The transformation of science systems in the Middle East and North Africa
El Ouahi, J.

Chapter 1

General Introduction

1.1 Introduction

1

The 'global science system' was in the 20th century relatively Anglo-American-focused (Aalbers, 2004). This global science system is becoming more and more focused on China (Marginson, 2022). In terms of the history of science, Delbourgo (2019) proposed the *knowing world,* a geopolitically pluralistic vision of history, to describe our planet as made up of multiple scientific cultures that still resonate in our global current science system. One of these cultures corresponds to the production of scientific knowledge in Islamic societies during the medieval and early modern eras in the Islamic Caliphates and the Ottoman Empire. Popular historian Jim Al-Khalili's *House of Wisdom: How Arabic Science Saved Ancient Wisdom and Gave Us the Renaissance* (2011) provides a rich description of scientific contributions in this specific region from roughly the 8th to the 14th century.

This region is currently commonly referred to as the so-called 'Middle East and North Africa'. As recently dismantled by Lynch (2022), such orientalist cartographic terminology of this area is not firmly rooted in premodern history as it may currently seem. The broad, multicultural Cherifian, Ottoman, and Safavid Empires included the Arab provinces of North Africa, the Levant, Arabia and Central Asia for centuries. There was a natural connection between Mesopotomia and China, between the Horn of Africa and the Arabian Gulf, and North African was well linked to Southern Europe and sub-Saharan Africa by Islamic networks. The French and British colonialism and great-power politics of the 19th and 20th centuries were the more recent sources from which the United States adapted its view of the Middle East and North Africa. The notion that this broad region shared a single culture was shaped by a series of ideological presumptions on the purported exoticism of Arabs, Persians, and Turks, a worldview that was memorably referred to as 'Orientalism' (Said, 2014). Although, I also refer to this specific region as the 'Middle East and North Africa' in this thesis, I remain critical towards this terminology and its history.

As mentioned earlier, there is already a strong history of science in the Middle Eastern and North Africa (MENA) region which was at the center of the global science system in the Middle Ages. During the Islamic Golden Age, spanning roughly from the 8th to the 14th centuries, significant contributions to various fields of science emerged from different knowledge centers in MENA. Islamic scholars of this period made groundbreaking advancements in mathematics, astronomy, medicine, optics, robotics and philosophy among many others. Al-Khawarizmi, Al-Biruni, Ibn Sina, Al-Razi, Ibn Al-Haytham, or Al-Jazari are a few examples of many scholars from the MENA region who made significant contributions to science, which influenced European scholars during the Renaissance and had a lasting impact on several fields. For instance, Al-Khawarizmi is known as the "father of algorithms and algebra," with these terms derived from his own name and the word *al-jabr* used in the title of his renowned book, *Al-kitab al-mukhtasar fi hisab al-jabr wa'l-muqabala* (The Compendious Book on Calculation by Completion and Balancing). Al-Jazari is considered as the "father of robotics" due to his numerous inventions in mechanical engineering and is known to have described in great details his work. Ibn Sina produced *al Qanun fi al-Tibb* (The Canon of Medicine), a comprehensive medical encyclopedia and Al-Razi composed *Al Kitab al-Hawi fi al-tibb* (The Comprehensive Book on Medicine), an extensive 23-volume textbook which served as a primary medical curriculum in European schools until the 14th century.

Speaking of books, it is important to note that these books were written in Arabic, the *lingua franca* of science in that period. Pedersen (2014) provides a history of the Arabic book covering various topics such as: books writing in Arabia before Islam, the Quran, the transmission of books, bookbinding, libraries, and printed books. The House of Wisdom in Baghdad, the Zahiriyya library in Damascus, the library of Fatimid Caliphs in Cairo, the library of Spanish Umayyad Caliphs in Cordoba or the library of University of al-Qarawiyyin in Fez (Morocco) are examples of such libraries hosting hundreds of thousands of scientific books and mansucripts but also serve as research centers of the Islamic Golden Age. Buringh and Van Zanden (2009) provide estimates of manuscripts and printed books in Europe from the 6th through 18th century and highlight the strong Islamic scientific heritage in Spain which dominated the European book production in 1600 (about a third) but then saw a relative decline to around 2% in 1800. Going through a detailed and comprehensive history of science in this dissertation is not the main objective and would still not do justice to the rich contributions of all the various civilizations. But it is still worth mentioning the dynamics and knowledge flows between different regions. For instance, several works of George Saliba (2007, 2008, 2009) clearly highlight the shifts and the dynamics of global science systems throughout history in terms of its changes such as the main language of scientific communication in different periods but also in terms of exchanges of techniques or ideas between regions. These dynamics include connections between the

Greek and Muslim scholars, between China and the Islamic civilization, and later Europe with the making of its renaissance. From a social and history of science perspective, Huff (2017) explores the rise of early modern science in the West and sheds light on necessary structural and cultural changes for the European scientific revolution to occur in comparison to prior scientific civilizations which lacked these factors.

A few centuries later, the current global science system has been expanding at an increasing rate (Marginson, 2020). The world is now experiencing a pluralization of science with a growth of international scientific collaboration (Wagner et al., 2017). When exploring the pluralization of global science, some research studies investigated various regional or national science systems. For instance, several publications covered the emergence of Asian science systems (Dang, 2015; Oldac & Lili, 2021; van der Wende et al., 2020), the post-Soviet regions (Chankseliani et al., 2021) and the Latin American systems (Koljatic & Silva, 2001). One way to track the growth of the research output of the MENA region is through the count of scientific manuscripts published as journal articles and reviews which serve as the primary medium for original academic research. According to the World Bank (2019), MENA consists of the following countries: Algeria, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestine, Qatar, Saudi Arabia (KSA), Syria, Tunisia, the United Arab Emirates (UAE) and Yemen. Pakistan, Afghanistan, and Turkey are also included in this dissertation as they are commonly included in the MENA region (MENAP[1] and MENAT[2]) as neighborhood Muslim-majority countries. Using the Web of Science (WoS) data, the collective publication output in MENA has experienced remarkable growth since 2008. In 2008, a total of 54,561 scientific publications (articles & reviews) were published, while in 2021, that number soared to over 260,000 papers. This represents an absolute growth of about five-fold. This rise led to a significant increase in the relative share of the MENA research output in the world, more than doubling from 4% to 9%. In contrast, as shown in Figure 1.1, Latin America, and the Asia-Pacific region (APAC, Australia & New-Zealand (ANZ) excluding China) experienced a slower growth in their global publication share. The European Union 27 countries and the United States witnessed substantial decreases in their share of the global scientific output.

[1]MENAP: https://www.imf.org/en/Publications/REO/MECA/Issues/2019/10/19/reo-menap-cca-1019#Sum
[2]MENAT : https://en.wikipedia.org/wiki/MENA

It is important to note that the exceptional growth trajectory of China heavily impacted these regions, as it tripled its share of world publications from about 8% in 2008 to over 24% in 2021.

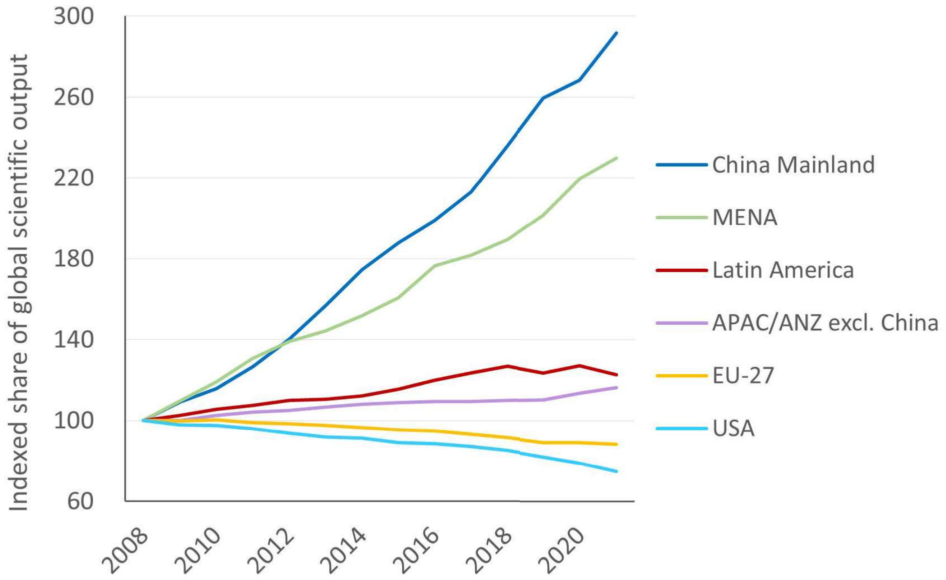

Figure 1.1 Change in publication output as the percentage of all journal articles and reviews indexed globally in the WoS between 2008 and 2021

In summary, the MENA region is becoming a more substantial actor in the current science system. The MENA region has experienced substantial growth in research publication, reflecting its increasing engagement with the global research community through scientific publishing. This growth is not only characterized by an absolute increase in scientific output but also by the region's expanding share of the global publication landscape. The significant role of international collaboration in driving this growth underscores the value of knowledge exchange and cooperation in tackling complex scientific and societal issues (Adams et al., 2021). Economic policies have stressed the role of science in innovation and economic growth throughout the previous few decades (Etzkowitz & Leydesdorff, 2000; Stephan, 2012). In this context, many Middle Eastern and North African economies have made considerable investments in science and technology capacity to promote research and innovation (Schmoch et al., 2016; Shin et al., 2012; Siddiqi et al., 2016). One common goal of the MENA nations is to become so-called ‘knowledge-based economies’ (OECD, 1996). Several studies investigated the recent increase in scientific output in many of these countries as a result of this transformation (Ahmad et al., 2021; Al Marzouqi et al., 2019; Cavacini, 2016; Gul et al., 2015; Sarwar & Hassan, 2015). However, the increase shown in Figure 1.1 may also highlight that some countries have science systems where resources have been recently under strong pressure to publish considerably. As a result, researchers strive to secure publication in indexed journals. These ‘publish or perish’ systems may lead to very minor incremental scientific contributions or superficial publications but may also raise questions about research integrity (Ataie-Ashtiani, 2018; Qiu, 2010).

The transformation of science systems in this specific region is largely under studied. This reflects a significant gap in our knowledge of how science evolves in various parts of the world. Geographically and economically, the MENA area is situated at a crossroads between established and emerging economies. With a strong history of commitment to science, MENA already shows signs of rapid scientific expansion and is an ideal location for studying the evolution of science systems in countries with different levels of development. Despite its closeness in terms of geography, this region is immensely diverse in terms of languages, economy, culture, as well as history. These aspects make the region exceptionally diverse in terms of social and cultural factors, which could prove to be a differentiating factor in the transformation of MENA's science systems.

The overarching aim of this PhD dissertation is to document the transformation of science systems in the Middle East and North Africa region in the recent years. It covers the story of how MENA research institutions have tried to become part of the so-called ‘global science system’, which is the English-language based science system. This system is primarily evident through indicators like Nobel prizes and university rankings, consistently placing Harvard as the top-ranked university globally. MENA research institutions have tried in various ways to adapt to this system by making themselves seen on the global stage. This transformation happened through adapting research evaluation methods developed in the West, collaborating more with other nations, asking scholars to publish in English, and adopting international journal publishing standards. This thesis analyses the process of institutional isomorphism and some of the mechanisms that have helped MENA institutions transform their science systems. Several facets of the scientific systems in MENA are examined. The thesis consists of five distinct studies, developed independently as stand-alone pieces of research that contribute to the existing body of knowledge. They complement each other and tackle the different research questions that are presented in this section.

Specifically, sections 1.2, 1.3 and 1.4 cover the transformation of science systems in terms of scientific collaboration, mobility, research funding and gender parity in science. Section 1.5 and 1.6 are concerned with the role of scientometrics in guiding this transformation process by adapting to global standards along with the adoption of regional scientific content as complements to adopting global standards. Finally, section 1.7 describes the research methods and the data used in this PhD dissertation and also outlines its structure.

1.2 Scientific mobility and collaboration

Scientific mobility refers to the outward and inward physical movement of students and faculty members to other domestic or international communities to engage in learning, collaboration, and research. Scientific mobility is not a recent phenomenon. Scientists have always been mobile, at least since the Middle Ages where famous scholars travelled around the universities in the Islamic world and in Europe. These flows enabled the diffusion of ideas and the exchanges of know-how to benefit larger groups of people. UNESCO recently highlighted the importance of collaboration to enrich science by making knowledge accessible to all but also to engage with diverse perspective and various science systems for a global benefit (UNESCO, 2023). For example, the COVID-19 pandemic is an excellent example of the significance of scientific collaboration in addresses major challenges (World Economic Forum, 2021b). Since the end of the 20th century, digital technology has made possible for more students and scientists to engage in academic activities outside of their own countries in a larger scale. The adoption of the Internet as well as an increased travel connectivity contributed to successful relationships in scientific collaboration (Geman & Geman, 2016). And as the academic labor market has become increasingly global (Slaughter & Leslie, 1997), international mobility is more and more seen as a necessary component of a successful academic career (Jepsen et al., 2014). Rodrigues et al. (2016) show that international mobility has a direct and positive impact on scientific discovery, careers and cultural maturity.

Scientific collaborations are described as interactions between several scientists that allow information transfer and the fulfilment of research activities toward a mutually agreed-upon goal. Co-authorship in journal papers is a specific type of scientific collaboration. In the past two decades, several studies have revealed a rise in the number of co-authored articles in all scientific fields as well as across countries (Moody, 2004; Wagner & Leydesdorff, 2005). Glänzel and Schubert (2004) argue that the structure of international co-authorships has changed since the 1980s.

Scientific collaboration and mobility can be viewed as communication networks (Wagner & Leydesdorff, 2005), allowing knowledge to be exchanged and innovation to be fostered. There have been many studies about scientific collaboration and mobility. Frenken et al. (2009) provide a review of mobility and collaboration studies that use the spatial dimension in scientometrics and propose a proximity approach for spatial scientometric research. Katz (1994) mentions that the distance of partners, the size of the country as well as political and economic factors affect the number of collaborations. More recently, Chinchilla-Rodríguez et al. (2018) compared the position of countries with regard to scientific mobility and collaboration to better understand the dynamics of scientific networks and their impact on scientific development. According to Beaver (2001), historical, cultural and linguistic proximity are the primary motivators for collaboration in scientific research. Cheol Shin et al. (2013) show that collaboration patterns differ across science systems with researchers in developed systems collaborating more than those in developing systems.

Collaboration and mobility are used to integrate global scientific networks (Nerad, 2010). But they are also means to transform national science systems by internationalizing them (Marginson, 2022). International mobility and collaboration are indeed perceived as two sides of internationalization, with the former triggering the latter (Kato & Ando, 2017). For example, Cao et al. (2020) have shown that Chinese returning scientists are instrumental in linking China to the global science network as they continue to actively collaborate with colleagues from their former host science system. Gureyev et al. (2020) offered a comprehensive analysis of the literature on scientific mobility over the past 30 years. They outline the main stages of scientific mobility and consider its advantages, disadvantages, and relationship to brain drain and brain circulation concepts. Netz et al. (2020) also provide a systematic review of the effects of international mobility on scientists' careers. They analyze several dimensions such as networks, productivity, impact or funding and many others. They mention the positive effects as well as the negative effects of international mobility and the expansion of scientists' networks caused by the scientific mobility. All these studies show the significance of scientific mobility and collaboration in driving scientific progress and innovation, and addressing global issues.

Recently, bibliometric methods have offered a few solutions to macro level analyses of international scientific mobility thanks to computational advancements and to author name disambiguation algorithms (Moed et al., 2012; Moed & Halevi, 2014). These methods provide insights into the mobility phenomenon especially in regions, like MENA, where there is a lack of mobility information (Fargues, 2006; Özden, 2006). Such analyses provide useful information to local policy makers that aim to understand the structure and the recent evolution of international scientific collaborations and mobility. At the same time, such information is also useful to comprehend the existing networks at the global and regional levels. Chapter 2[1] sets out to answer the following research question (RQ):

RQ1. *What are the main characteristics of the scientific mobility and collaboration networks at the regional and country levels in MENA? What are the personal characteristics of the mobile scientific workforce in MENA, particularly in terms of academic age and gender?*

Chapter 2 answers this research question by presenting a large-scale analysis of the scientific mobility flows and collaboration linkages based on a total of 1 million WoS-indexed papers published between 2008 and 2017 by 1.5 million authors. Changes in affiliations are used to define a taxonomy of different types of scientific mobility. And co-authorship is used as a proxy measure of scientific collaboration. Three complementary approaches are used to provide a comprehensive analysis of scientific mobility in MENA during this period. This chapter investigates the main destinations and origins of mobile scholars. Then, the academic age and the gender of mobile scientists are determined. Next, findings of this chapter reveal the structure and the mobility and collaboration linkages at the regional and country level and characterize the mobile scholars. Finally, the policy relevance of the scientific mobility and collaboration aspects of these linkages are discussed.

[1] **Chapter 2 is based on:**
El-Ouahi, J., Robinson-García, N., & Costas, R. (2021b). Analyzing scientific mobility and collaboration in the Middle East and North Africa. *Quantitative Science Studies*, *2 (3)*, 1023–1047. https://doi.org/10.1162/qss_a_00149

1.3 Gender differences in science

Gender disparities still persist in our society, and scientific research is no exception. There is an abundance of literature on the gender differences in academia. Expanding higher education is clearly identified as a specific aim in the SDGs: "by 2030 ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university" (UNESCO, 2015b). Many nation states' research institutions explicitly target the SDGs as part of the UN-led transformation process. Research institutions are now increasingly adopting the SDGs as a set of principles to shape their missions and translate aspirational goals into quantitative and long-term benefits (Bautista-Puig et al., 2021).

Participation of women in research from their early career has been relatively well analyzed (Boekhout et al., 2021; Huang et al., 2020; Kulis & Sicotte, 2002; Larivière et al., 2013; Larivière et al., 2011; Macaluso et al., 2016; Su et al., 2009). In most countries, national educational systems make such information publicly available which allows researchers to learn more about the state of female participation from elementary school to higher education and research. Many studies have also attempted to explain the underrepresentation of women in the science system (Ceci & Williams, 2007; Williams, 2018). Such studies cover psychological, sociological and educational factors which may contribute to the gender differences. These studies are useful in covering some of the barriers to the female participation in science since they are mainly related to the organization of the national science systems (political, cultural, economic and geographical environments). More recently, Ceci et al. (2023) analyzed the empirical evidence for gender bias in seven important contexts in the tenure-track academy: tenure-track hiring, grant funding, teaching ratings, journal acceptances, salaries, recommendation letters and, lastly, journal productivity which can moderate bias in the six other contexts. These authors found no evidence of bias between women and men in three contexts (grant funding, journal acceptances, and recommendation letters), but they also found that women were advantaged in the hiring domain. However, bias was found against women in the other three contexts. Considering the significant resources dedicated towards gender bias in academic science, they also suggest that it becomes crucial to gain a precise comprehension of the situations and locations where these research efforts are justified.

Bibliometrics methods have been developed to obtain a complementary picture of female participation in science across geographies from a scientific publication perspective. Halevi (2019) provided an overview of such bibliometric literature on gender differences in science. Halevi groups the bibliometric studies into four main areas: Productivity, Impact and visibility, Performance, and Academic standing. Research has shown that differences between women and men in science are visible in several ways, for instance in terms of participation, productivity, collaboration, authorship, grant applications and citations (Fox et al., 2017; Holman et al., 2018; Larivière et al., 2013; Ley & Hamilton, 2008; Lincoln et al., 2012; Shen, 2013). Women underrepresentation in science in specific regions or countries has been documented in Italy (Abramo et al., 2009), France (de Cheveigne, 2009), Québec (Larivière et al., 2011), and Russia (Paul-Hus et al., 2015). It has also been shown some of the MENA nations show the worst performance globally (Larivière et al., 2013).

Cross-sectional assessments of gender disparities at a given period were common in early bibliometric research (Halevi, 2019). Recent research has begun to provide longitudinal studies that examine scientists' careers across time (Boekhout et al., 2021; Huang et al., 2020). Huang et al. (2020) provide empirical evidence that suggest significant gender differences in total productivity across science, technology, engineering, and mathematics commonly referred to as STEM fields. They note that the increase in the number of women researchers over the past 60 years has paradoxically increased gender differences. Recently, the UNESCO also reported the low representation of women fields and advised that efforts need to be made to address these gaps (UNESCO, 2021b). Understanding these gender differences in science can help in identifying barriers and develop strategies to enhance inclusivity. The gender issue is extensively addressed in the scholarly literature and a diversity of perspectives is needed to better understand gender differences in science (Waltman, 2021).

Comprehensive information at the regional and country level is required to identify region specificities as well as to provide nuanced responses to the concerns that inspired prior research on gender disparities in science. Bibliometric studies offer valuable insights into gender differences in science and are useful to promote gender equity in science. However, it is important to remember that bibliometric indicators alone do not capture nor explain the full complexity of such differences. In chapter 3[1], we start by exploring how countries in MENA have recently engaged with gender policies. We also address the question of how some gender differences have evolved over time in MENA from a bibliometric perspective. More specifically, the research question of chapter 3 is as follows:

RQ2. *What is the proportion of women scientific authors in MENA by country and field? What is the relationship between gender and productivity and lead authorship in this specific region? And how has the participation and performance of women in science changed recently?*

Chapter 3 answers this research question by analyzing a set of 1.7 million WoS papers published between 2008 and 2020 by 1.1 million authors affiliated to institutions in MENA. Based on the first names of the authors of these publications and their suspected country of origin, chapter 3 algorithmically infers a gender for authors. This chapter also defines three cohorts of researchers who started their research career in three different years (2008, 2012 and 2016) by using the year of first publication as the start of their research career. To gain a comprehensive picture, the shares of women authors are computed at the country and field of research level. Next, the number of published papers by authors during the study period, as a measure of publishing productivity, is also examined for each cohort. Combining the gender information with the bibliometric information of scientific papers, and more specifically the authorship position used as a proxy measure of leadership and seniority, chapter 3 compares the probabilities of being first or last author for each cohort. Findings of this chapter provide evidence of the gender disparities in science in MENA in terms of women participation and performance. The work presented in chapter 3 differs from other longitudinal bibliometric studies in that it focuses on the MENA region, with the goal of better understanding the representation of women in scientific research in this specific region in recent years.

[1] **Chapter 3 is based on:**
El-Ouahi, J., & Larivière, V. (2023). On the lack of women researchers in the Middle East and North Africa. *Scientometrics*, *128*(8), 4321-4348. https://doi.org/10.1007/s11192-023-04768-5

1.4 Funding in research from a scientometric perspective

Countries expect research and development to drive their economies and funding agencies play a crucial role in driving scientific progress (Braun, 1998) by supporting universities, research institutions but also researchers through specific and selective programs. Ways and means of allocating research funding are considered influential in governing public science (Aagaard et al., 2021). Such allocations affect the topics addressed, their scope, the scientific outputs with some potential impact on public research. Public funding for research and development plays a crucial role in promoting economic growth and competitiveness (Congress, 1991), while also contributing to the advancement of society (Sarewitz, 1997). Understanding how research funding influences the scientific production and dissemination can provide insights into the effectiveness of funding mechanisms and their potential implications (Heyard & Hottenrott, 2021). Many studies have sought to determine the returns of investment of the research and development enterprise. To do so, different strategies have been used. One way of quantifying such returns at the macro level include for instance studies of the effects of public research and development expenditure on economic growth (Paasi, 1998; Salter & Martin, 2001; Terleckyj, 1985). Another approach consists of measuring the effect of public funding on the volume, impact and quality of scientific output (Adams & Griliches, 1998; Campbell et al., 2010; Cronin & Shaw, 1999; Harter & Hooten, 1992; Payne & Siow, 1999; Sandström, 2009).

Since 2008, the WoS has been capturing funding acknowledgements mentioned in scientific publications (Clarivate, 2022). This data allowed researchers to analyze funding sources (Alvarez-Bornstein & Montesi, 2021). Early studies on funding acknowledgements focuses on analyzing the coverage by country, disciplines, and document types, as well as methods to standardize funding organizations (Álvarez-Bornstein et al., 2017; Costas & van Leeuwen, 2012; Sirtes & Riechert, 2014). Later, research explored funding mechanisms, such as Möller's identification of major funding agencies and their contributions to national performance in five European countries (Möller, 2019). These studies are example of analyses which provide insights into bibliometric approaches to studying research funding and funding acknowledgements. They explore the patterns of funding, the impact of funding on research outcomes, and the significance of acknowledging funding sources in scientific publications. Such analyses allow the comparison of research funding systems across different countries. For instance, a study points at the internationalization of research funding through the involvement of various funding sources at organizational and national levels (Wang & Shapira, 2011). In a different study, it is observed that China is dominated by a single funding agency while the United Kingdom has more diverse funding sources (Wang et al., 2012). Similarly, other studies focused on specific groups of countries, such as the G9 countries (Huang & Huang, 2018), the Global South (Chankseliani, 2023), the BRICS (Shueb & Gul), or individual countries (Costas & Yegros-Yegros, 2013; Díaz-Faes & Bordons, 2014; Gao et al., 2019; Gök et al., 2016). More recently, Chataway et al. (2019) analyzed the trends in science funding support in Sub-Saharan Africa. They found relatively low levels of funding and cross-country collaboration, and they highlighted the need for capacity building.

The literature on funding acknowledgements found in scientific publications highlights the importance of recognizing the financial support received to conduct research. Although such acknowledgements come with limitations, they provide some insights into funding sources and drive accountability in research. It is worth noting that, in some cases, it is possible to conduct research without funding support. Also, some researchers might not acknowledge the financial support received from their direct employer. It is difficult to map the relationships between scientific research funding and the outcome of the individual related research activities. Another complex aspect of such relationship concerns the scientific collaborations drawing on multiple and various types of sponsorships to produce research outputs. Nevertheless, the exploration of funding acknowledgements can still contribute to improve our understanding of the funding landscape and its influence on scientific research.

As mentioned in section 1.1, countries in MENA have been looking to become knowledge-based economies (OECD, 1996). Economies in MENA have relied heavily on natural resources (World Bank, 2019). A report from UNESCO (2015c) provides some insights into science funding in MENA, indicating low investment in Research and Development at that time. However, some changes were already taking place. Significant investments in science and technology capacity have been made in recent years to foster research and innovation in MENA countries (Schmoch et al., 2016; Shin et al., 2012; Siddiqi et al., 2016).

The availability of funding acknowledgements in large bibliometric databases such as the WoS has received a lot of attention from funding agencies and researchers due to their potential to demonstrate the contribution that funders have had to scientific research and society at large. This raises some important questions such as how the funding structure and the recent investments in science may reflect in scientific publications at various levels (specific regions, countries, fields, research institutions, scientific journals and authors). However, there is not yet enough evidence in the literature to show how funding acknowledgements can actually determine the contribution of research funders in MENA. In this context, chapter 4 addresses the following research question:

RQ3. *To what extent has the funding structure in MENA evolved over recent years? What are the characteristics of the major funders in MENA, in terms of type and location?*

Chapter 4 answers this research question with an extensive analysis of funding acknowledgements found in 2.4 million scientific papers published between 2008 and 2021 by authors affiliated to institutions located in MENA countries. The primary objective is to uncover valuable insights into funding activities in the region by identifying the major funding agencies and evaluating their contributions to national scientific publications. To achieve this, publication data from the WoS is examined. A major step consists of applying a data unification process to variant names of funding agencies. Chapter 4 also classifies the funders following the typology of funding agencies found in InCites. The location of the funder is also considered to determine the source of funding granted to researchers. Findings of this chapter shed light on the structure of the research funding in MENA. The results of this chapter also confirm the complex nature of funding in research, especially when co-authorship and co-funding are involved.

1.5 Scientometrics data and science systems

Science systems and research management are currently strongly linked. As documented by Franssen and Wouters (2019), bibliometric methods have increasingly been developed and used within the context of science policy as a tool for research evaluation since the 1980s (Moed et al., 1985). Today, bibliometric methodologies are primarily understood in the context of research performance evaluation (Moed et al., 2004). Through economic incentives, promotion, research funding, and reputation, research management and evaluation have a significant impact on knowledge development. The shifting academic landscape has resulted in an increase in research on staff assessment and, more particularly, the criteria used to evaluate faculty members (Lamont, 2012; Musselin, 2009; van Arensbergen et al., 2014).

Scientific research has primarily been judged on the basis of scientific publications, which are the major and presumably measurable product of science (Geuna & Martin, 2003; Hicks, 2012).Various criteria of publication performance are used to evaluate and rank academics and universities (Hirsch, 2005; Narin & Hamilton, 1996), which involves the distribution of research money and the assignment of academic roles (Geuna & Martin, 2003; Hicks, 2012). The accuracy of modern performance metrics, such as productivity, citation indexes, and peer review, has been the focus of a massive literature (Anninos, 2014; Basu, 2006; Werner, 2015). *Scientometrics*, term coined by Nalimov, are methodological approaches in which scientific publications become the subject of analyses (Vlèduts et al., 1959). For instance, mathematical and statistical approaches are used to assess the structure of a discipline, investigate emerging trends in a specific discipline, research networks, authorship, citation impact, future trends or funding patterns (Moed, 2006). These quantitative methods aim at analyzing and measuring various aspects of scientific literature and they have the potential to inform research policies and decision-making. And such analyses have gained popularity in the business of research in recent years (Donthu et al., 2021). H-indices, citation counts, and Journal Impact Factors (JIF) are few examples of bibliometric indicators used when evaluating research (Wouters, P. et al. 2015). Recent research has shown how such indicators are used as 'judgment devices' (Karpik, 2010) in research evaluation (Hammarfelt & Rushforth, 2017).

Many researchers argue that bibliometric methods have become overly influential (de Rijcke et al., 2016; Hammarfelt & Rushforth, 2017; McKiernan et al., 2019), and their utility is still controversial. This may result in a situation where "we risk damaging the system with the very tools designed to improve it" (Hicks et al., 2015). Another study of university changes found that there is now a bias in favor of research quantity rather than quality, a bias in favor of traditional methodologies and a bias towards short-term performance rather than long-term research capability (Marginson & Considine, 2000). Weingart (2005) argues that the introduction of bibliometric approaches is a response to the pressures on science systems to legitimize themselves. Once these metrics are applied as indicators on which distributive decisions are based, they have intended and unexpected consequences, as well as a high political significance. According to sociology of science and ethnographic studies, scientists do react to non-epistemic factors (Gläser et al., 2002). National governments in both developed and developing countries have frequently sought to improve their science systems in recent years, for example, through large-scale reforms (Huang et al., 2015), research assessment programs (Weingart, 2005), or changing funding mechanisms to incentivize priority research areas (Roco, 2011).

For example, individual researchers are rewarded with bonuses by the Spanish National Commission for the Evaluation of Research Activity (CNEAI) for publishing in prestigious journals. According to Jiménez-Contreras et al. (2003), researchers have responded by boosting their research output, achieving the CNEAI policy's goal. Another study from Australia found that after formula-based financing was implemented, the quantity of publications increased, but the quality of the papers, as evaluated by citations, did not (Butler, 2003). This policy-driven incentive resulted in predictable counter-productive outcomes. In Argentina, the tenure evaluation system of the National Scientific and Technical Research Council (CONICET) heavily relies on publication records, often prioritizing publications indexed in the Web of Science and Scopus (Beigel, 2014).

These strategies and policies, which originated mainly from North America and Europe, are translated locally in other regions of the world. They are being copied and adopted by emerging countries as ‘global standards’. For example, Fischman et al. (2010) have observed that many editors of Latin American journals define ‘quality’ in terms of ‘international standards’ which are hegemonic in nature. Another instance of global standards correspond to independent international ranking systems, widely used by students to gauge the performance and reputation of universities globally, are being discussed and adopted strategically by research institutions in MENA (Damar et al., 2020; Gaillard & Bouabid, 2017; Karabchuk et al., 2021; ریگی et al., 2020). Such rankings are usually based on quantified measures of institutional performance, teaching quality and research productivity and impact (Cashell, 2012; Hazelkorn, 2014). In Turkey, Tonta and Akbulut (2020) examined whether monetary support increase citation impact of scholarly papers and found that such programs should be reconsidered. In the case of Iran, Sotudeh (2011) argues Iranian scientists have adopted a relatively poor publication strategy and that Iran needs to watch more carefully the functioning of its science system. She also suggests this may be the case for similar science systems, where the emphasis is given to quantity rather than quality. In this context, chapter 5 explores the recent adoption of bibliometric data by policy makers and research managers in MENA and how such adoption has transformed their science systems. Chapter 5 tackles the following research question:

RQ4. *How do research managers in MENA adopt global scientometric standards in local contexts? In which local processes are scientometric data and indicators used and what specific functions do they serve?*

This chapter explores the usage of scientometric data and indicators by research managers in different science systems in the MENA region. Chapter 5 answers the research question through a qualitative study of transcripts of semi-structured interviews with research managers. The research managers come from 12 distinct research institutions spanning across 9 different countries in MENA. Chapter 5 proposes that the implementation of scientometric-based rules plays a crucial role in driving this transformation. Firstly, this chapter highlights how research managers adopt scientometrics as global standards. Additionally, it demonstrates the adoption of various scientometric data and indicators through a 'glocalization' process. Lastly, the study illustrates how research managers leverage this data to inform decision-making and policymaking. By exploring research management and evaluation based on publication activities in specific contexts, this study contributes to a better understanding of the usage of scientometric data by research managers. Moreover, it explores how such data facilitates the adaptation and transformation of local science systems to align with global standards.

1.6 Databases of regional scientific literature

De Swaan (2013) describes the global language system as a *galaxy of languages* with five or six thousand languages spoken on earth. As mentioned earlier, one way to be part of the current global science system is to publish in English, the current *lingua franca* of international science. Publishing research papers in English is necessary to reach the broadest audience and have the most impact in the international scientific community (Garfield, 1989). However, this represents a barrier for non-native English speakers. There have been some debates about the hegemony of English in academia and the importance of supporting scholarship in diverse languages. For instance, Beigel (2014) calls for a more nuanced system that recognizes the value of scientific research published outside the mainstream ‘international’ publishing circuits sustained by the WoS, Scopus and Google Scholar, particularly in the social sciences and humanities where local contexts and languages are important. In the Arab world, Hanafi and Arvanitis (2014) argue that Arabic is being marginalized in social science research. They identify structural constraints and dependency by choice as the two main reasons for this marginalization. In terms of structural constraints, Hanafi and Arvanitis (2014) acknowledge external factors such as the dominance of English in academic publishing, as well as the globalized nature of higher education with Western models. As for the dependency by choice, they argue that Arab scholars themselves contribute to this marginalization because of the perceived lack of prestige associated with Arabic scholarship.

There have also been some debates about the inclusiveness of bibliographic databases such as the WoS. For example, Gibbs (1995) claimed that the Science Citation Index in WoS was biased toward Global North English-language scientific journals. Garfield (1997) responded that a statistically valid definition of bias was needed to conclude whether WoS was biased against so-called ‘Third World’ journals, referring to the law of concentration applied to science journals or Bradford law (Garfield, 1996). Several studies show that the Institute for Scientific Information (ISI) reinforced the concept of prestigious "core journals" and established the impact factor as the gold standard for success in an increasingly English-dominated publishing system (Guédon, 2011; Heilbron, 2002; Ortiz, 2009). ISI's databases were the only option for evaluating research output of countries and institutions until 2004, when Google launched Google Scholar and Elsevier launched Scopus, offering a comparable alternative to the WoS (Archambault et al., 2009). In a few recent large-scale studies, the coverage of the WoS was not only compared to Scopus but also to other databases such as Google Scholar, Dimensions, Microsoft Academic, Crossref or OpenCitations' COCI (Martín-Martín et al., 2021; Visser et al., 2021). Martín-Martín et al. (2021) found that different bibliographic databases sources have some overlap, but no source covers all documents and Visser et al. (2021) emphasize the need to combine different sources for a broad coverage to properly take into account locally relevant research.

Earlier, Hicks (1999) discussed the difficulty to achieve full comprehensiveness of international social science literature and the bibliometric consequences this might have. She argues that the polyglot character of the social sciences might make their indexation in a single database difficult. To overcome this challenge, many institutions have set up translation departments to assist scientists in publishing in English. But translating is difficult, time consuming and expensive. Another way is to build national databases to complement the international citation database like the WoS. National databases have common missions such as the global promotion of local research, dissemination of local findings and objective evaluation of national institutions, researchers, and journals across all fields. Several databases have been developed with these intents such as the Chinese Science Citation Database, the Korean Journal Database, the Serbian Citation Index, SciELO, the Russian Science Citation Index, the Taiwan Humanities Citation Index and the Citation database for Japanese papers (Chen, 2004; Choi et al., 2013; Jin & Wang, 1999; Lucio-Arias et al., 2015; Moskaleva et al., 2018; Pajic, 2015).

Sivertsen (2018a) argues that the usage of local language in scholarship is crucial to promote engagement with stakeholders and the general public. This is also necessary for science to fulfil its social responsibilities or have localized impacts (Garcia-Ramon, 2003; Hasse & Fischer, 2003; Huang, 2011; Samers, 2000). However, if evaluation processes have an impact on publication practices and if they change research agendas, researchers may decide to prioritize English- speaking audiences over locally relevant research (Bianco et al., 2016). It is also worth noting that different languages and communication venues affect different audiences (Hicks, 2004). Non-English journals play unique roles that are distinct from those of mainstream English journals, as highlighted by Chavarro et al. (2017): they offer researchers opportunities to enter into the world of scientific publishing and address topics that might receive less attention in mainstream publication titles. There is an extensive literature about the development of regional or national databases.

In Europe, Jappe (2020) examined the practice of performance assessment of public research in 21 EU countries between 2005 and 2019. She found that the bibliometric evaluation in Europe was mostly reliant on the citation indices provided in the WoS. Citation data was supplemented in other studies by national databases, which are more extensive in terms of research output (Sīle et al., 2018). Kulczycki et al. (2018) have studied the publication patterns in SSH from the Czech Republic, Denmark, Finland, Flanders (Belgium), Norway, Poland, Slovakia, and Slovenia. They have shown it is possible to get a more complete view of academic publications by using national databases built upon institutional research information systems.

In Latin America, Velho and Krige (1984) critiqued the Web of Science (WoS) for favoring industrialized nations and fundamental sciences, arguing that this bias skewed the representation of publication trends in other disciplines and contexts. Meneghini et al. (2006) conducted an analysis of the Scientific Electronic Library Online (SciELO) database and found that with SciELO's comprehensive database, researchers can now conduct studies that were impossible using the limited coverage of the WoS. Aguado-López et al. (2014) examined the coverage of Ibero-American journals in the Journal Citation Reports (JCR) and SCImago Journal and Country Rank (SJR). Aguado-López et al. found that the coverage of Ibero-American journals in these databases was relatively low and concluded that there was a need to increase their visibility. More recently, Bohórquez (2017) studied the emergence and growth of regional journal indexing systems in Latin America in light of the dominance of the WoS. Bohórquez argues that alternative journal indexing systems emerged to cover journals that do not pass the standards required by the WoS but also as a response to geographical, linguistic, and disciplinary biases in the coverage of the WoS.

In the MENA region, specifically in Iran, Jafar Mehrad and Mohammad Ghane (Shiraz University) led the development of the Islamic World Science Citation Center (ISC) after 2002, in collaboration with the Ministry of Science, Research, and Technology. The ISC aims to provide an evaluation of Iranian scientific journals based on principles used in the WoS and scientometric indicators. ISC now covers more than 1,800 peer-reviewed Iranian journals, with a total of 3,400 titles published in the Organization of Islamic Cooperation (OIC) countries. ISC has indexed approximately one million records, with more than 40 million references and one million citations. In terms of language distribution in ISC, about 54% are English language journals, 38% are in Farsi, and 8% are published in Arabic journals (Mehrad & Ghane, 2020).

In Turkey, the Turkish Academic Network and Information Centre (ULAKBIM) created the TR index in 2000 in accordance with international standards in terms of evaluation and selection process. Its aim is to better reflect the scientific knowledge of Turkey by indexing national scientific journals in all fields including Science and Social Sciences, and subfields of Dentistry, Pharmacy, Engineering, Basic Sciences, Health Sciences, Veterinary Medicine, Social Sciences and Humanities. As of May 2023, the TR index consists of 2,645 titles published in several languages: English (47% of the titles), Turkish (40%), German (4%), French (3%), Arabic (3%) and other languages (3%).

The Arabic Citation Index (ARCI) was funded by the Egyptian government and launched in 2020 in partnership with Clarivate. It gives access to bibliographic information and citations to scientific publications from more than 600 selected Arabic journals (http://arcival.ekb.eg).

ARCI aims at bridging the gap between local and global scientific production. It also aims at enabling scholars to work with a broader audience on national, regional, and worldwide research projects by providing access to local scholarly information. Another goal is to increase the exposure of research published in Arabic scientific publications, therefore expanding the Arabic academic footprint on the global scientific map. These significant advances from MENA nations confirm the significance of national indexes as a regional addition to worldwide citation indexes like the WoS or Scopus.

As documented by Franssen and Wouters (2019), the development of comprehensive national databases has opened a new direction of bibliometric studies allowing for a more complete view of the scientific publication practices by covering more sources, and research with local relevance than highly selective international databases. Thanks to these regional and national databases, a more comprehensive view of each country's research output is now covered, and local research evaluation systems can benefit from more inclusive evaluation of scientific work not indexed in international citation databases (Jin & Wang, 1999). From this perspective, chapter 6 sets out to address the following research question:

RQ5. *What are the predominant research domains and topics represented in the literature indexed in the Arabic Citation Index (ARCI)? How can the usage of this bibliometric database contribute to a more comprehensive and inclusive assessment of research activity in the MENA region?*

Chapter 6 answers this research question on the basis of bibliometric analyses of ARCI. This chapter provides a comprehensive overview of the scientific literature indexed in ARCI and explores its potential applications in research evaluation. About 140 thousand scientific papers published between 2015 and 2020 are analyzed by using their available metadata. Besides, chapter 6 presents the distribution of such literature at various levels such as research domains, countries, languages but also open access availability. In addition, unsupervised machine learning techniques and text mining algorithms are employed to reveal the main topics covered by ARCI. Findings of this chapter show how ARCI can complement so called ‘global standards’ in the context of more inclusive research assessment processes. Finally, this chapter also discusses the study findings and open up to several research opportunities suggested for further exploration.

1.7 Research methods and structure of this dissertation

1.7.1 Reflections on research methods

In this dissertation, a mixed-methods approach is employed, combining both quantitative and qualitative methods to comprehensively address the research questions. This approach was chosen to ensure a pragmatic investigation, allowing us to leverage various data analysis skills and to benefit from a diverse range of data sources.

The research questions of this thesis require data from diverse perspectives, including scientific publications, surveys, and interviews. Additionally, access to various data sources played a crucial role in shaping the methods employed in this dissertation. Rigorous quantitative techniques were applied to extract meaningful insights from large-scale datasets, perform statistical analyses, and identify patterns and trends relevant to the research questions related to scientific mobility, collaboration, and gender differences. And qualitative methods allowed a deep dive into the nuances and complexities of specific phenomena such as research management and research funding through in-depth semi-structured interviews, email surveys as well as manual analyses of random samples of scientific publications.

In general, some research questions are best answered through quantitative analysis, while others demand a more interpretive and contextual understanding, achievable through qualitative methods.

I want to highlight here the pragmatism underlining the research design of this PhD thesis, incorporating a diverse set of data skills, and using various data sources. At the same time, liberal choices in employing different methods also show a commitment to addressing research questions from multiple perspectives, as well as an ability to handle a wide range of research questions, contributing to the broad focus of the dissertation.

1.7.2 Structure of this dissertation

This thesis consists of seven chapters and is organized as follows. There are five chapters after this introduction, each of which presents a distinct study. The first three studies, primarily descriptive, show the recent transformation of science systems in MENA. Scientific mobility and collaboration in MENA are examined in Chapter 2. Chapter 3 explores gender differences in the scientific workforce from the United Nations sustainable development goals' perspective. And chapter 4 covers the research funding structure based on funding acknowledgements found in scientific publications. The other two studies document how scientometrics are used as global standards. Chapter 5 is concerned with the role of scientometrics in guiding the transformation of science systems in MENA, as research institutions in the region increasingly adopt scientometric data as 'global standards' and integrate them into local research systems, thus influencing decision-making and policy formulation. Chapter 6 provides a better understanding of the Arabic scientific literature as a complement to adopting global standards. Finally, chapter 7 concludes the thesis by summarizing and contextualizing the main findings presented in chapters 2 to 6. Based on the research findings, chapter 7 also elaborates on the implications of the findings of the thesis and puts forward prospects for future research.