

## GLOBAL INSTITUTIONAL PROFILES PROJECT

# Data Collection Guidebook—General Notes

## SECTION 2 – GENERAL NOTES

The more data you are able to provide, the clearer and fuller your institution profile will be. This facilitates a more meaningful comparison between institutions. Some data are essential if we are to create a complete and comparable profile for your institution. Other data will help to ensure that we provide a richer profile across subjects and years. Please make note of the below minimum data requirements.

### ESSENTIAL INFORMATION - (minimum requirements)

The minimum requirements for data submission to the *Institutional Profiles* project are:

- Institutional and contact information
- “OVERALL” data about your activities
- 1 full year of data (recent-most year available in data collection tool)
- Minimum of six priority data points across each of the six subject areas

Institutional and contact information should be as complete as possible. This information will be a valuable resource for people wishing to find out more about your institution.

University departmental and affiliation information should be as complete as possible. This information will assist Clarivate Analytics in the ongoing [Name Unification Process](#) whereby paper and citation counts are properly attributed.

- Number of academic staff
  - Of which are international/overseas origin
- Number of research staff
- Number of students
  - Of which are international/overseas origin
- Undergraduate - new student intake
- Number of undergraduate degrees awarded
- Doctoral - new student intake
- Numbers of Doctorates awarded
- Institutional income
- Research income
  - Research income from industry and commerce

**Note:** data field definitions are available in Section 3 — [Data Definitions](#)

### SUBJECT BREAKDOWNS

In addition to the OVERALL section, data in the six subject areas are essential for us to generate meaningful performance indicators. Failure to supply the minimally required subject level data may

result in your institution being left out of subject specific comparisons, and will prevent you from showcasing your institution's areas of strength.

We are committed to overcoming subject bias in our comparisons, and will use “normalized” data for many indicators. Each subject often possesses different characteristics. For example, research in the medical sciences typically has more funding than research in the social sciences. This reflects the costs of undertaking research, such as the requirements of research materials and infrastructure. Normalizing data allows us to account for varying subject characteristics and represent institutions comparatively regardless of their subject mix. If you do not provide subject level data, we may utilize approximations that may not represent your institution favorably.

As an absolute minimum, please provide data for the following six fields in each of the six subject areas:

- Number of academic staff
- Number of research staff
- Number of students
- Number of undergraduate degrees awarded
- Numbers of Doctorates awarded
- Research income

If you do not have activity in a subject area, you should enter “0” rather than leaving it blank. For example if you have academic staff only in the Social Sciences and Arts & Humanities, you should enter the correct numbers for those fields and “0” for the other fields. Leaving a field blank tells us that you choose not to supply the data whereas entering a “0” tells us that you do not have activity for that subject.

## RETROSPECTIVE DATA

The data collection tool provides columns for you to enter five years of data. By submitting data for all five years you will provide a more meaningful impression of the changing nature of your institution. This is useful for identifying trends. As an absolute minimum you must complete the OVERALL and subject level minimum requirements for the most recent year.

## FAILURE TO MEET MINIMUM REQUIREMENTS

If you are having difficulty supplying the minimally required data, please contact your regional data editor or send an inquiry to [science.profilesproject@thomsonreuters.com](mailto:science.profilesproject@thomsonreuters.com).

To assist you with data collection process, we have also provided regional support materials such as subject mapping tables for local subject classification schema. These are available at: <http://ip-science.thomsonreuters.com/globalprofilesproject/datacollection/support/>.

Failure to meet the minimum requirements may mean that we will have to rely on estimations of your data. The result may not be a fair representation of your institution since, where there is an element of doubt; we will err on the side of caution and use the lower estimate.

## ANNUAL UPDATES

The Global Institutional Profiles Project data collection process is an annual activity. Information that you submit this year will be retained in the system and will be used to pre-populate the historical record of your profile for future updating activities.

## PRE-POPULATED DATA

You may find that some fields have already been populated with data. This data has been gathered on your behalf from publicly available sources such as IPEDS-USA or HESA-UK. It is essential that you confirm the accuracy of this information before submitting your completed profile. If you are unclear on the source of this data please contact your local data editor for details.

## AFFILIATED INSTITUTIONS AND HOSPITALS

We recognize that many institutions are complex organizations with many constituent parts such as overseas campuses and affiliated hospitals and that it is frequently difficult to distinguish these independently.

To guide your decisions about the inclusion or exclusion of data relating to these affiliated institutions we recommend that you consider whether the affiliated institution is included or excluded from your existing Annual Reports and Financial Reports and follow the same approach to reporting of data here. Once you make that decision, efforts should be made to be consistent with all data reporting. For example if you excluded an overseas campus for reporting the amount of income, you should also exclude it from reporting of student and staff numbers.

Often times it is impossible to distinguish the activities of affiliated hospitals from the activities of the university. It is acceptable to include affiliated hospitals within the various data points that are reported, but this must be done consistently. Note that Academic Staff (see the data definitions document) should NOT include clinicians from affiliated hospitals unless they also have an academic post and a significant component of their workload is dedicated to traditional academic activities such as teaching or research.

## PUBLICATIONS

We are not asking you to supply data about your institution's publications. We will use Clarivate Analytics' in-house systems to generate data about your publications and citations to those publications. This is designed to be a fair system that will minimize any potential to manipulate the statistics.

If you wish to inform us about potential name variants for your institution you may do so in the "Institution's name variants" field of the INSTITUTIONAL & CONTACT INFORMATION page.

Additional information on the name unification process that is used to identify all the publications from your institution is available at

<http://ip-science.thomsonreuters.com/m/pdfs/UnifyingNameVariants.pdf>

You may be contacted by Clarivate Analytics staff for additional information about your affiliations.

## LANGUAGE

All textual information must be entered in English.

The information you enter will be made publicly available in many different countries to readers with a wide range of languages. If you enter all your text in clear English then you will maximize the visibility of your institution. You may enter the local language name variant of your institution in the “Institutional name variants” field of the INSTITUTIONAL & CONTACT INFORMATION page, however, please enter this name using the Romanized alphabet.

## REPORTING OF NUMBERS

To avoid confusion that may be created by varying treatments of thousand separators and decimal places, please do not use any thousand separators when reporting large numbers.

Numbers should be in the format: 1000.23

Do not use the formats: 1,000.23 or 1.000,23 or 1 000.23

You do not need to use decimal points when reporting whole numbers. In fact we do not expect the use of decimal points in most numbers that you will report. Please see below for details of the level of estimations.

## ESTIMATIONS

We recognize that it is not practical to supply exact figures for some of the data fields. We ask that you input a minimum of 3 significant figures (also known as *significant digits*). This means that there should be a minimum of 3 significant figures in all numbers regardless of where the decimal point falls.

For example, a large number such as 256332.1 can be reported as 256000

Alternatively, 76.48 can be reported as 76.5

In both cases the 3rd figure is rounded to the nearest whole number.

You may supply more detailed data should you wish. We will include the detailed data in the completed profiles. However, most comparisons will be based on 2 or 3 digit scoring system and more detailed data will not make any significant difference to the end results.

To learn more, visit:

<http://ip-science.thomsonreuters.com/globalprofilesproject/>