Gender Breakdown in Irish Research Funding - SFI Grants

Author: Derek O'Callaghan (@dgocallaghan)
Revision: 1.2
Date: 2018-11-22

Contents

1 Executive Summary **3**

2 Introduction **5**

3 SFI Grants Overview **6**

4 SFI Programmes **9**

5 SFI Research Centres **14**

1 Executive Summary

Irish research funding bodies such as Science Foundation Ireland (SFI), the Irish Research Council (IRC) and the Health Research Board will *"require higher education Institutions to have Athena SWAN Gender Equality accreditation in order to be eligible for research funding"*[1]. In a major national initiative supported by the Higher Education Authority, the Athena SWAN Charter was launched in Ireland in early 2015, where *"figures published by the Higher Education Authority highlighted gender inequality as an issue for the sector"*[2]. In light of this, it is interesting to take a took at existing datasets to get an idea of specific gender breakdown details. SFI has provided a dataset that *"details STEM research and ancillary projects funded by Science Foundation Ireland since its foundation in 2000"*[3]. This is a useful dataset for such an exercise, and is the focus of this exploratory analysis. It is envisaged that similar analysis will be performed with other available datasets.

In this analysis (see here for full details[4]), the comma-separated values (CSV) file[5] provided by SFI (accessed 2018-10-09 and 2018-11-12, no changes at latter date) has been used as the source dataset, where each entry in the dataset corresponds to a particular successful grant. Gender details of the applicants are not currently provided in the SFI data. Consequently, a *Gender* column was added by means of a manual analysis of the unique *Lead Applicant* names. As more than a thousand applicants are in the dataset, the column values are currently restricted to *Male* and *Female*. However, this could be extended as required. It is emphasised that this is a best effort at the present time, where a small number of inaccuracies may exist. In addition, it is recommended that this process be validated by SFI, by providing updated data that includes gender details, along with similar data (anonymised) for all original applicants.

There are currently 359 (23%) female and 1,203 (77%) male successful applicants in the SFI data, who received 910 (20%) and 3,569 (80%) grants respectively from 2001 to 2017 (note: only four 2017 grants are included in the data). The corresponding grant totals were EUR 0.3 billion (13%) and EUR 2.1 billion (87%) awarded to female and male applicants respectively. The top 20 applicants ranked according to aggregated grant totals includes a single female and 19 males, who received 7 (4%) and 183 (96%) grants respectively, with corresponding grant totals of EUR 56.07 million (7%) and EUR 716.69 million (93%). Additional details on the number of grants and corresponding grant totals awarded to universities may be found in the full analysis.

[1] http://www.sfi.ie/research-news/news/irish-funding-bodies-to-require-athena-swan-gen

[2] https://www.ecu.ac.uk/equality-charters/athena-swan/athena-swan-ireland/

[3] https://www.sfi.ie/about-us/governance/open-data/

[4] https://github.com/derekocallaghan/genderbreakdown/blob/master/Gender%20Breakdown%20-%20SFI.ipynb

[5] https://www.sfi.ie/about-us/governance/open-data/Open-Data-to-30-11-2016-for-Publica csv

Each entry in the SFI data is associated with a particular *Programme Name*. The full analysis contains an overview of the gender breakdown for each of the 60 unique programmes, by means of aggregated grant totals plots. For most programmes, male applicants have received the majority of the corresponding grant totals, with female applicants more successful in a minor number of programmes having lower aggregated grant totals. For example, looking at the "SFI Discover" programme, 77 (56%) and 61 (44%) grants were awarded to female and male applicants respectively, with corresponding grant totals of EUR 3.25 million (51%) and EUR 3.12 million (49%). Grants for certain programmes were awarded exclusively to male applicants, where notable examples include the "SFI Research Professorship" programme, with EUR 72.65 million awarded to 24 male applicants.

The *Lead Applicant* column in the SFI data is restricted to a single applicant name per grant, and does not specify other people involved in the grant application. However, further details may be found by looking at the SFI Research Centres[6]. As each centre typically provides a website containing researcher profiles, a separate dataset was created, where the researcher HTML pages were retrieved for 16 of the 17 current centres (accessed 2018-10-12 and 2018-11-12), and a corresponding *Centre*, *Name*, *Role*, *Gender* record was created for each researcher parsed from the website pages. *Role* values are typically provided in the websites, while *Gender* was populated using the same process employed for the original SFI data. The *Role* values vary between centres, with certain roles not always specified. Consequently, the *Role* column has between restricted to the *Funded Investigator* $->$ *Director* roles (or similar). It can be assumed that these indicate potential candidates for *Lead Applicant* in the SFI dataset.

Grant details for only 12 centres (and their lead applicants) are provided in the SFI data. Of these, 2 centre applicants are female, with corresponding grant totals of EUR 39.36 million (11%) and EUR 320.71 million (89%) awarded to female and male applicants respectively (excluding supplements and related grants). It is noted that one of the two centres originally awarded to female applicants is currently led by a male director, according to the corresponding centre website. Of the other five centres for which details are not included in the SFI data, four are currently led by male directors, with data unavailable for the final centre. According to the website data, centre researchers having *Investigator* or *Director* roles (or similar) include 113 (19%) female and 490 (81%) male researchers.

[6] `http://sfi.ie/sfi-research-centres/`

2 Introduction

Irish research funding bodies such as Science Foundation Ireland (SFI), the Irish Research Council (IRC) and the Health Research Board will *"require higher education Institutions to have Athena SWAN Gender Equality accreditation in order to be eligible for research funding"*[7]. In a major national initiative supported by the Higher Education Authority, the Athena SWAN Charter was launched in Ireland in early 2015, where *"figures published by the Higher Education Authority highlighted gender inequality as an issue for the sector"*[8].

In light of this, it is interesting to take a took at existing datasets to get an idea of specific gender breakdown details. SFI has provided a dataset that *"details STEM research and ancillary projects funded by Science Foundation Ireland since its foundation in 2000"*[9]. This is a useful dataset for such an exercise, and is the focus of this exploratory analysis. It is envisaged that similar analysis will be performed with other available datasets.

This document provides a summary of the current analysis, where full details, including all steps used to generate the outputs may be found here[10]. This summary is structured as follows:

- Section 3 contains an overview of the SFI data and corresponding gender breakdown.
- Section 4 provides details based on individual SFI programmes.
- The SFI research centres, including corresponding data not currently provided by SFI, are discussed in Section 5

[7] http://www.sfi.ie/research-news/news/irish-funding-bodies-to-require-athena-swan-gen

[8] https://www.ecu.ac.uk/equality-charters/athena-swan/athena-swan-ireland/

[9] https://www.sfi.ie/about-us/governance/open-data/

[10] https://github.com/derekocallaghan/genderbreakdown/blob/master/Gender%20Breakdown%20-%20SFI.ipynb

3 SFI Grants Overview

In this analysis, the comma-separated values (CSV) file[11] provided by SFI (accessed 2018-10-09 and 2018-11-12, no changes at latter date) has been used as the source dataset. SFI also provide a corresponding Explanatory Note[12]. Each entry in the dataset corresponds to a particular successful grant, with the grant totals in the *Revised Total Commitment (including overhead)* column. The explanatory note states that *"This is the total current value of the award (direct costs) plus overhead (a contribution to the host Research Body towards the indirect costs of hosting SFI-funded research programmes)."*

Gender details of the applicants are not currently provided in the SFI data. Consequently, a *Gender* column was added:

- This was achieved by means of a manual verification of the unique first names of the *Lead Applicant* column.
- **As more than a thousand applicants are in the dataset, the column values are currently restricted to *Male* and *Female*. However, this could be extended as required.**
- For the names where the gender was unclear, it was determined by checking publicly available data, for example, university profiles or photos.
- In cases where the *Lead Applicant* was *"VP/Dean of Research..."*, the gender was either that of the *Lead Applicant* at the time, or that of the person mentioned in the *Title of Research* column (if available).

Note: the *Gender* column has been manually validated a number of times. However, it should be emphasised that this is a best effort at the time, where a small number of inaccuracies may still be present. Ideally, it would be useful if this can be validated by SFI, where they would provide updated data with gender details, along with similar data (anonymised) for original applicants. Availability of the latter data would permit normalisation and comparison of unsuccessful and successful applicant rates

As the original SFI data *"is available to use free of charge and licensed under a Standard Irish PSI Licence (CC-BY 4.0)"* (see also Irish Open Data license[13] and CC BY 4.0[14]), the updated CSV has been made available here[15].

[11] https://www.sfi.ie/about-us/governance/open-data/Open-Data-to-30-11-2016-for-Publica
csv

[12] https://www.sfi.ie/about-us/governance/open-data/Science-Foundation-Ireland-Grant-Co
docx

[13] https://data.gov.ie/pages/opendatalicence

[14] https://creativecommons.org/licenses/by/4.0/legalcode

[15] https://raw.githubusercontent.com/derekocallaghan/genderbreakdown/
master/data/20181112/20181112_gender_Open-Data-to-30-11-2016-for-Publication-v2.
csv

Table 1 looks at *Gender* according to unique *Lead Applicant*, number of grants, and corresponding grant totals, while Table 2 lists the top 20 *Lead Applicants* according to aggregated grant totals.

Gender	Number of unique Applicants	Number of Grants	Grant Totals
Female	359 (**23%**)	910 (**20%**)	EUR 0.3 billion (**13%**)
Male	1203 (**77%**)	3569 (**80%**)	EUR 2.1 billion (**87%**)

Table 1: Overview of SFI data according to *Gender*. Note: SFI currently only provide details of successful applicants.

Gender	Number of Grants	Grant Total (EUR million)
Male	13	87.84
Male	10	60.05
Female	7	56.07
Male	3	48.68
Male	7	45.55
Male	13	45.11
Male	19	43.83
Male	7	37.23
Male	14	37.22
Male	6	36.89
Male	9	36.74
Male	12	32.47
Male	10	30.09
Male	5	29.27
Male	20	27.06
Male	8	25.12
Male	2	25.03
Male	13	23.35
Male	5	22.61
Male	7	22.54

Table 2: Top 20 *Lead Applicants* according to aggregated grant totals.

Figure 1 plots the number of grants per month from 2001 to 2017 (most recent date in SFI data), with the corresponding grant totals in Figure 2.

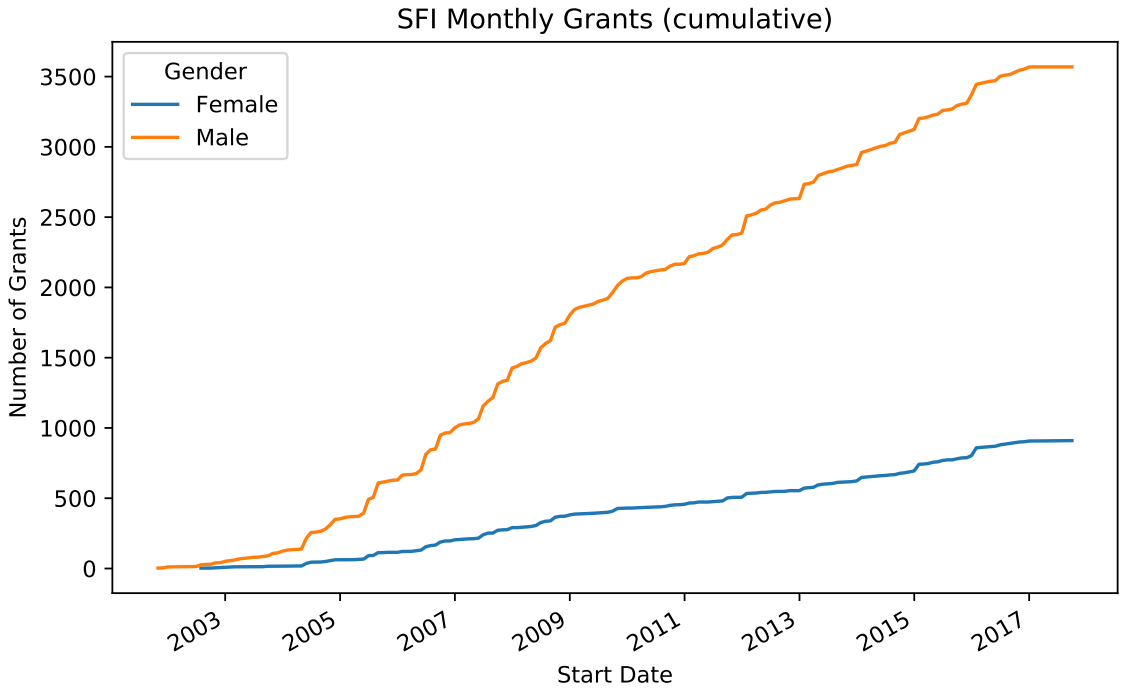

Figure 1: Number of grants per month (cumulative), from 2001 to 2017.

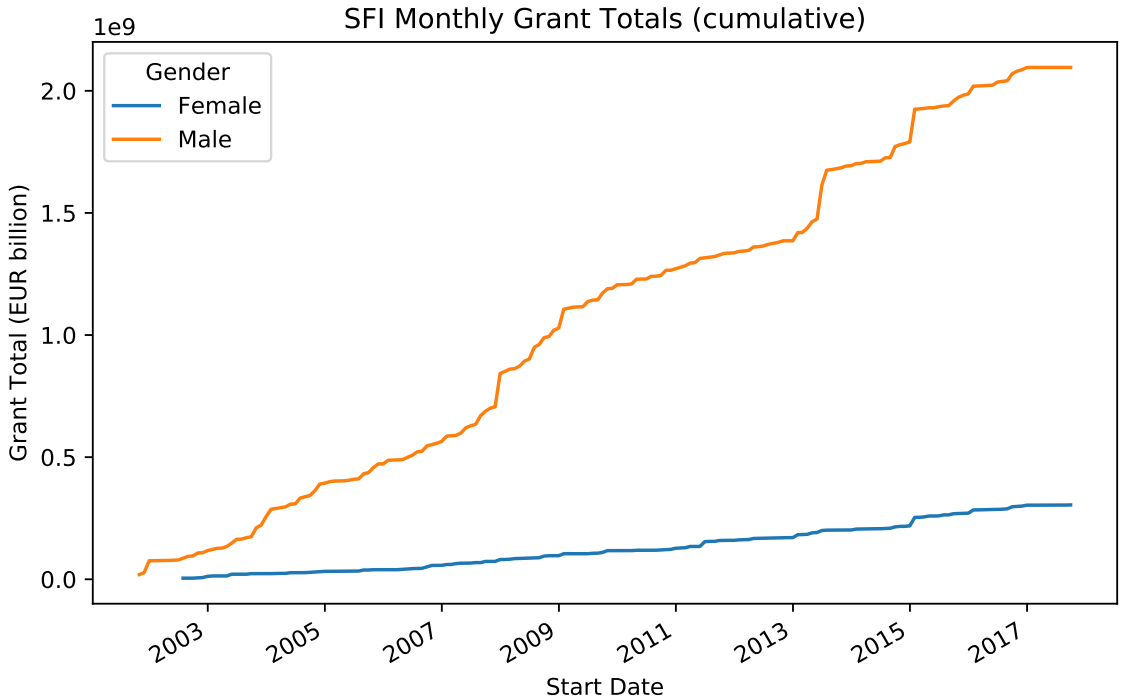

Figure 2: Grant totals per month (cumulative), from 2001 to 2017.

4 SFI Programmes

Each record in the SFI data is associated with a particular *Programme Name*. Figure 3 contains an overview of the gender breakdown per programme, using aggregated grant totals.

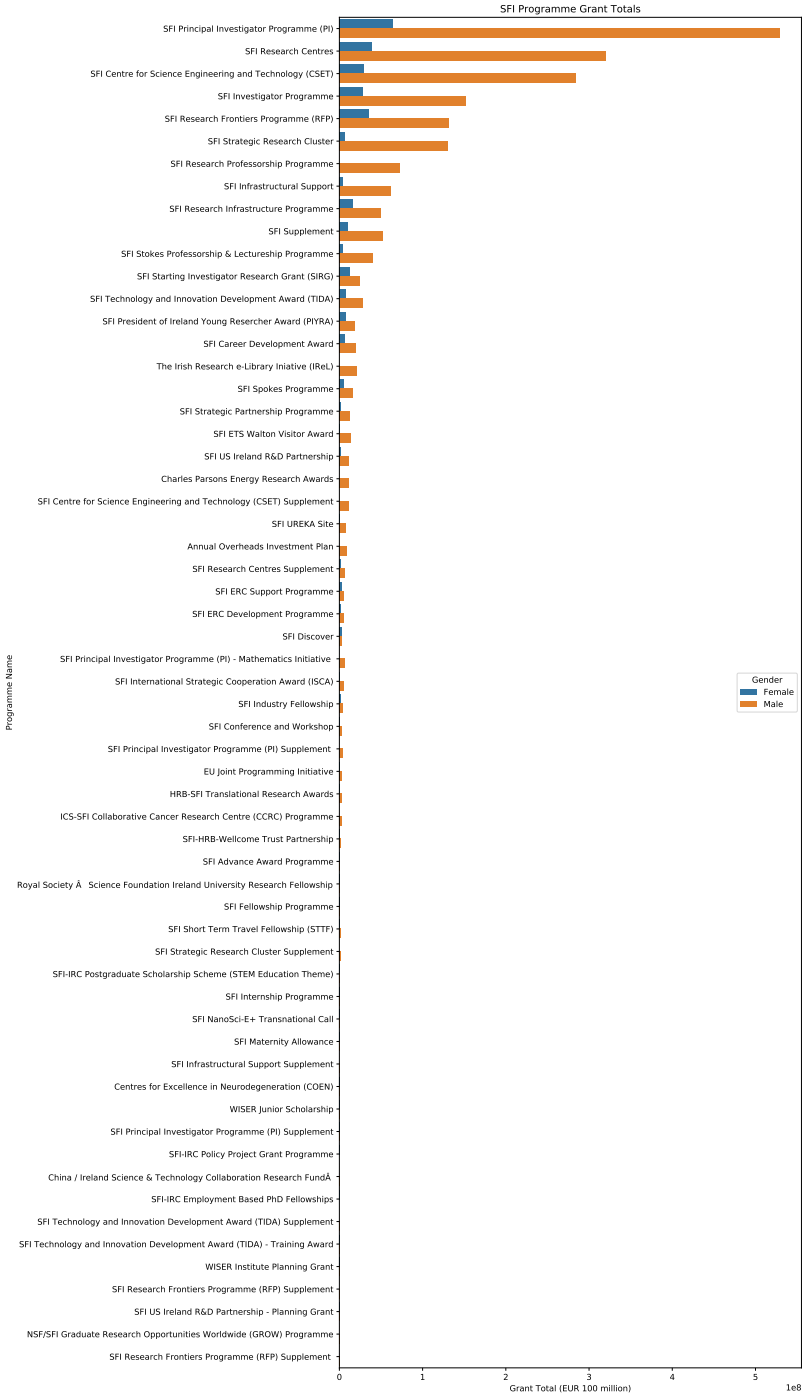

Figure 3: Gender breakdown per SFI programme, using aggregated grant totals.

The single plot in Figure 3 is insufficient due to the number of unique SFI programmes and corresponding aggregated grant totals. Consequently, multiple plots have been created (with different totals ranges), and can be found in Figure 4, Figure 5, Figure 6, and Figure 7.

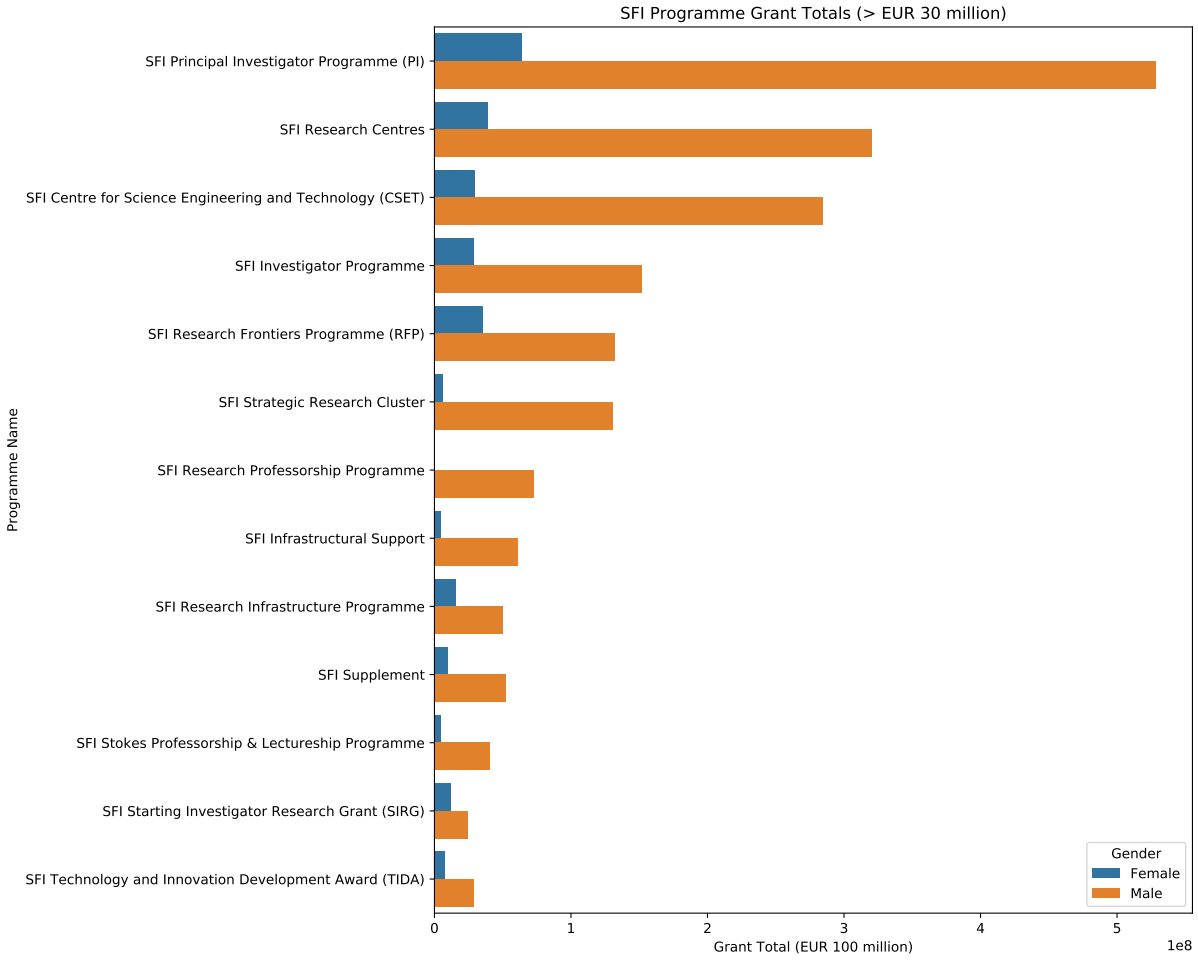

Figure 4: Gender breakdown per SFI programme, focusing on programmes having aggregated totals greater than EUR 30 million.

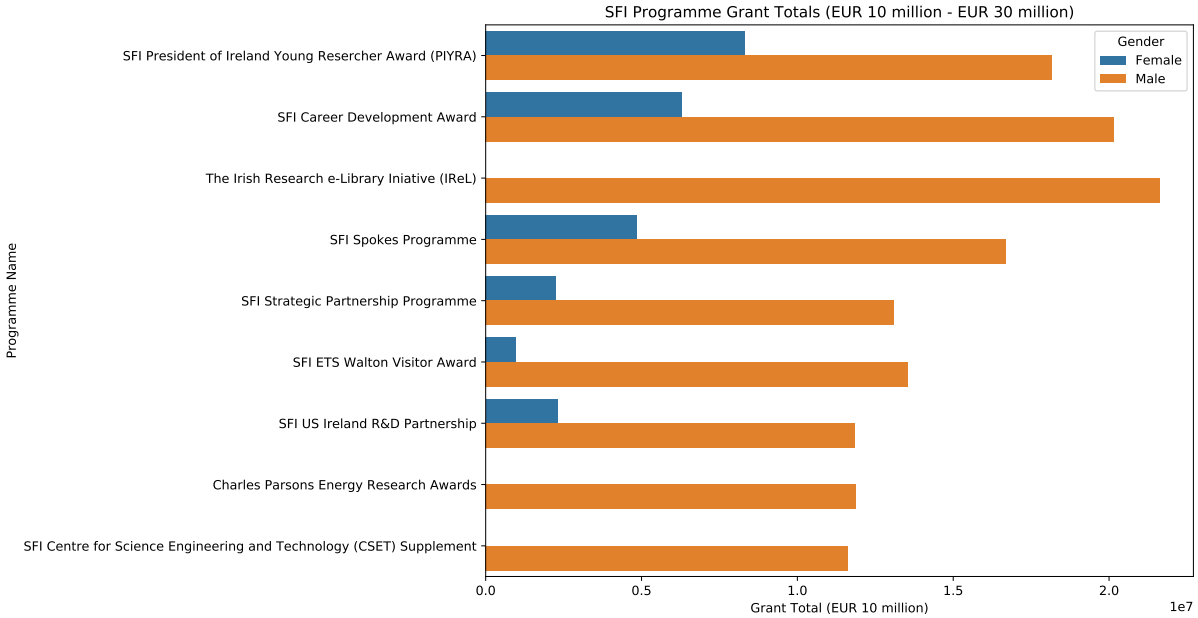

Figure 5: Gender breakdown per SFI programme, focusing on programmes having aggregated totals between EUR 10 million and EUR 30 million.

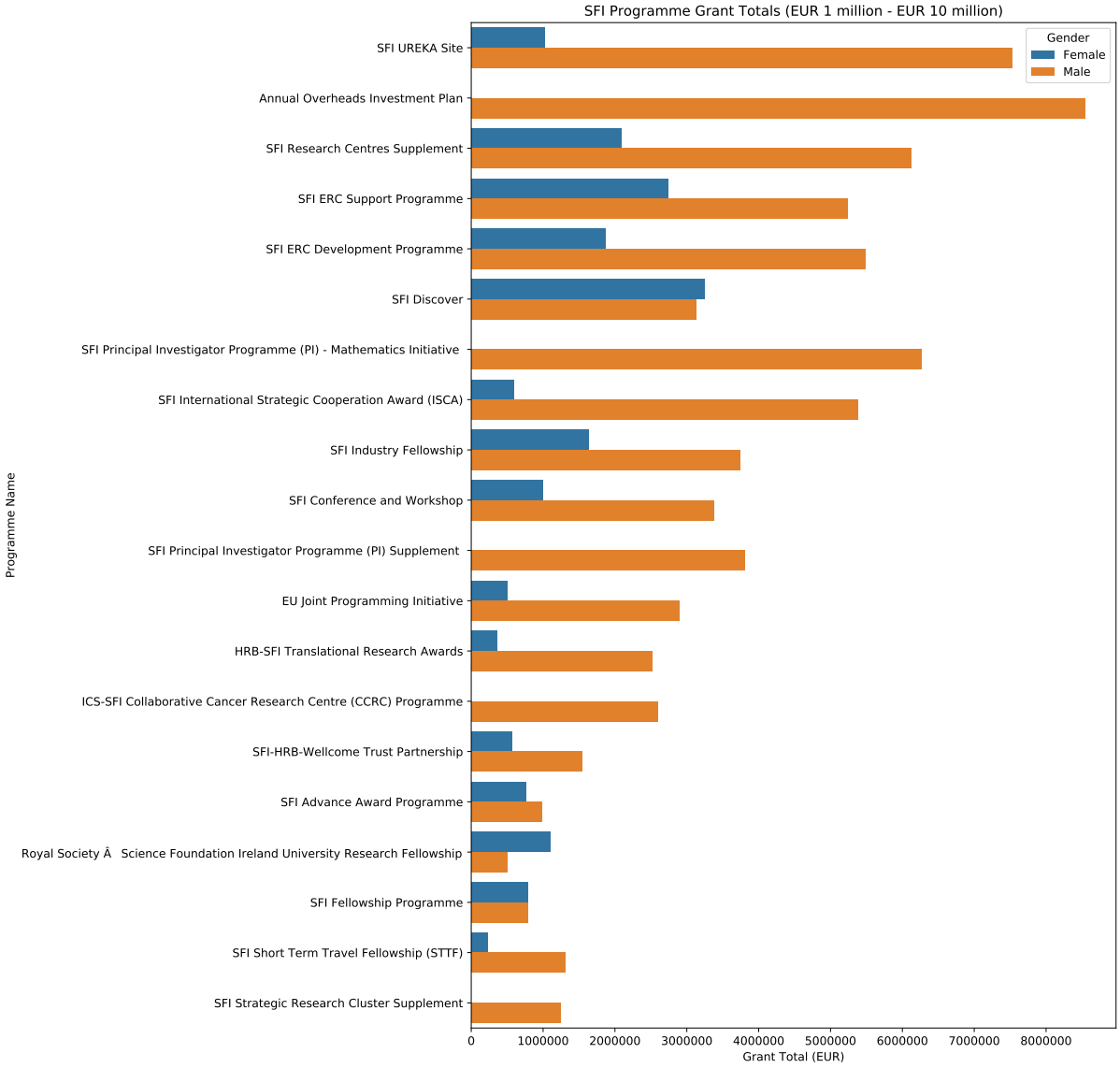

Figure 6: Gender breakdown per SFI programme, focusing on programmes having aggregated totals between EUR 1 million and EUR 10 million.

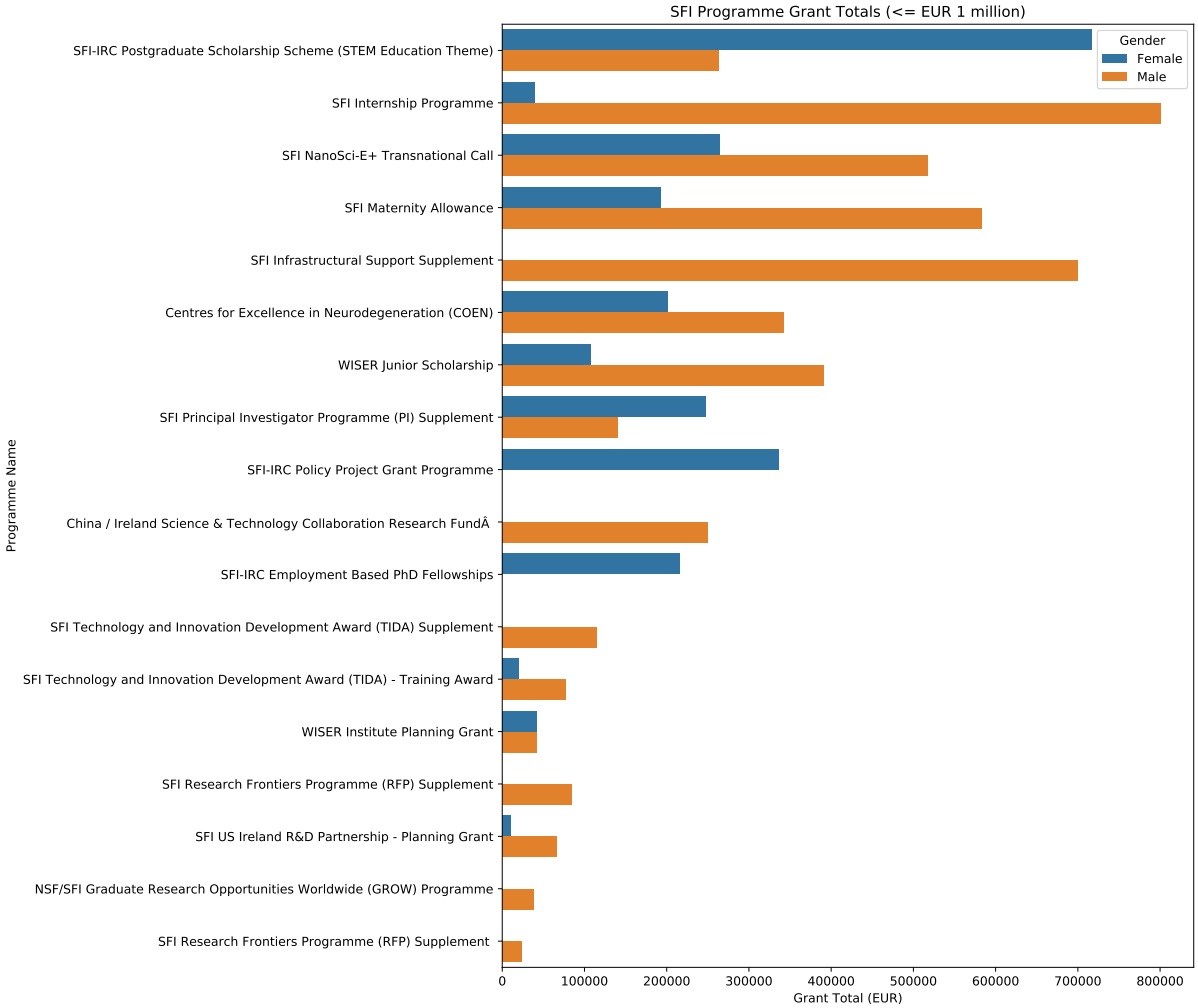

Figure 7: Gender breakdown per SFI programme, focusing on programmes having aggregated totals up to EUR 1 million.

5 SFI Research Centres

As mentioned in Section 3, the SFI dataset contains a *Lead Applicant* column, where other people involved in the grant are not specified. According to the Explanatory Note[16]: *"The Lead Applicant is the person who holds and is responsible for the award. This is, in the vast majority of cases, the person who applied for and won the grant. However, in a small number of cases, the award may have been transferred to another researcher"*.

However, further details may be found by looking at the SFI Research Centres[17]. As each centre typically provides a website containing researcher profiles, a separate dataset was created, where the researcher HTML pages were retrieved for each centre (accessed 2018-10-12, with a subset also accessed on 2018-11-12), and a corresponding *Centre*, *Name*, *Role*, *Gender* record was created for each researcher parsed from the HTML. *Role* values are typically provided in the HTML, while *Gender* was populated using the manual process described in Section 3. The *Role* values vary between centres, with certain roles not always specified. Consequently, the *Role* column has between restricted to the *Funded Investigator* $->$ *Director* roles (or similar). It can be assumed that each of these indicate potential candidates for *Lead Applicant* in the SFI dataset.

Note:

- On the case of at least one centre, the current director (and gender) is different to the corresponding *Lead Applicant*.
- Each centre is associated with a single *Lead Applicant* in the SFI data. **No further details are provided in the SFI data regarding other researchers involved in the application process and subsequent funding awards within the centres themselves.**
- **Details for only 12 centres are provided in the SFI data.**
- It is possible that certain researchers are missing from this derived dataset, for example, centre researchers who were not listed on the corresponding website pages on 2018-10-12 or 2018-11-12.

Table 3 contains a ranking of the SFI centres according to aggregated grant totals, with a gender breakdown of researchers in Figure 8 based on data from the corresponding websites.

[16] `https://www.sfi.ie/about-us/governance/open-data/Science-Foundation-Ireland-Grant-Co`
`docx`

[17] `http://sfi.ie/sfi-research-centres/`

Centre	Start Date	Lead Applicant Gender	Grant Total (EUR million)
Insight	2013-07-01	Male	57.71
CÚRAM	2015-01-01	Male	37.09
APC	2013-06-01	Male	35.89
AMBER	2013-06-01	Male	35.17
Lero	2015-01-01	Male	32.60
CONNECT	2015-01-01	Female	31.15
ADAPT	2015-01-01	Male	30.71
SSPC	2013-06-01	Male	28.13
iCRAG	2015-01-01	Male	24.89
IPIC	2013-06-01	Male	19.61
MaREI	2013-06-01	Male	18.91
INFANT	2013-06-01	Female	8.21

Table 3: SFI centres ranked according to aggregated grant totals (focusing on the *"SFI Research Centres"* programme in the data). Supplements and other grants are excluded. **Details for only 12 centres are provided in the SFI data.**

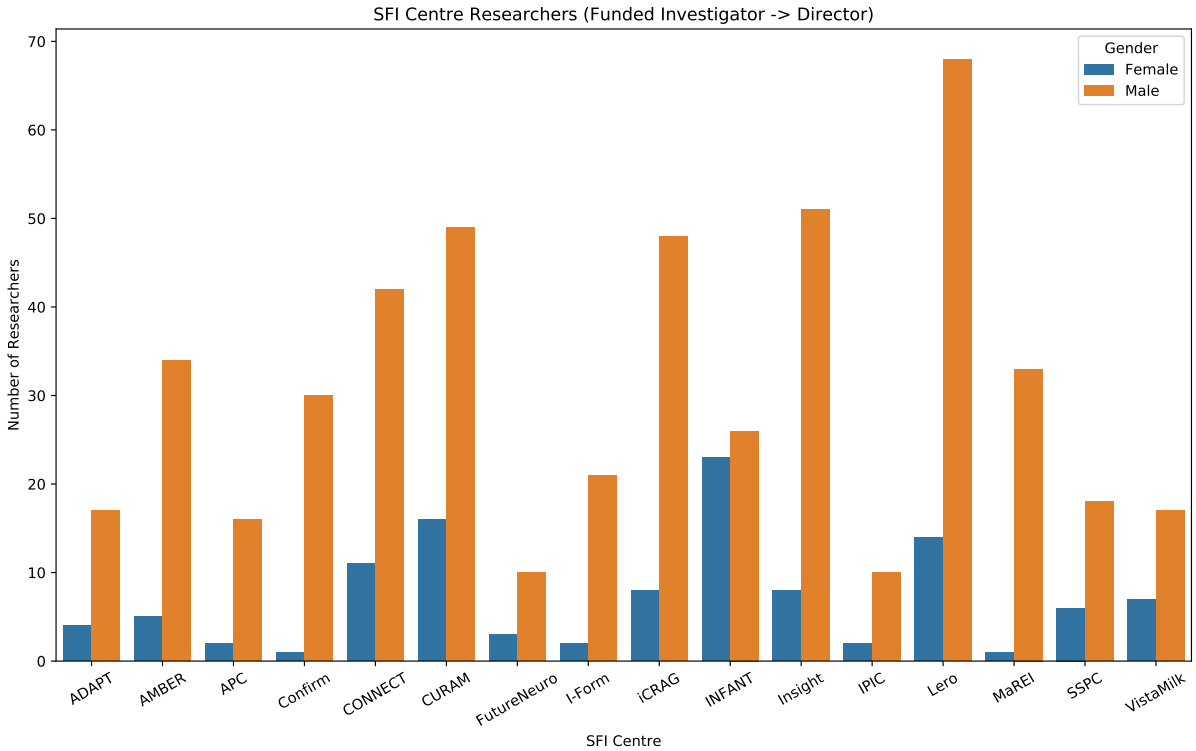

Figure 8: SFI centre researchers, using data from the corresponding centre websites. Note that this includes additional centres not provided in the SFI data (see Table 3). The BEACON centre was excluded as its website PEOPLE page was not working as of 2018-11-12.