



# **STRUCTURAL CHARACTERISTICS OF THE MTA RESEARCH NETWORK'S SCIENTIFIC PERFORMANCE 2015-2017**

**BASED ON WEB OF SCIENCE DATA**

## Introduction

The aim of the present report is to provide an overall picture of the MTA research network's (henceforth MTA) internationally measurable scientific (publication) performance and its structure. The analyses reveal the characteristics of the MTA research network's internationally measurable scientific (publication) performance with regard to all disciplines based on international standards. The report focuses on the 3-year-period ending in 2017 – thus it contains the internationally visible publication output from 2015 to 2017. The working sample has been sorted from high-ranking journal articles (reviews, articles) bearing in mind that relevant international standards rely on these. The analysis is based on Web of Science (WoS) publications assigned to MTA (this is a thoroughly cleaned data set created with the purpose of a deeper structural analysis and the analysis of MTA affiliations). The overview relies on various other international databases with regard to referential values, disciplinary classification, etc.

The analyses rely on relative and normalised indices thus enabling the commensurability of disciplines. International comparison is also based on relative indices. Disciplinary classification is based on the Fields of Science system, the OECD standard introduced by the Frascati Manual.

Applied databases and sources

**Web of Science:** citation database (SCI, SSCI, A&HCI): bibliographic data

**JCR:** journal impact factors (JIF) and their disciplinary rankings (Q1-Q4)

**InCites:** indicators and limit values for each discipline based on WoS data

## Quantitative and Qualitative Characteristics of Publication Output

The output for each field and discipline of science is represented via three relative indices. In order to compare different fields, publication volume (1) is represented through shares in the national output, and quality (2) is revealed through the proportion of publications in Q1 journals for each field (Q1%). The latter indicator refers to publications appearing in journals that belong to the highest 25% percent in each field with regard to impact factor (JIF) (in case of multiple category affiliation, the most favourable option was taken). For the

**Fields of Science:** the hierarchic classification of disciplines introduced in the Frascati Manual, the scheme applied in the OECD

**Output indicator:** propor-

sake of completeness, the present analysis includes the proportion of cited publications within the overall output (cited%). As this indicator is greatly influenced by disciplinary characteristics it has only informative value.

The above mentioned characteristics of performance are summarized on Figure 1 (main disciplines) and Figure 2 (other disciplines). Indices show a yearly average with regard to the 2015-2017 period. The research network has the highest national share in natural sciences (~45%), dominated by physics and astronomy, yet mathematics, chemistry and biology have similar proportions. Natural sciences are followed by engineering and social sciences (~30% each), in case of the former, mechanical engineering, material science, industrial biotechnology and nanotechnology are the most prominent fields, regarding the latter, psychology leads the quantitative ranking. It is important to highlight that humanities lead the next larger group, closely followed by the volume of agricultural sciences and medical and health sciences (~20% each). As for humanities, each sub-discipline contributes to their overall share with roughly similar proportions, whereas agricultural sciences are led by agronomy-forestry-fisheries, and medical sciences by theoretical medicine.

A qualitative approach to the MTA output greatly changes the above detailed picture. 45-70% of the main disciplines' output belongs to Q1 (except for humanities): the quantitatively low-ranking medical and health sciences come on top (almost 70%), especially within clinical medicine. Natural sciences (physics and astronomy, chemistry, environmental science, biology), engineering (relatively even distribution among sub-disciplines) and agricultural science (agronomy-forestry-fisheries) also rank high (~60%). Almost half (~45%) of the social sciences output appears in Q1 journals, the most important fields are psychology and communication science (information science and scientometrics). Considering the journal article output of humanities, 20-25% appears in Q1 journals.

tions within the total national output

**Q1%:** the proportion of those publications that appear in journals belonging to the first 25% of the impact factor (JIF) ranking (if a journal is ranked for more than one discipline, the more favourable result was taken)

Figure 1 The output of major disciplines and their qualitative indices

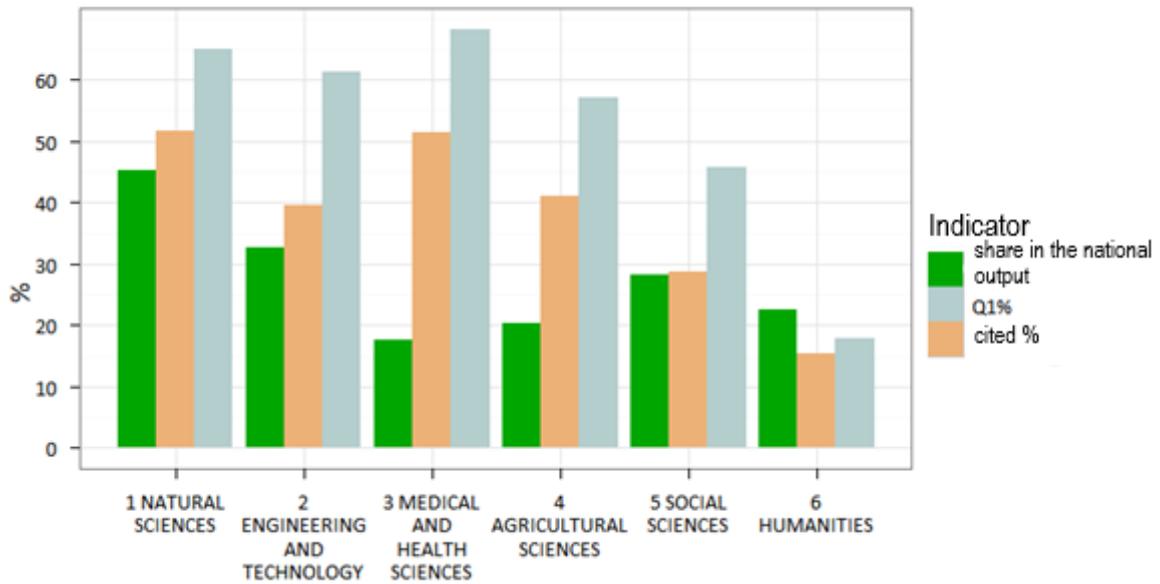
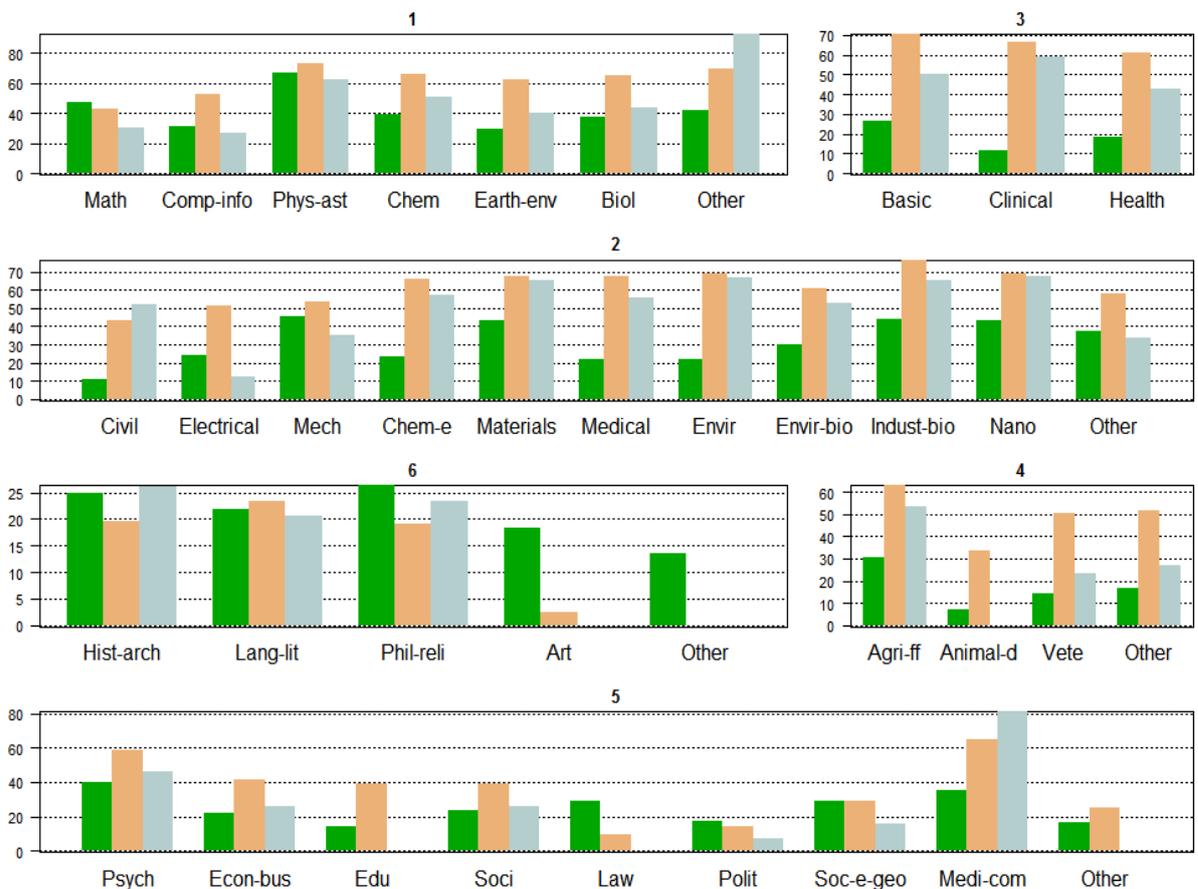


Figure 2 The output and qualitative indicators of sub-disciplines (abbreviation list available in the appendix, numbers denote the major disciplines of Figure 1)



## The (Short-Term) Impact and Excellence Indicators of Disciplines

Citation impact – a basic form of measuring a research network’s success – is represented through size and area independent indicators. Following the best international practice, impact is shown via (1) mean normalized citation score (MNCS) and (2) the 10% excellence index, denoting the most cited 10% of publications for each discipline. The latter enables us to compare disciplines and publication types, as well as to relate the MTA output to international standards. In order to measure impact, the present analysis focuses on the current citation scores of 2015-2016 publications (the minimum period during which the citation rate of most disciplines becomes visible is 2-3 years). Since the normalized index of scientific impact correlates the number of citations to the average impact of international output for each discipline, the possibility arises to apply regions and country groups as benchmarks for MTA output. The below analysis contains multiple comparisons based on international reference values and on the (aggregated) performance of EU13 and EU28 countries.

Figure 3 represents the average (normalized) impact of the major disciplines. In addition to the average citation index of each discipline, the figure compares the MTA citation impact to the EU13 and EU28 average (=integrated international). It is essential to notice that the research network ranks above the EU13 average and close or above the EU28 average. This is especially true for natural sciences (~1.6 fold over international average), whereas medical and health sciences, social sciences and agricultural sciences rank around the international average. The latter group is important to highlight as it illustrates that low quantitative measures do not equal a performance deficit: these disciplines impose a remarkable scientific impact (and attract international attention) despite their smaller size.

This relatively steady picture slightly changes if the excellence index (PP10%) is considered. Figure 5 compares the major disciplines along these two indicators (MNCS – integrated international, PP10%). Despite the fact that natural sciences rank high in both respects, other disciplines with citation indexes around world average have different shares in the international “elite.” It is important to highlight that social sciences achieve the same result as natural sciences (12-13%), moreover, medical and health sciences also rank above the internationally expected result (10%). Similarly, the disciplines with a lower share stay close to the expected value (humanities rank closest to it).

Figure 4 and 6 shows the impact structure of the sub-disciplines with regard to MNCS and MNCS vs PP10%, respectively. This more detailed

**MNCS:** It correlates the annual number of MTA publications to the disciplinary average (the average citation number within the discipline in the given year). Reference value (it refers to the international average) = 1.

**PP10%:** The proportion of MTA output belonging to the most cited 10% within a discipline in each year. Reference value = 10%

picture of natural sciences confirms the above represented situation: the outstanding results of physics and astronomy are mostly responsible for the high scores in case of both indicators, yet other fields of natural sciences also rank around the international average (mathematics has the same result as EU28 and EU13 countries). With regard to excellence, medical and health sciences also score around the international standard. As for engineering, environment and biotechnology (Envir-bio), environmental management (Envir) and chemical engineering have the most outstanding result (if we consider both indicators). The average impact of social sciences (on a global scale) is dominated by sociology, psychology and communication science (information science, scientometrics). Pedagogy, economics and law (together with sociology) scores around the world average, and have a very high excellence index (~20%). Most subfields within humanities (history, archaeology, philosophy and religious studies, linguistics and literary studies) reach the EU28 and also the world average. The excellence index reaches the internationally expected value in linguistics and literary studies (10%).

Figure 3 Mean normalized citation scores (MNCS) with regard to three reference values

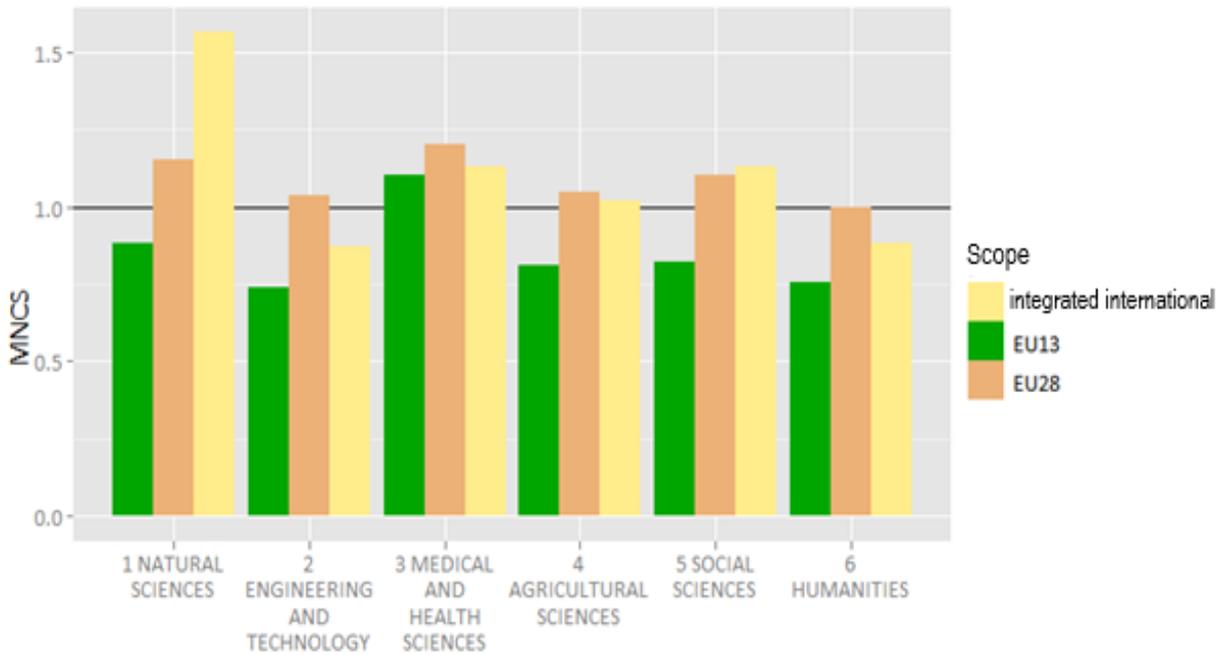


Figure 4 Normalized citation impact of the disciplines with regard to three reference values (see the abbreviation list in the appendix, numbers denote the major disciplines)

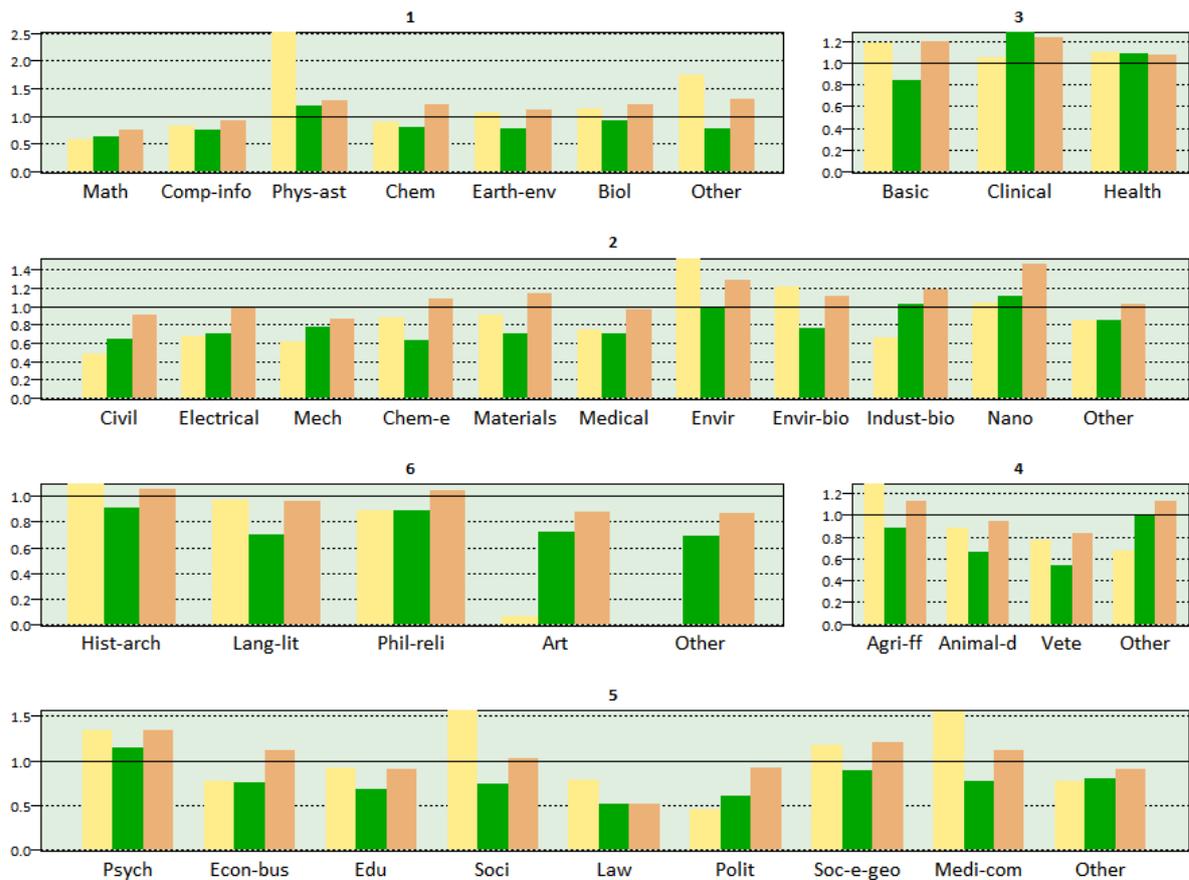


Figure 5 Mean normalized citation score of the major disciplines (MNCS, integrated international) vs. excellence index (PP10%)

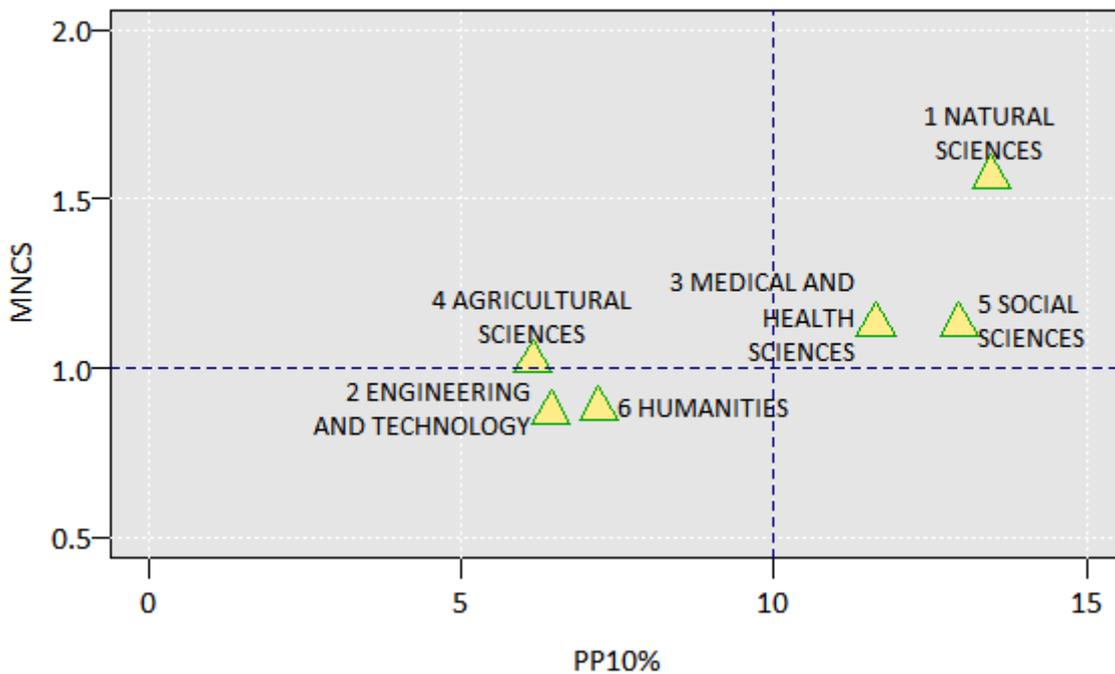
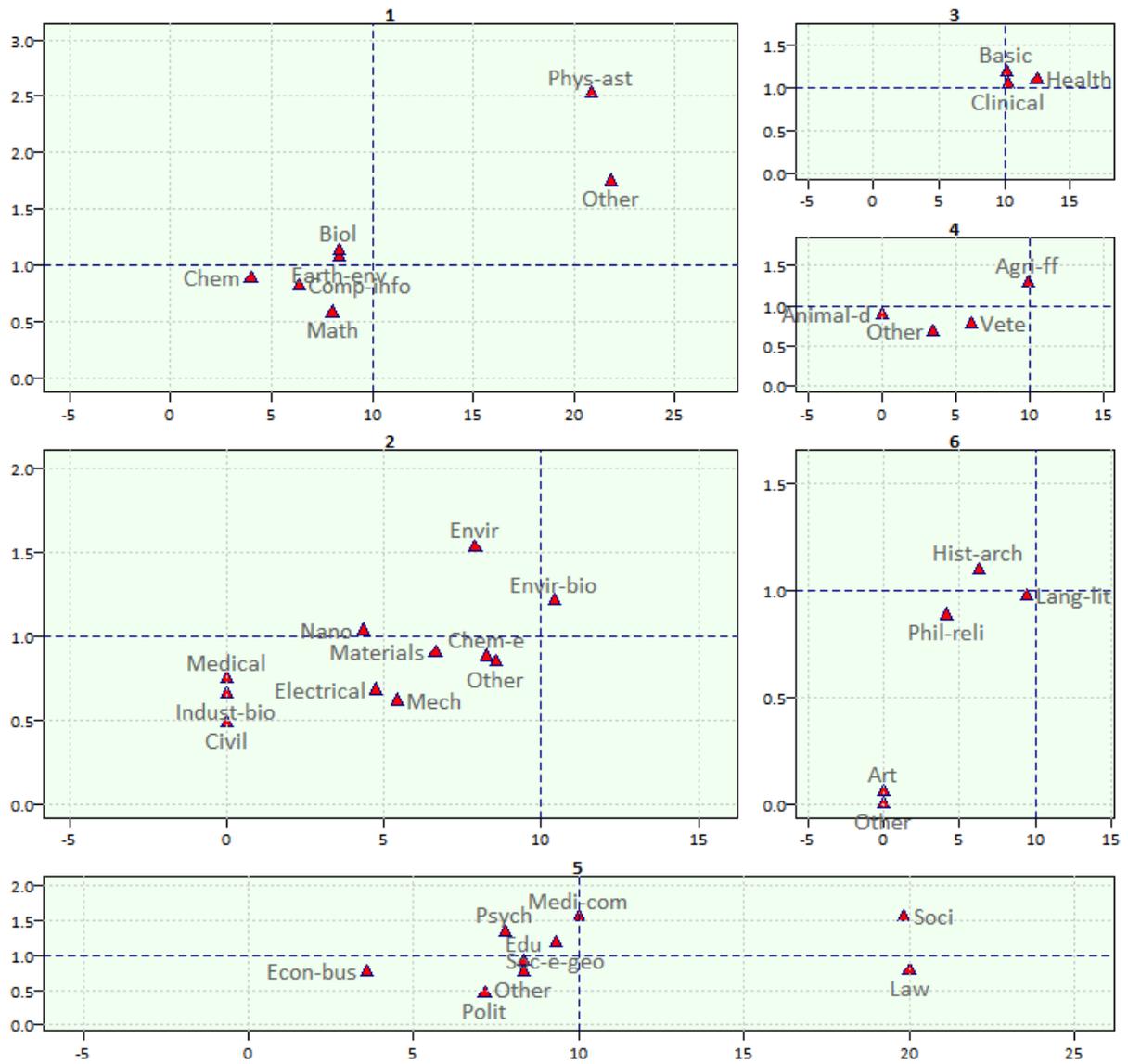


Figure 6 Mean normalized citation scores of sub-disciplines (MNCS, integrated international) vs. excellence index (PP10%) (see the abbreviation list in the appendix, numbers denote the main disciplines).



## The Internationally Visible Research Network of the MTA

The significance and success of the research network within the national and international R&D system can easily be illustrated through the MTA co-operation network. The basic indicators of scientific co-operation are co-authorship and its indexes. The research network's international relations can be characterised by the portion of co-operation-based publications in the total output. In addition, it is also revealing to go beyond scientific impact and track social impact. Publications based on academy & industry relations (the proportion of publications produced in co-operation with the public sector) can be an indicator for social impact.

Figure 7 represents these two basic indicators of collaboration with regard to the 2015-2017 total output. According to the figure, collaboration is significant on all fields, yet a few self-evident differences arise due to the region specific characteristics of co-authorship (and internationality). Despite the latter, social sciences (a discipline with traditionally low co-authorship rates) achieves a high percentage (~40-50%), only slightly below that of natural sciences, medical sciences and engineering (60-70%). Humanities achieve 20%. Publication produced in collaboration with the private sector reveal an even more interesting structure: the most academy & industry collaborations are carried out in medical and health sciences (~3.5%), whereas natural sciences, engineering, agricultural sciences and even social sciences have similar results.

Another important characteristics of the network are the names, impact and role of collaborating countries. The number of co-operation-based publications among partner countries and the percentage of such publications in the total output is a wide-spread indicator of collaboration. To this end, two methods are applied: (1) the conventional approach marks the impact of each country with the number of collaborations they take part in. This is called "full counting." According to contemporary bibliometrics, however, (2) "fractional counting" provides a more realistic picture regarding the intensity of co-operation, as it considers the share of each country in the common research, and assigns publications accordingly. This method enables us to somewhat counteract distortion due to hyperauthorship, which strongly over-represents countries with large author consortiums in high output disciplines (see high-energy physics).

The impact of the most important co-operation partners within the international (publication) network is shown on Figure 8 (share >1%). As before, the most significant collaboration partners in the 2015-2017 period are the USA and Germany (>10%), the next group consists of Italy, France and England (6-8%). This section, dominated by "west-

The most basic indicator of international scientific collaboration is the proportion of publications produced in international co-operation.

**Full counting:** the traditional way of measuring co-operation intensity, it allocates the co-authored work to each contributor (thus it multiplies the number of international publications)

**Fractional counting:** a recently suggested way of measuring co-operation intensity, that divides co-authored publications in relation to the authors' participation (this method does not multiply the number of international publications)

ern countries”, is followed by a – in a broader sense – Western European group (Switzerland, Sweden, Spain, Belgium, Austria), two significant Asian countries (Japan, China) and Russia (3%). The so called post-socialist countries also form a distinct cluster with a 2% share.

The picture representing the impact of partner countries significantly changes, if in addition to proportions in the total output, we consider disciplinary variability. Figure 10 reveals the disciplinary impact and dispersion (SD) of partner countries. It is clearly visible, that with the exception of Germany that has an equally high impact in all disciplines ( $10\pm 1\%$ ), even countries with the highest overall volume (such as the USA and England) have varying proportions if we look at the data for each discipline. Figure 9 shows the partner network for the six major disciplines. The integrated pattern fits natural sciences, medical and health sciences, as well as engineering, yet the volume of the USA is lower in engineering and in medical sciences, since Western European countries are almost exclusive. With regard to agricultural sciences, post-socialist countries (Slovenia, Slovakia) appear alongside Middle Eastern and Far Eastern countries (United Arab Emirates, India), yet Italy has the highest proportions (alongside Germany). Considering social sciences, the USA, Germany, England and France are the most significant partners, followed by Belgium and Austria – these results go against the generally shared preconception that this discipline limits partnerships to regional collaboration (note however, that this is not true for international journal publications). The above mentioned preconception is even more frequently applied for humanities, yet – as the visualisation clearly shows – collaboration work is led by England and Germany (the presence of the USA is not characteristic in this discipline), followed by China and the Netherlands with high proportions compared to other disciplines. Nevertheless, the relatively high score of Austria demonstrates the importance of regional co-operations.

Figure 7 The proportion of publications produced in international collaboration vs. the proportion of publications in collaboration with the private sector with regard to the major disciplines

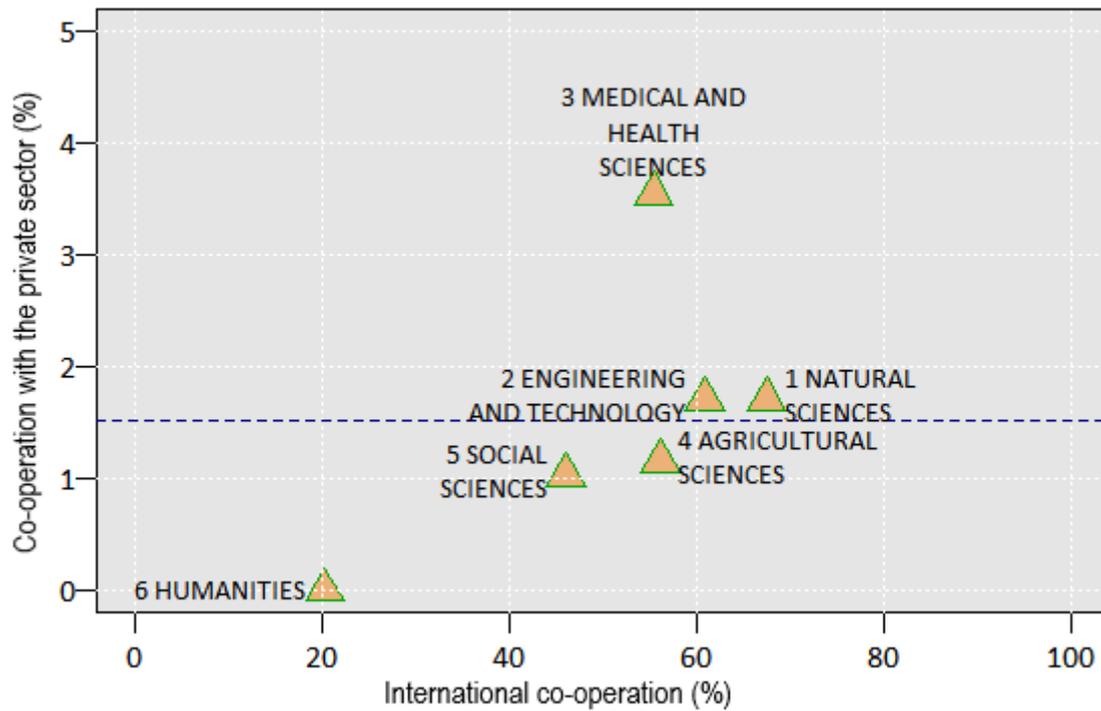


Figure 8 The shares of the most important partner countries from the national output produced in international collaboration for the main disciplines (fractional counting)

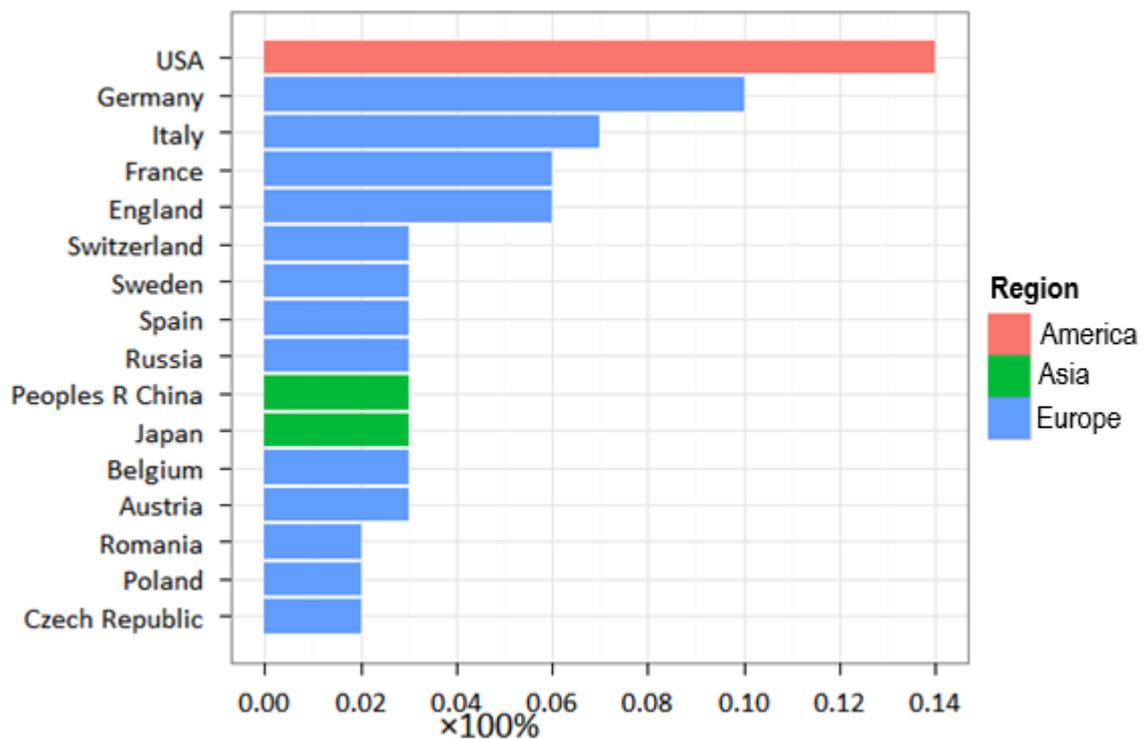


Figure 9 The shares of the most important partner countries from the national output produced in international collaboration for the sub-disciplines (fractional counting)

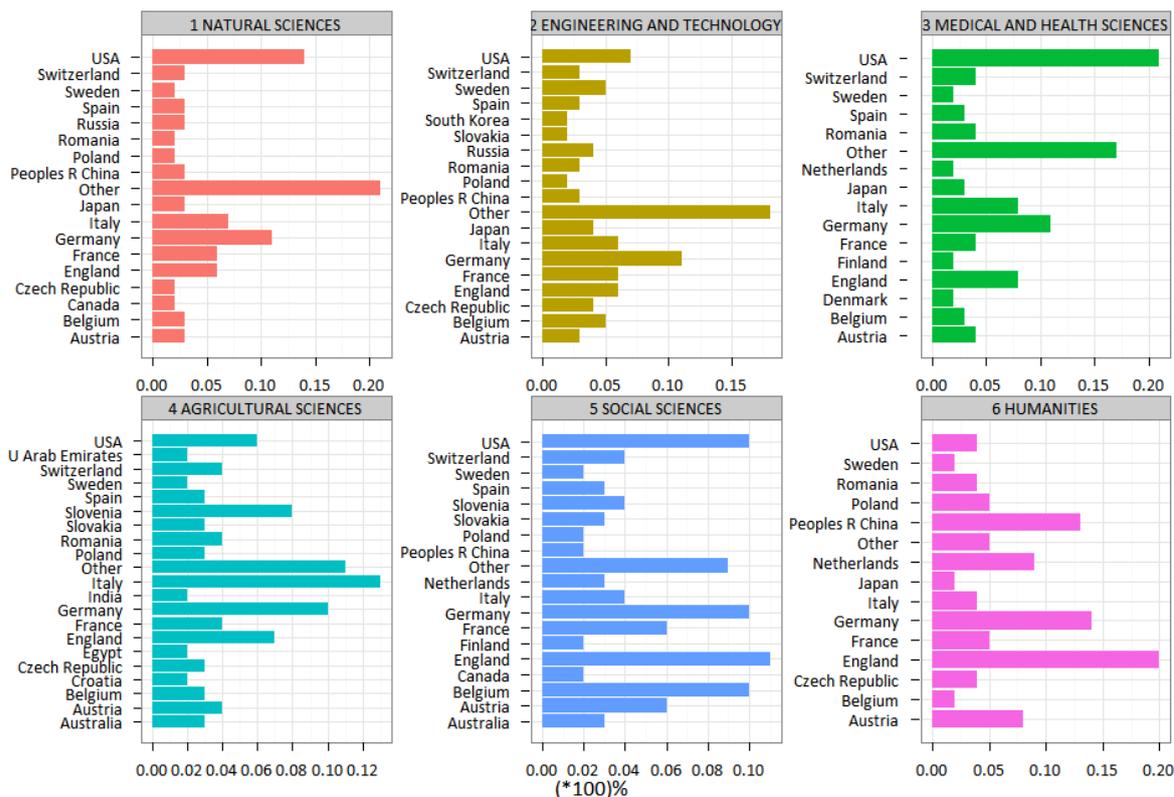
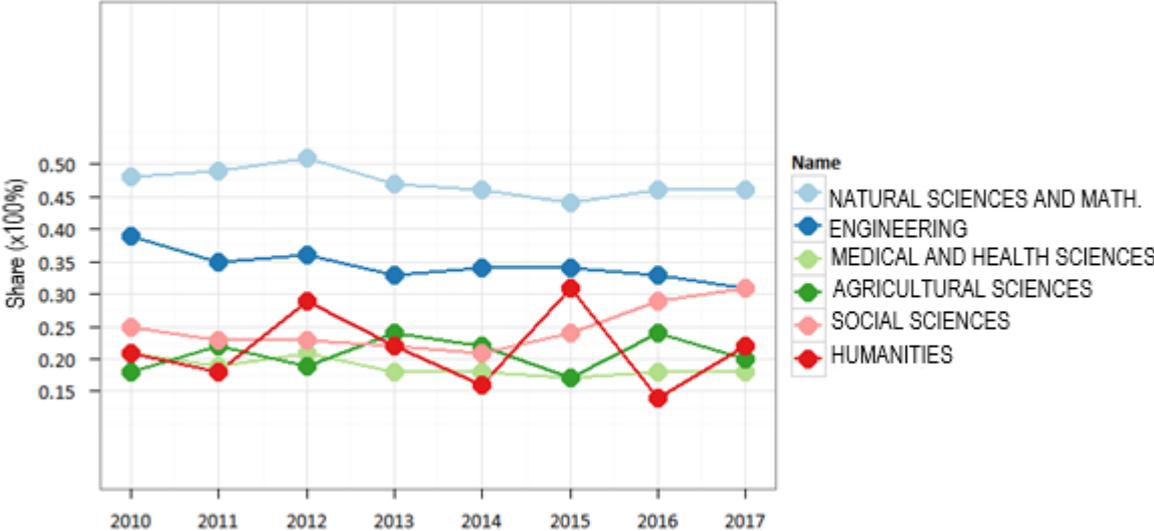


Figure 10 The average disciplinary volume and variability of countries

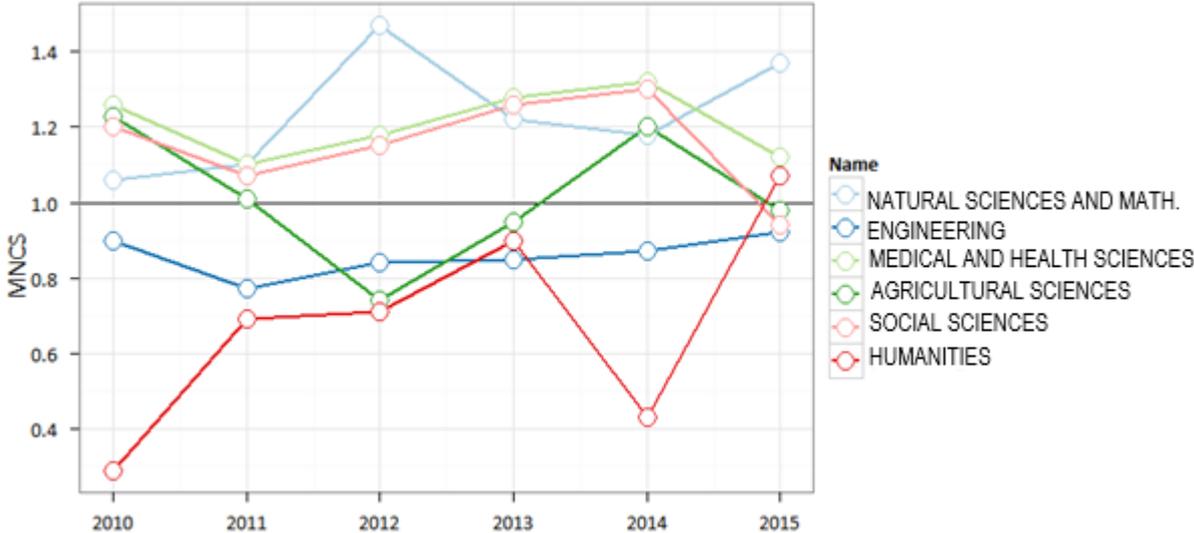
Country	Average	SD
Germany	0,11	0,01
England	0,1	0,05
USA	0,1	0,06
Italy	0,07	0,03
Slovenia	0,06	0,03
Austria	0,05	0,02
France	0,05	0,01
Netherlands	0,05	0,04
Peoples R China	0,05	0,05
Belgium	0,04	0,03
Russia	0,04	0,01
Switzerland	0,04	0,01
Australia	0,03	0
Czech Republic	0,03	0,01
Japan	0,03	0,01
Poland	0,03	0,01
Romania	0,03	0,01
Slovakia	0,03	0,01
Spain	0,03	0

# Additional Figures: the Structure of the MTA Output on a National Level (background material for the MTA government report)

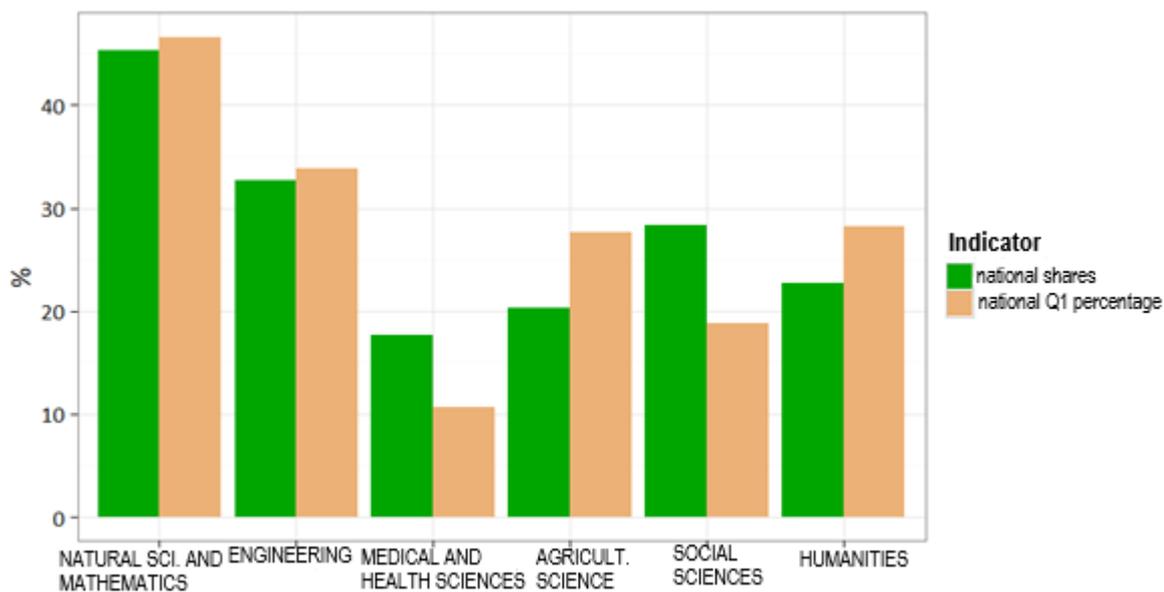
1. Proportions of the MTA research network from the national output of the major disciplines, 2010-2017 (Web of Science data, OECD classification, x100%)



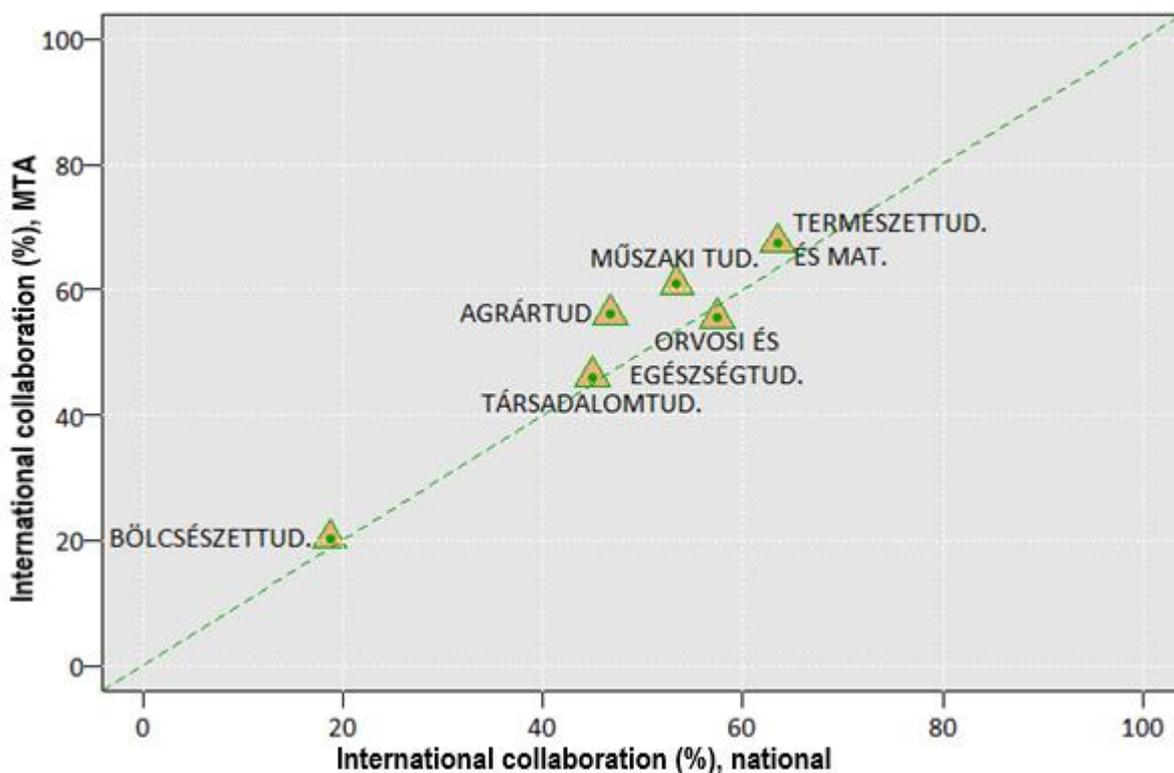
2. Citation impact of the MTA research network for the major disciplines, 2010-2015 (Web of Science data, OECD classification, x100%. The citation indicator is the number of citations normalized with regard to year and discipline, the reference line, MNCS=1, denotes the international average).



3. The shares of the MTA research network from the national output published in Q1 journals for the past three years (2015-2017) in the major disciplines (Web of Science data, OECD classification, x100%)

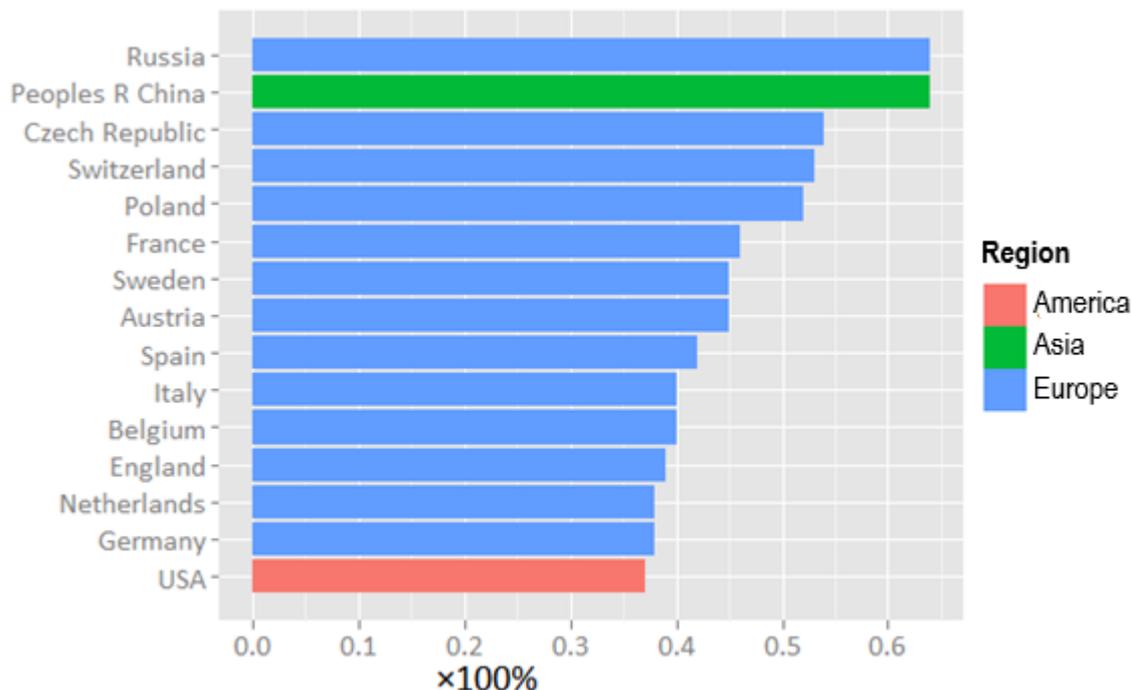


4. Proportions of publications produced in international collaboration for the MTA and the overall national output in the last three years (2015-2017) for the major disciplines (Web of Science data, OECD classification). The two values are the same for the disciplines appearing along the diagonal line, above/below results mean that the values are higher/lower for the MTA.

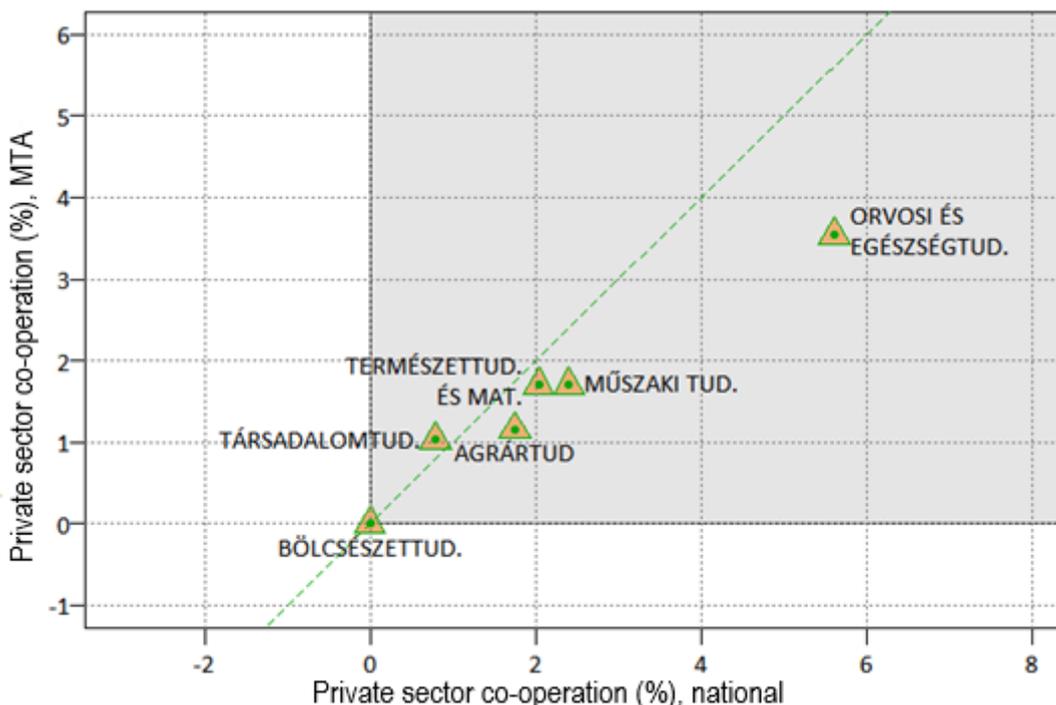


Disciplines in English: Bölcsészettud. – Humanities; Társadalomtud. - Social Sciences; Orvosi és Egészségtud. - Medical and Health Sciences; Agrártud. - Agricultural Sciences; Műszaki tud. – Engineering; Természettud. és Mat. - Natural Sciences and Mathematics

5. Proportion of MTA publication in the international co-authored output with regard to the most important partner countries (proportion of co-authored publications  $\geq 10\%$ ) for the past three years (2015-2017, Web of Science data, OECD classification)



6. The proportion publications for the major disciplines in collaboration with the private sector within the MTA total output and the national output for the past three years (2015-2017) based on Web of Science data and using the OECD classification. For the disciplines along the diagonal line the two scores are the same, above/below means higher/lower results for the MTA.



Disciplines in English: Bölcsészettud. – Humanities; Társadalomtud. - Social Sciences; Orvosi és Egészségtud. - Medical and Health Sciences; Agrártud. - Agricultural Sciences; Műszaki tud. – Engineering; Természettud. és Mat. - Natural Sciences and Mathematics

## Summary

With regard to the analysis above, the following statements can be made about the publication output of the MTA research network for the 2015-2017 period:

- There is significant difference between quantitative and qualitative structure of the network's internationally visible output (publications indexed in WoS). Considering the national output, natural sciences have the highest share (~45%), followed by engineering and social sciences (~30% each), then agricultural sciences and medical sciences (~20% each, due to co-authorship the sum of the proportions is more than 100%). As opposed to this, the qualitative measures of the output are high for all disciplines, especially for medical and health sciences, natural sciences, engineering and agricultural sciences (proportion of Q1 publications is 60-70%). This proportion is also high for social sciences, almost half of the total output is Q1.
- With regard to short-term citation impact, each discipline scores over the EU13, just below or slightly over the EU28 and over the international average. More than 10% of the output of natural sciences, social sciences and medical sciences belong to the most cited 10% (above the internationally expected result). The other three major disciplines (humanities, agricultural sciences, engineering) score 5-10%.
- The importance and expanse of the co-author network can be revealed through the fact that the proportion of collaborative publications is not only high in those disciplines that are traditionally region independent or carry out research in teamwork (though these disciplines have the highest results, with 60-70% in natural sciences, engineering and medical sciences) – social sciences also produce relatively high scores (~40-50%), and even humanities reach (~20%). Based on MTA publications, the most important partner countries are the USA and Germany, succeeded by Italy, France and England, but significant collaborations are carried out with eastern, Far Eastern and East Central European countries as well (Japan, China, Russia and post-socialist countries). The volume of these countries, however, differs for each discipline (except for Germany, the most important partner). Additionally, MTA collaborations characteristically focus on contact with eastern and East Central European countries (as well as Switzerland), this group is followed by Western European countries and the USA with 10-20% (average proportions for the formerly mentioned countries range between 40 to 70%).
- The intensity of private sector co-operation (academy & industry relations) is highest in medical and health sciences (3.5% of the publications), closely followed by natural sciences and engineering (1.5-2%), then agricultural and social sciences (around 1%).

## Appendix: Frascati (OECD) classification (Fields of Science, FOS)

<b>Abbrev</b>	<b>FOR code</b>
<b>1 NATURAL SCIENCES</b>	
Math	1.01 Mathematics
Comp-info	1.02 Computer and information sciences
Phys-ast	1.03 Physical sciences and astronomy
Chem	1.04 Chemical sciences
Earth-env	1.05 Earth and related environmental sciences
Biol	1.06 Biological sciences
Other	1.07 Other natural sciences
<b>2 ENGINEERING AND TECHNOLOGY</b>	
Civil	2.01 Civil engineering
Electrical	2.02 Electrical eng, electronic eng
Mech	2.03 Mechanical engineering
Chem-e	2.04 Chemical engineering
Materials	2.05 Materials engineering
Medical	2.06 Medical engineering
Envir	2.07 Environmental engineering
Envir-bio	2.08 Environmental biotechnology
Indust-bio	2.09 Industrial biotechnology
Nano	2.1 Nano-technology
Other	2.11 Other engineering and technologies
<b>3 MEDICAL AND HEALTH SCIENCES</b>	
Basic	3.01 Basic medical research
Clinical	3.02 Clinical medicine
Health	3.03 Health sciences
<b>4 AGRICULTURAL SCIENCES</b>	
Agri-ff	4.01 Agriculture, forestry, fisheries
Animal-d	4.02 Animal and dairy science
Vete	4.03 Veterinary science
Other	4.05 Other agricultural science
<b>5 SOCIAL SCIENCES</b>	
Psych	5.01 Psychology
Econ-bus	5.02 Economics and business
Edu	5.03 Educational sciences
Soci	5.04 Sociology
Law	5.05 Law
Polit	5.06 Political science
Soc-e-geo	5.07 Social and economic geography
Medi-com	5.08 Media and communication
Other	5.09 Other social sciences
<b>6 HUMANITIES</b>	
Hist-arch	6.01 History and archaeology
Lang-lit	6.02 Languages and literature
Phil-reli	6.03 Philosophy, ethics and religion
Art	6.04 Art
Other	6.05 Other Humanities

