

Instructions for creating and managing applications for SCAPIS data

Contents

| | | |
|--------|--|---|
| 1 | Access without creating an account..... | 1 |
| 1.1 | Create an account to apply for data and to follow status for your applications | 2 |
| 1.2 | Log in | 2 |
| 2 | Start an application | 2 |
| 2.1 | Create a new application..... | 2 |
| 2.2 | In general regarding your application for data | 2 |
| 2.3 | Step 1 "Contact form" | 3 |
| 2.4 | Step 2 "Describe the research project" | 4 |
| 2.5 | Step 3 "Approval from ethical review"..... | 4 |
| 2.6 | Step 4 "Population" | 4 |
| 2.7 | Step 5 "View SCAPIS core projects" | 5 |
| 2.8 | Step 6a "Select core variables" | 5 |
| 2.8.1 | Temporary restricted access to some data variables..... | 5 |
| 2.8.2 | Approval required for variables from local sub studies | 5 |
| 2.9 | Step 6b "Select registers" | 5 |
| 2.10 | Step 6c "Select pilot variables" (optional)..... | 6 |
| 2.10.1 | Approval required for variables from the pilot study | 6 |
| 2.11 | Step 6D "Select genotype data" | 6 |
| 2.11.1 | Access to imputed genotype data | 7 |
| 2.11.2 | Access to directly genotyped data | 8 |
| 2.12 | Step 7 "Review and submit" | 8 |
| 3 | See and follow your application in the Portal | 9 |

1 Access without creating an account

Without creating an account, you can visit the SCAPIS Portal at any time and see which variables are available. You can search for variables in the SEARCH field at the top of the page or see the full list of variables categorized into subcategories. For complete information on individual variables, click on the variable and further information will appear. Note here in particular if a variable belongs to a local sub study, as access to such variables will need the approval of the responsible researcher.

1.1 Create an account to apply for data and to follow status for your applications

To apply for data from SCAPIS and to see your application status, you must create an account and log in. Click "CREATE USER" in the field at the top of the page.

A red rectangular button with a white user icon on the left and the text "CREATE USER" in white capital letters.

Register with your name and email address, create a password and accept terms of use and privacy policy.

Please note that attempts to register an account with domains such as gmail or hotmail will be denied. In the event of an application to SCAPIS, the email address will be validated as belonging to a domain that is a potential research authority such as a university or a regional health care authority.

1.2 Log in

After creating an account, log on to the SCAPIS Portal via "LOGIN" in the field at the top of the page. To log in, enter your email address and password.

A red rectangular button with a white right-pointing arrow icon on the left and the text "LOGIN" in white capital letters.

2 Start an application

In logged in mode, start creating an application by clicking on "APPLICATIONS" in the field at the top of the page.

A red rectangular button with a white list icon (three horizontal lines) on the left and the text "APPLICATIONS" in white capital letters.

2.1 Create a new application

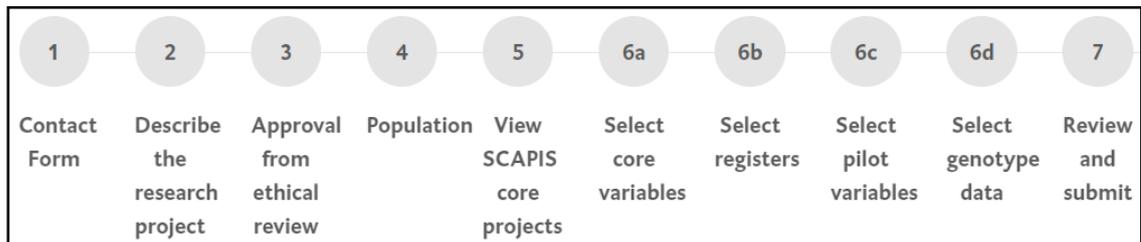
You can create a new application by clicking on "Add new". In next step you enter the title of the project application.

2.2 In general regarding your application for data

The steps of the application

The application consists of seven main steps, where step 6 consist of four sub-steps. The steps are visualized at the top of the application page in a process overview that is exemplified below. Step 6c is only displayed in cases where the SCAPIS pilot population has been selected in

step 4.



You can fill in the form in any order you like although we recommend the given order. You can easily go between the different steps by clicking on the number in the process overview or by progressing sequentially via "Next" and "Back" at the bottom of each page. If you have not already reviewed which data that are relevant to your project, it is recommended to start with reviewing the variables (step 6a for core and 6c for the pilot). This in order to see the variables currently available and whether any of the variables you are interested in belongs to a local sub study.

Language

The application can be written in Swedish or English. However, the title must be in English and in step 2 where the project is described, the "Project summary" must be in English. The project title and summary will be posted on scapis.org upon application approval. The "Popular science summary" should be written in Swedish if possible, and will be published on scapis.se upon application approval.

Clarification of information requested in each step

In each step in the application form there is information or answers that are mandatory, marked with an *. When all mandatory information has been filled in for each step, the step symbol in the process overview turns green. Only when all step symbols are green is it possible to submit the application in step 7. For each mandatory field in the application form, there is more detailed information about what information is to be provided. Read the information carefully.

The following sections clarify certain parts of each section in the application form:

2.3 Step 1 "Contact form"

The Principal Investigator (PI) for the applying project must be a PhD and must be included in the approval from the Swedish Ethical Review Board. If the person writing the application is not the person responsible for the research, the applicant must answer the question "Is the applicant the principal investigator?" by choosing "No" where upon fields to fill in information about the applicant appear.

We ask for personal identity number for the applying project's PI as it will be required for the data delivery process if your application is approved.

Under the section "Invoicing", billing information must be entered. All submitted applications are invoiced a fee due to initiating the application process. If the application is granted data, a cost will be added for this. Current prices are specified at www.scapis.org. We ask you to check

if your organization wants electronic invoice. Please note that it is important that you enter an invoice reference and that you enter the reference in the format required by your organization.

2.4 Step 2 “Describe the research project”

Note the language requirements in section 2.2.

At the end of this page, if the applicant project is linked to a local sub-study within SCAPIS, choose “YES” and then select which local sub-study in the drop-down list. If the sub-study you are looking for is not in the list, contact the SCAPIS office at scapis@scapis.org. Beneath the drop-down list, please briefly describe the connection between your project and the local sub study.

2.5 Step 3 “Approval from ethical review”

To process your application, the ethics application and its approval must be attached to the application. Observe that these requirements for the ethics application need to be fulfilled for your application to be able to be approved:

- the approval must clearly refer to the project applying for data
- the population described in the ethics application must match the population in the application
- the project’s Principal Investigator must be involved in the ethics application
- the organization responsible for the project (“forskningshuvudman”) applying for data from SCAPIS must be specified in the ethical application

Please note that **to access register data** (selected in step 6 b), **specific and mandatory requirements apply regarding the ethics application**. All information you need to write in the ethical application is available at www.scapis.org. Also note that in addition to specific wordings in the ethics application, it is **also required that specified information has been attached to the ethical application regarding which register data that is intended to be retrieved from SCAPIS**. Specified information about register data to attach your ethical application is provided by the SCAPIS office, contact scapis@scapis.org if you need this information.

2.6 Step 4 “Population”

Choose whether your population selection applies to the core study and / or the pilot study. If you choose the core study, you must specify whether it applies to the entire core cohort or certain sites. Since SCAPIS requires that your project offers collaboration for at least one researcher from each site from whom variables are selected from, you will be asked to fill in the name for one researcher from each site you ask for variables from.

All data distributed from SCAPIS is pseudonymized where each SCAPIS study participant is represented by an ID that is unique for each specific dataset delivered from SCAPIS. If you for any reason need a specific type of ID in your dataset, you must clarify this under “Data ID”.

2.7 Step 5 “View SCAPIS core projects”

As an applicant, you are responsible for carefully reviewing the titles for core publications within SCAPIS. If the project scope covers a topic that interferes with an ongoing SCAPIS core publication, the application will not be approved. If the core publication title does not provide sufficient information to evaluate this overlap, there is a link to a PDF where all Topic Research Forms (TRF:s) for all ongoing SCAPIS core publications can be found. When you have reviewed the core publications, you state either "I believe, to the best of my abilities, that my research question does not interfere" or "I am not sure if my research question interferes with, or is the same as, one or" and then mark in the list which title or titles you are unsure about.

2.8 Step 6a “Select core variables”

In this step, you select the variables that you are interested in from the SCAPIS core study and/or local sub studies associated with SCAPIS core study. Remember that requested data must be relevant to the project based on the description in the ethics application.

To view data from local sub studies, you need to select from which sites you want see local sub studies that currently have variables to choose among. If you do not select a site here only core variables will be shown. Note that you only can ask for sub study data that corresponds to the population selected in step 4. Regarding local sub study data, see also section 2.8.2.

Note that selection of genotype variables requires that your approved ethical review application states that you will use genetic data in your project.

2.8.1 Temporary restricted access to some data variables

Every now and then new data is added to the SCAPIS database and it may be relevant that certain data is restricted for a limited period. If there currently is data with temporary restricted access, you will find it in own subdomains named "restricted access". For updated information, see www.scapis.org.

2.8.2 Approval required for variables from local sub studies

If you choose variables from one or more local sub studies, you will need an approval from the responsible researcher. This will be handled by an email that will be sent to the respective responsible researcher when the application is submitted. The researcher responsible for the sub study is given information about which variables are requested, project title, project summary, Principal Investigator and the PI's email. The researcher responsible for the sub study variables will confirm to the SCAPIS office that data sharing is accepted, or the researcher will contact your project's Principal Investigator directly by email to make an agreement about certain conditions. The agreement is then reported to the SCAPIS office. If the local sub study does not accept that variables are to be shared for your project at all, the SCAPIS office will contact you when processing the application.

2.9 Step 6b “Select registers”

If register data is not relevant to your application, just click “I do not need any register data for this application” and proceed to the next step.

If your application requires register data, select one or both of the following:

- “I would like access to raw register files provided directly by SCAPIS” and select the registers you wish to apply for.
- “I would like to initiate a new register linkage” and specify your linkage request.

Please note that in order **to grant register data from SCAPIS, it must be clearly stated with specific wordings in your ethical application/approvement** that you will use data from the register in question. Also note that applying for register data from SCAPIS is associated with a separate additional cost per register. See also section 2.5 in this instruction and information in detail at www.scapis.org.

2.10 Step 6c “Select pilot variables” (optional)

If you have selected data from SCAPIS’ pilot study or data from both core and pilot in step 4, variables from SCAPIS’ pilot will appear in a specific step, step 6c.

2.10.1 Approval required for variables from the pilot study

If you choose data from the pilot study the same procedure will take place as for the local sub study data, see section 2.8.2. Responsible for pilot data is SCAPIS Gothenborg’s site Principal Investigator.

2.11 Step 6D “Select genotype data”

If genotype data is not relevant for your application, click “I do not need any genotype data for this application” and proceed to step 7.

If you are applying for genotype data, please read the section “If you apply for genetic data, please note:”, and then click “I have read and understood the above information”.

You will now see the following two main alternatives:

- “I would like access to imputed genotype data”
- “I would like access to directly genotyped data”

See detailed descriptions of each main alternative in the following sections

I have read and understood the above information

I would like access to imputed genotype data
Imputed data based on the HRC 1.1 panel, suitable for analysis of individuals of European ancestry. Imputed genotype data is available for 40 million SNPs and approximately 29,400 subjects. File size for the full population is 113 GB.

I would like access to directly genotyped data
More than 700,000 directly genotyped SNPs are available for approximately 29,400 subjects. File size for the full population is 2 GB.

It is important to note that

- it must be stated in your ethical review application that you intend to include analysis of genetic data in your research.
- if you apply for genotype data, the variables under “Genotyping – assay information” and “Genotyping – supporting results” in section 6a will be automatically added to your application.

2.11.1 Access to imputed genotype data

Under this alternative you can choose between

- Complete dataset
- SNPs based on names
- SNPs based on position

Selection of complete dataset

If you select “Complete dataset of imputed genotype data” you cannot select genotype data based on names or position. However, it is possible to make selections also for the “directly genotyped data”.

I would like access to imputed genotype data
Imputed data based on the HRC 1.1 panel, suitable for analysis of individuals of European ancestry. Imputed genotype data is available for 40 million SNPs and approximately 29,400 subjects. File size for the full population is 113 GB.

Select the alternatives for the imputed genotype data
Note that selection of both SNPs based on names and SNPs based on genomic position will generate two datasets.

Complete dataset of imputed genotype data (File size for the full population is 113 GB)

Select SNPs based on names in the imputed genotype data

Select SNPs based on genomic position in the imputed genotype data

Selection of SNPs based on names

Select this option in order to apply for specific SNPs. The names are either rsIDs, e.g. rs123456789, or custom names in the format, chromosome:position:REF_ALT (where REF and ALT refers to reference allele and alternative allele) e.g. 1:842461:G_T. Custom names are only available for SNPs that lack rsID.

List the names of the SNPs in the text-box. Available SNPs and unavailable SNPs are listed below the box. The names need to be exact. Wildcards will not work.

Complete dataset of imputed genotype data (File size for the full population is 113 GB)

Select SNPs based on names in the imputed genotype data

List the SNPs with rsIDs or custom IDs (for SNPs that lack rsID), separated with commas.
The genomic position corresponds to Genome Reference Consortium Human Build 37 (GRCh37).

rs571093408, rs12345, rs753763699

Available SNPs (2)

| Id | Chromosome | Position | Ref | Alt | Info score |
|-------------|------------|----------|-----|-----|-------------|
| rs571093408 | 1 | 13380 | C | G | 1.24992e-09 |
| rs753763699 | X | 40840080 | G | A | 0.247446 |

Not available SNPs

Id

rs12345

Selection of SNPs based on genomic position

You can apply for SNPs within a range with a given start and end position. Click “Add selection”, select chromosome, and specify start and end positions. Click “Add selection” again to add more ranges. You can remove a range by clicking “Remove”. Click “Download SNPs (csv)” if you would like to download a csv-file with a list of all the SNPs within that range.

Select SNPs based on genomic position in the imputed genotype data

Click “Add selection” and specify chromosome number, start and end positions.

The genomic position corresponds to Genome Reference Consortium Human Build 37 (GRCh37).

If you would like to know which SNPs that are included in each selection, click the button to download a csv file.

| Chromosome number | Start position | End position | No. of SNPs | Remove | |
|--------------------------------|---------------------------------------|---------------------------------------|-------------|---------------------------------------|--|
| <input type="text" value="x"/> | <input type="text" value="40839981"/> | <input type="text" value="40840131"/> | 3 | <input type="button" value="Remove"/> | <input type="button" value="Download SNPs (csv)"/> |

2.11.2 Access to directly genotyped data

The different alternatives for the directly genotyped data are identical to the alternatives for the imputed genotype data described in section 2.11.1. There is however a difference regarding the naming of the SNPs for the directly genotype data. Some will have conventional rsIDs, but several SNPs will have different type of names, or rsIDs with a prefix. Since names need to be exact when selecting SNPs based on names, we recommend that you download and refer to either “Illumina’s chip manifest for GSA MD 3”, or “list of custom and base panel SNPs” from the list of further documentation.

- **Further documentation:**
 - [Download example data](#)
 - [Link to the HRC 1.1 panel](#)
 - [Download Illumina’s chip manifest for GSA MD 3](#)
 - [Download list of custom and base panel SNPs](#)
 - [Download Quality Control report SCAPIS genotype data](#)

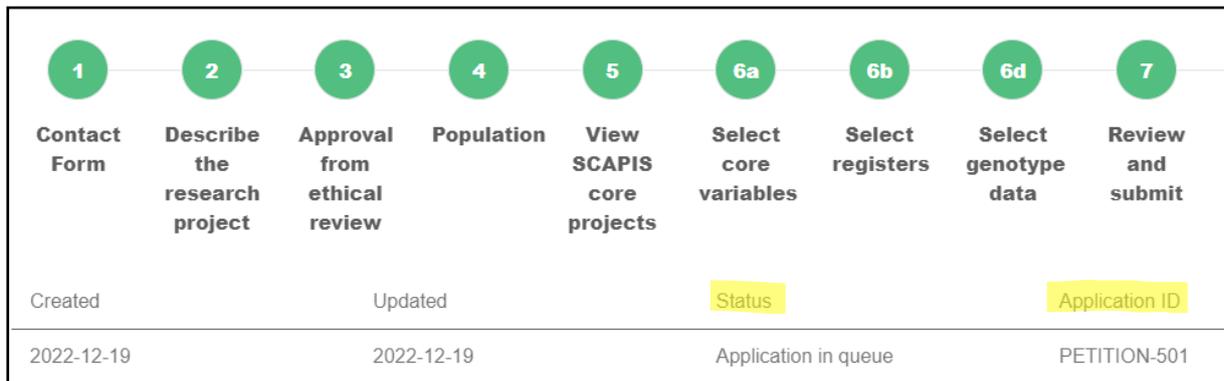
I have read and understood the above information

2.12 Step 7 “Review and submit”

In this step, you can review your application before submitting it by clicking on “Submit”. After an application has been submitted, information in the application is locked for editing. Note that if you have chosen variables from a local sub study and / or from the SCAPIS pilot, you will see under “Embargo emails” which emails will go to the relevant responsible researchers for approval.

3 See and follow your application in the Portal

When you are logged in to the SCAPIS Portal, you can view and click on both ongoing applications and applications that have already been submitted by clicking on "APPLICATIONS" at the top of the page. For submitted applications, you can see under "Status" where your application is in the processing process. There are three different statuses:



- **"Application in queue"**: The application is pending and has not yet started the processing process.
- **"Application under review"**: The application is actively being processed by the SCAPIS office.
- **"Application review completed"**: The application has been processed and a decision has been made.

In the overview, you can also see an **"Application ID"** for submitted applications. This ID is your application serial number at the SCAPIS Office and is useful in any communication with the SCAPIS office.