial startegies towards webometrics ranking (Aguilo 2008)

METADATA
Dublin Core® Metadata Initiative
INNOVATION

	DC Element Name	Definition
1.	Title	A name given to the resource.
2.	Creator	An entity primarily responsible for making the resource.
3.	Subject	The topic of the resource.
4.	Description	An account of the resource.
5.	Publisher	An entity responsible for making the resource available.
6.	Contributor	An entity responsible for making contributions to the resource.
7.	Date	A point or period of time associated with an event in the lifecycle of the resource.
8.	Type	The nature or genre of the resource.
9.	Format	The file format, physical medium, or dimensions of the resource.
10.	Identifier	An unambiguous reference to the resource within a given context.
11.	Source	A related resource from which the described resource is derived.
12.	Language	A language of the resource.
13.	Relation	A related resource.
14.	Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
15.	Rights	Information about rights held in and over the resource.

Source: SpringerLink

Extended

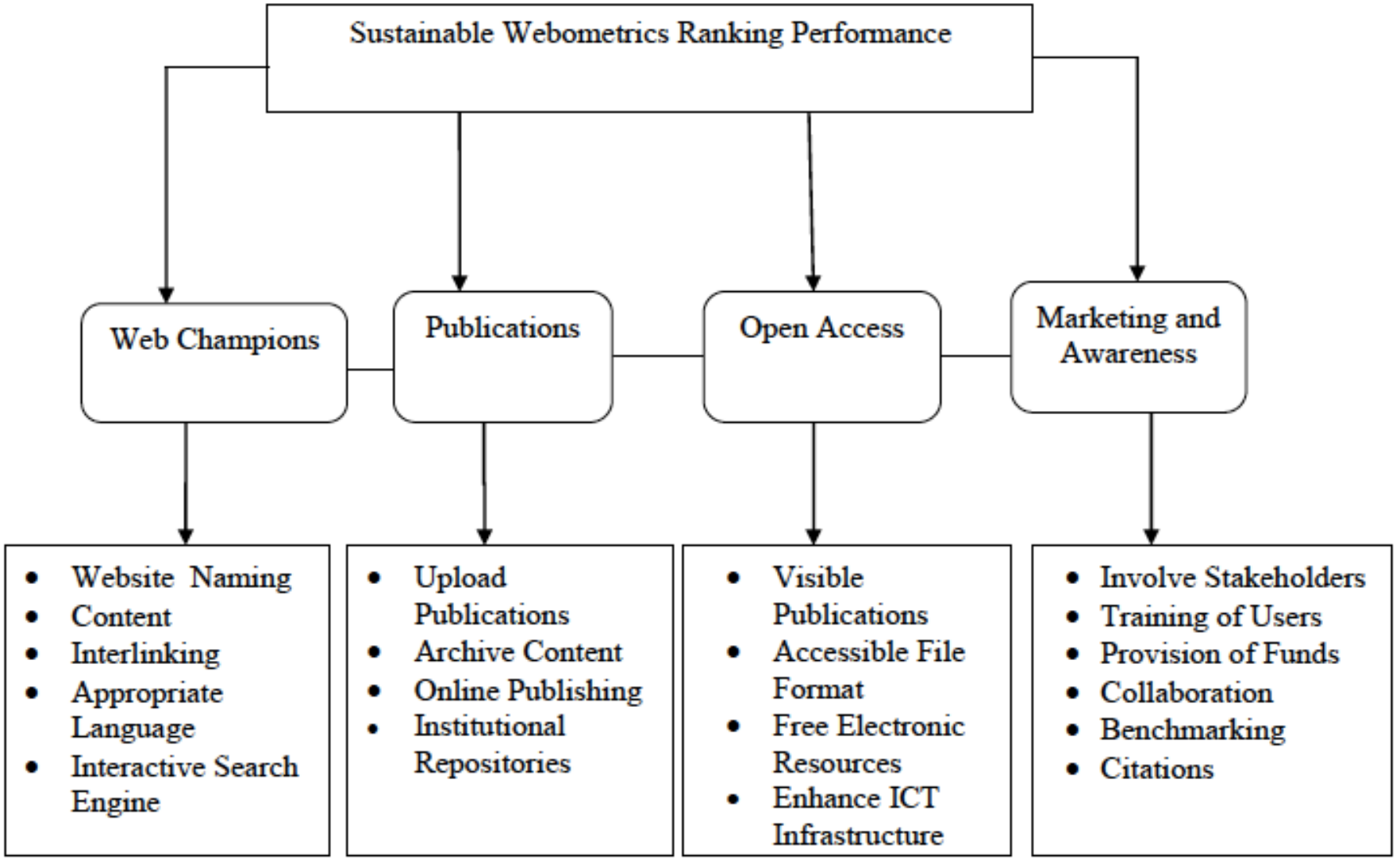


Figure 1: Strategies for Webometrics Ranking (Researcher, 2015)

University of Nebraska Lincoln D.F: Khamala

FORESIGHTING in Digital Mindset

UNIVERSITAT DE BARCELONA

SPHERE
Support and Promotion for Higher Education Reform Experts

Erasmus+

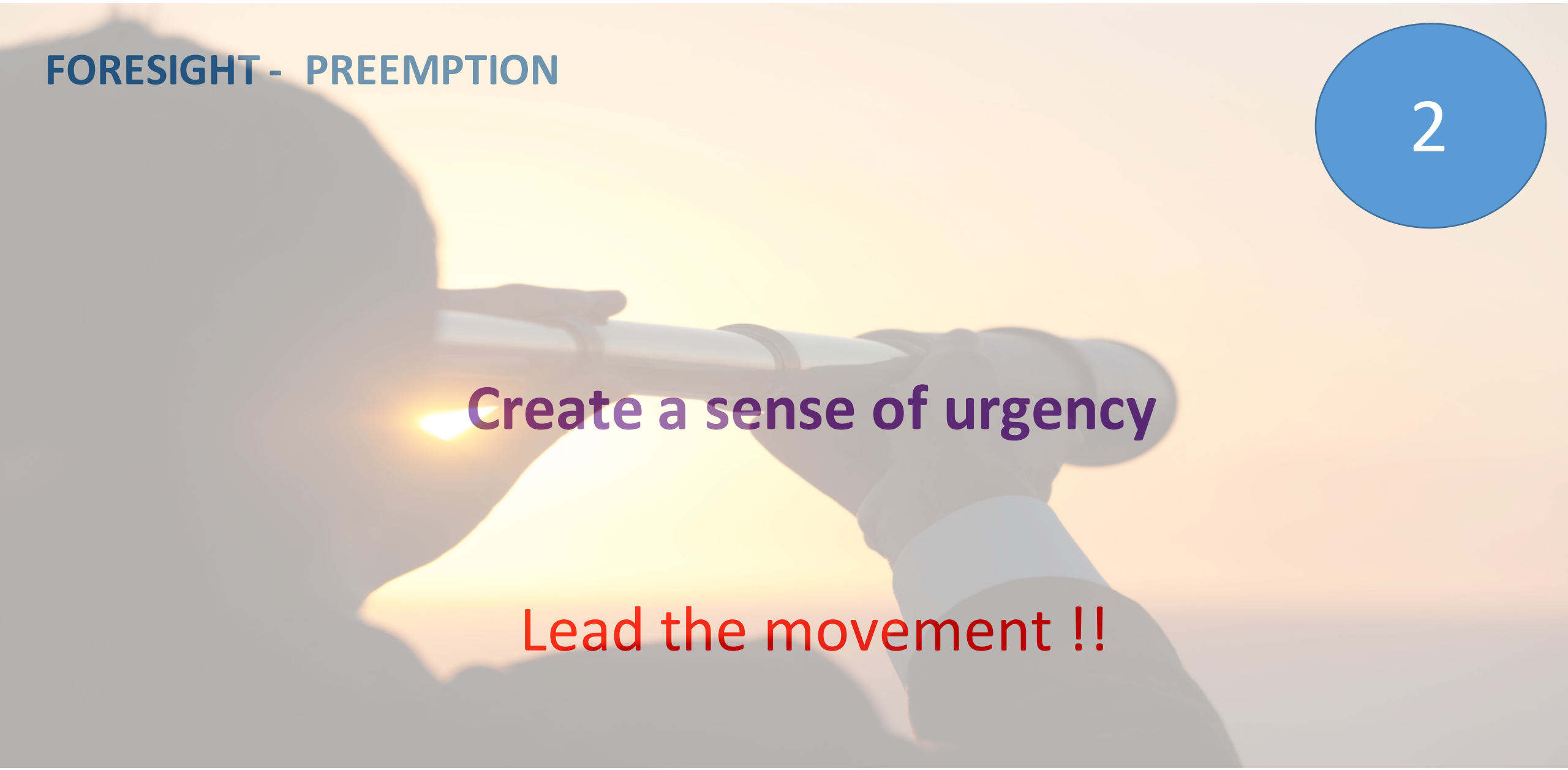
FORESIGHT - PREEMPTION

1

Constantly explore

FORESIGHT - PREEMPTION

2



Create a sense of urgency

Lead the movement !!

FORESIGHT - PREEMPTION

3

Watch out for early-warning signals

Measure it !!!

FORESIGHT - PREEMPTION

4

Create transformation capabilities

Ambidexterity !!!

FORESIGHT - PREEMPTION

5

Control the narrative

Avoid frictions !!!

FORESIGHT - PREEMPTION

6

Choose the right adaptive change

Learn!!!


LIBERTAS PERFVNDET
OMNIA LVCE
THANKS !!!!