SIB Open Research Licensing Guideline

Preamble

The SIB Swiss Institute of Bioinformatics (**SIB**) promotes Open Science as a core principle for software and research data. In particular, SIB supports the sharing of research data, software and their meta data by implementing the FAIR principles according to which they should be findable, accessible, interoperable and reusable. Furthermore, SIB endorses free access to scientific publication (Open Access) and encourages the use of open public licences wherever possible.

The purpose of this SIB Open Research Licensing Guideline (the **Licensing Guideline**) is to provide recommendations and guidance related to the distribution of SIB's research data, datasets, databases, publications and software generated or developed by SIB (the **SIB Materials**).

Despite the strong encouragement to use this Licensing Guideline, this document is non-binding and its application must be evaluated on a case-by-case basis, in particular by taking into account possible legal (e.g. personal data regulations) or contractual barriers (e.g. confidentiality obligations, IP rights). In this respect, the separate document "**SIB Open Science Checklist**" provides key considerations for determining whether you can apply the licensing solutions provided in this Licensing Guideline.

Training materials: please note that the Copyright Guidelines have been published to provide basic knowledge regarding copyright law and specific guidance on how to be acknowledged and re-use third party content for training materials. If you would like to know more about this matter, please refer to this document.

This Licensing Guideline was defined with a pragmatic approach, focusing on minimum requirements. For the sake of conciseness, some legal and technical aspects have been voluntarily omitted. If you want to know more about open licensing or FAIR principles, you will find further reading below:

- Elixir FAIR Cookbook
- FAIR principles (for research data)
- FAIR4RS principles (for research software)
- OpenAIRE Guidelines
- IMI2 Project guidelines for open access to publication and research data
- FAQ Creative Commons licences
- Free Software Foundation Licensing & Compliance Guide

I.   Licensing recommendations for SIB Materials

Important: this section I ("Licensing recommendations for SIB Materials") establish concrete recommendations of open public licences to apply to SIB Materials. However, depending on the circumstances, adoption of the recommended licences may be prevented by various obstacles, such as applicable laws, pre-existing internal policies and/or provisions of an agreement with a third party. If you have any doubts or concerns about the adoption of the recommended licences on SIB Materials, please contact the SIB Legal and Technology Transfer Office (legal@sib.swiss).

A.   THIRD PARTY MATERIALS

As a general rule, if you re-use third party materials and/or make any derivatives from them, you must make sure you are allowed to do so.

If the third party materials are licensed under **copyleft licences** such as the Creative Commons Attribution-ShareAlike or GNU GPL licence, it is strongly recommended to maintain an internal list tracking these materials and their respective licences. These licences grant permissions to use, copy, modify and distribute work, provided that any modifications or derivative works are also made available under the same terms. Unless otherwise agreed with the rights holder(s), this means that if you include or combine them with SIB Materials, you may have to apply the same terms and conditions for the distribution of your own works.

In the event you wish to reuse materials and/or software code generated by **artificial intelligence tools (AI Tools)**, it is important to carefully review their terms and conditions. Some AI Tools may grant you a free worldwide unlimited licence to use the outputs, while others may assert ownership or impose restrictions on their use. They may also state that you can freely re-use the outputs without any guarantees, particularly from an intellectual property perspective. As these outputs may contain or rely on copyrighted materials or other content owned by third parties, it is important to ensure that you have the permission to re-use them and under what conditions.

B.   DATA, DATASETS AND DATABASES

CC BY 4.0. SIB recommends to licence databases and their processed data in the form of texts (transcript, text), graphics (table, chart, diagrams, animations, models) and images (whether fixed or moving such as pictures, photos, visuals) included in the database under the **Creative Commons Attribution 4.0 International Public Licence** (CC BY 4.0).

Databases are composed of different elements such as the data itself, metadata, structure and data model of the database, visualisations and works derived from the data. You should avoid applying different licences to these different parts of databases. A single public licence should be preferred, i.e. the CC BY 4.0 licence.

Creative Commons Attribution, Share Alike 4.0 International Public Licence (CC BY-SA 4.0) may be considered in specific cases. Distributing derived datasets and databases originally licensed under CC BY-SA terms implies to release them under the same conditions as the original. A downside of the share-alike provision is that it can hinder open data adoption as it can result in data being less open than intended. Since CC the BY-SA 4.0 licence may reduce interoperability and automation, it should be used primarily for ensuring that all derivatives of the original content will be licensed under the CC BY-SA 4.0 licence.

CC0 Public Domain Dedication. For raw data (data collected from a primary source and that has not been processed in any way), and unless required by law or contractual obligations, SIB recommends applying the **Creative Commons Zero 1.0 Universal Public Domain Publication Licence** (CC0 1.0) to ensure and promote their accessibility and dissemination.

Please note that when datasets and databases combine both raw data and processed data, different licensing solutions may come into play (i.e., CC0 1.0, CC BY 4.0 and CC BY-SA 4.0 licences). In this case, SIB recommends applying the **same licence** to the database and its content, namely the CC BY 4.0 licence. This provides clear and concise terms for the access and reuse of all data made available in the database, ensures that proper acknowledgement and citation are given to the rights holder(s) and ultimately, helps to maximize the widest possible dissemination and reuse of the data.

Public Repositories. If you publish your research data in a public data repository, make sure to include the licence under which the research data is available and its Digital Object Identifier (DOI). If you feel lost in the heterogenous landscape of repositories, the website www.re3data.org provides an overview of existing international repositories for research data.

C.   OUTPUTS PROVIDED BY WEB-BASED TOOLS

Principle. SIB Materials can be incorporated and used by web-based tools to generate Outputs (the **Outputs**). For example, SIB Materials (e.g., in the form of annotated data) may be integrated into web-based tools to generate spider graphs to determine physicochemical and medicinal properties of molecules or 3D models of a protein. Regardless of their representation (diagrams, tables, figures, modelling), the ownership over such Outputs should be clarified so that end-users can easily understand on what terms they may re-use them. The decision to licence Outputs provided via the use of web-based tools must be assessed on a case-by-case basis.

Restrictions related to third party rights. Where Outputs are built upon by querying and collecting data from third party websites or resources, the licensing of Outputs may require the explicit consent of those third parties and may require specific attribution.

CC BY 4.0. If no restrictions apply and no third party can claim rights over the Outputs, SIB recommends applying a CC BY 4.0 licence. The CC **BY-SA** 4.0 licence should only be considered to ensure that Outputs and their derivatives can be reused by any third party under the same rights.

D. SOFTWARE

Principle. SIB supports the release of software under Open source licences. However, the decision to release software under an Open source licence must be evaluated on a case-by-case basis.

Absolute restrictions. The Open source licensing of software **is not allowed** if such licensing would infringe applicable laws, pre-existing internal policies and/or provisions of an agreement with a third party.

Relative restrictions. In certain situations, a case-by-case evaluation may lead to the conclusion that licensing software under Open source licences **should be avoided**. This may be the case in the following situations:

(i) a substantial effort is required to package the software appropriately for public release, and it is considered that this effort significantly outweighs the expected benefits of open source distribution; or

(ii) the quality of the software (in terms of performance, reliability, functionality and programming) is considered insufficient, and its public exposure risks to convey a negative image; or

(iii) there is an overriding interest in retaining control of the source code of a software and not releasing it, it being understood that SIB encourages the dissemination of Outputs generated by software under free open licences (e.g., results generated by software and web-based tools); or

(iv) based on a thorough analysis of the case, a non-Open source approach is considered more effective for maximizing dissemination.

GPL-2.0-or-later. In the absence of restrictions, SIB recommends licensing software under the **GNU General Public Licence v2.0 or any later version** (GPL-2.0-or-later) which are copyleft licences. For more information about GNU GPL licences, please refer to GNU website.

The following licence models may also be considered for software:

(i) Apache-2.0: this licence may be used when the copyleft effect of the GNU GPL licences are too restrictive or when restrictions are imposed by existing agreements, or when no control over the possible commercial exploitation of the program by third parties is necessary.

(ii) LGPL-3.0-or-later: this licence may be considered for software libraries when the primary objective of the open-source distribution is the rapid and widespread adoption of those software libraries.

(iii) AGPL-3.0-or-later: this licence should be applied for Software-as-a-Service (SaaS) model to reinforce the copyleft effect. This ensures that any source code snippet remains available and prevents third parties from repacking and selling it through a SaaS model.

E.   SCIENTIFIC PUBLICATIONS

CC BY 4.0. To the possible extent, SIB employees **are strongly encouraged to retain their rights** in scientific works. This can be done through an addendum to the publishing contract (see for instance, the template addendum available underline here).

To promote Open Access, publications and other scientific outputs (e.g. posters, presentations) should be licensed under an open licence, preferably, the CC BY 4.0 licence, and published in open access journals or by an open access publisher (listed in DOAJ, DOAB, COPE, OASPA). Disseminating the full text of the publications in an appropriate public non-commercial repository (listed in OpenDOAR) should also be considered.

Where applicable, ensure that the publications state the funding source and other formal acknowledgement according to the applicable regulations, in particular applicable funding regulations.

F.   IMPLEMENTATION CONSIDERATIONS

Minimum requirements. SIB Materials should include the name of the rights holder(s), the date of first release/publication and the mention of the applicable licence.

If your research data is packaged in a ZIP, RAR or TAR file, a legal statement should be included in an introductory document at the top level of the directory structure. The same approach should be followed for code snippets.

The legal statement may also include the names of third parties who contributed to the development of the research data/software code and/or from whom you received funding. Please note that this could also be a contractual requirement arising from an existing agreement.

SIB templates. To ensure proper implementation of these SIB licensing recommendations, SIB encourages the use of templates attached to this Licensing Guideline. They also include additional considerations such as privacy, liability, governing law and jurisdiction provisions. Specific comments have been included in italics in red so that they can be easily adapted to specific cases.

II. GENERAL OVERVIEW

Data, Datasets and Databases	Outputs provided by web-based tools	Software	Scientific publications
Default licence **CC BY 4.0** *(Section I.B)*	*Default licence* **CC BY 4.0** *(Section I.C)*	*Default licence* **GPL-2.0-or-later** *(Section I.D)*	*Default licence* **CC BY 4.0** *(Section I.E)*
Alternate licence **CC BY-SA 4.0** *(Section I.B)* *Exception for raw data* **CC0 1.0** *(See Section I.B)*	*Alternate licence* **CC BY-SA 4.0** *(Section I.C)* *Restrictions* *(See Section I.C)*	*Alternate licence* **Apache-2.0** **LGPL 3.0-or-later** **AGPL-3.0-or-later** *(Section I.D)* *Restrictions* *(See Section I.D – Absolute and relative restrictions)*	**X**
Minimum Requirements A) Copyright or attribution statement and licence notice B) Identify third party materials and their respective licence	*Minimum Requirements* A) Include terms of use in web-based tools used to generate Outputs (including copyright and licence) B) Identify third party materials and their respective licence	*Minimum Requirements* A) Copyright and licence notice B) Publish the software code in the dedicated SIB repository (GitHub or GitLab)	*Minimum Requirements* According to publisher's rules. Usually: A) Copyright and licence notice B) Include citing rules
Legal Notice Template *Annex A*	*Legal Notice Template* *Annex B*	*Legal Notice Template* *Annex C*	**X**

Annex A – Legal notice template for data, datasets and databases

The following template notice provides a non-exhaustive list of key points to include that will govern the access to and use of databases and their contents. For databases accessible online, it may be useful to implement a click-through notice (with an "I agree" button) to strengthen the enforceability of the applicable terms. On balance, this may also limit the ability of automated tools to access the data available on the online database (API). If you wish to allow queries to your site via APIs without implementing a click-through notice, we recommend that you include a specific statement (see Section 1 below).

---

1. Copyright. © < *year of public release* > < *right holder(s)* > *[only for copyright materials licensed under CC BY 4.0 or CC BY-SA 4.0]*

2. Licensing. Unless otherwise noted, any content or data available on [XXX] is licenced under a *[applicable licence]*

[Optional – only if data is incorporated into database accessible online]
In addition, we request that any content or data taken from [XXX], either by the API, scraping the site or other means, is attributed and cited appropriately.

Please note that the use of the name of [XXX] [*optional*: and its logo] must not be associated with any publicity or business promotion, without our prior written approval.

[Optional - to exclude third party materials from the CC licences]
3. Exclusion. The *[applicable licence]* licence terms do not apply to the following elements included in the [XXX]: *[list all the materials included in [XXX] that are licensed under specific licensing terms and the name of the rights holder(s)].*

4. Acknowledgment and Citing Rule.

[*For content licensed under CC BY // or CC BY-SA licence*]
- Please cite the following references papers when reusing or building upon our [XXX], or when you publish or distribute results that were fully or partially based on research materials obtained from the [XXX]: [XXX]

[For content licensed under CC0 1.0 licence]
- If you wish to cite [XXX] or when you publish or distribute results that were fully or partially based on the data obtained from the [XXX], please cite the following references papers: [XXX]

[Optional addition depending on the nature of research data]
5. Medical Use. [XXX] is not designed, intended, or made available as a medical device, or a substitute for professional medical advice, diagnosis, treatment, or judgment.

6. As is. [XXX] and its content are provided "as is" and "as available". We do not guarantee their availability, correctness, accuracy, reliability, and completeness, nor the suitability for any specific purpose. We assume no liability for any indirect, or consequential loss or damages.

[Optional – only if accessible online]
7. Personal Data. See [XXX] Privacy Policy. *[Reference must be made to the applicable privacy policy of the institution hosting the [XXX] and which is responsible for the processing of the personal data of users].*

8. Governing Law and Jurisdiction. The access and use of the [XXX] are governed by Swiss substantive law, without reference to its conflict of laws provisions. The competent courts in *[city name]*, Switzerland have exclusive jurisdiction.

9. Contact. If you have any questions, feel free to contact [XXX] at [XXX@XXX.XXX]

[Optional – only if accessible online]
10. Modifications. We reserve the right to amend and update these terms of use at any time.

Annex B – Legal notice template for web-based tools

Where web-based tools are used for providing computational services to third parties, clear terms and conditions should be established so that end-users can easily understand the rules governing their access and use as well as those governing the right to use the Outputs. To this end, a non-exhaustive list of key points is provided below. As adjustments may be required depending on how the web-based tools operates, we have included specific comments in italics in red to help you adapt this template to your needs.

Note that the legal notice template below **should not be applied** to regulate the use, modification, and distribution of the source code of web-based tools (refer to Annex C). This template should only be considered to describe user obligations and restrictions when using web-based tools to generate Outputs.

[XXX] is committed to open science and open knowledge. Through its [*ToolDescription*], [XXX] [*DoesWhat*] and is operated by *[right holder(s)]*.

Your access to, and use of, the [XXX] and any content or data available on it, or generated through it (together the **Content**) are subject to the following terms and conditions:

1- Licensing

SIB and our licensors own the copyright and all intellectual property rights in and to [XXX] and the Content. You have the right to use them, provided you comply with the following:

a) [XXX]. Your use of the [XXX] is subject to the acceptable use policy below (see Section 2 below).

b) User Output. Unless we mention otherwise, if you download a file or generate any output and download it (for instance by making an individual query) as part of your acceptable use of the [XXX](together the **User Output**), such User Output will be licensed to you under the [Creative Commons Attribution licence (CC BY 4.0) // or // the Creative Commons Attribution-ShareAlike licence (CC BY-SA 4.0)]. This licence gives you the rights to share and adapt the User Output under a few core conditions, and in particular:

- Attribution: You must give appropriate to [XXX] when reusing or sharing the User Output. Do so in accordance with best practice in the scientific community, clearly indicating that you used the [XXX]. You may do so in any reasonable manner, but not in any way that suggests [XXX] endorses you (see Section 1(c) below).

When you publish, patent or distribute results that were fully or partially based on User Outputs or models obtained from the [XXX], please cite the corresponding papers.

[*Only if CC-BY SA 4.0 licence applicable*]
- ShareAlike: if you remix, transform, or build upon the User Output, you must distribute your contributions under the same licence as the original.

Visit the Creative Commons website for information about your rights and obligations under this licence.

c) Branding. We do not authorize the use of our name, logos and other branding of [XXX]. [XXX] must not be associated with any publicity or business promotion, without our prior written approval.

[Optional - only if your software code has been publicly released under AGPL licence]
d) Open source. Certain components of the [XXX] may be covered by so-called open source/free cultural licences. Nothing in

these terms of use restricts, limits or otherwise affects any rights or obligations that you may have, or conditions to which you may be subject, under any applicable open source (free cultural licences).

2. Acceptable Use Policy

You must only use the [XXX] for lawful and legitimate reasons. The following is a non-exhaustive list of practices that would not be considered a legitimate use:

a) accessing or using any part of the [XXX] for the purpose of building or improving a competitive product or service or copying its features or user interface, including as an example, by enabling others to access or use the [XXX] through a separate portal.

b) activities that may impact the stable operation of the [XXX] or impair anyone else's use of it, such as systematic or automated queries and downloads, automated submissions, or otherwise using other mechanisms than the ones provided by us.

c) using any automated means, such as bots, robots, spiders, crawlers, or scrapers, or otherwise circumventing any restrictions in any robot exclusion header, for any purpose, including for the purpose of automatically collecting and reconstructing a Substantial Part of the [XXX]. For the purposes of these terms, a "Substantial Part of the [XXX] means any subset of the [XXX] or its content containing more than 20% of the [XXX].

d) circumventing technological measures intended to prevent direct access to the databases and/or software code underlying the [XXX].

3. Acknowledgement and Citing Rule

When you publish, patent or distribute results that were fully or partially based on

Content obtained from the [XXX] please cite the following references papers: [name]

[*Optional*]

4. Medical Use

[XXX] is not designed, intended, or made available as a medical device, or a substitute for professional medical advice, diagnosis, treatment, or judgment.

5. As Is

[XXX] and Content (including all User Output) are provided "as is" and "as available". We do not guarantee their availability, correctness, accuracy, reliability and completeness, nor the suitability for any specific purpose. We assume no liability for any indirect, consequential loss or damages. You use them at your own risk.

6. Personal Data

See the SIB's Privacy Policy *[Reference must be made to the applicable privacy policy of the institution hosting the [XXX] and which is responsible for the processing of the personal data of users].*

7. Contact

If you have any questions, you can contact the [XXX] at [XXX@XXX.XX].

8. Governing Law and Jurisdiction

These terms of use and the use of the [XXX] are governed by Swiss substantive law, without reference to its conflicts of laws provisions. The competent courts in Lausanne, Switzerland, have exclusive jurisdiction.

9. Modifications

SIB reserves the right to amend and update these Terms of Use at any time.

Annex C – Legal notice template for Software

Appropriate legal notice must be embedded in the software code when it is ready for distribution on public platforms like GitHub or GitLab. Such notice is typically placed in a file named "LICENSE" or "README" in the root of the repository that contains the full copyright and license statement of the distributed software code. Here is an example that you may include in your "LICENCE" or "README" file:

---

1. Copyright. © < *year of public release* > < *right holder(s)* >

2. Licensing. [*XXX – software name*] is licensed under *[name of the applicable software licence]*.

[*Optional*]
3. Medical Use. [XXX] is not designed, intended, or made available as a medical device, or a substitute for professional medical advice, diagnosis, treatment or judgment.

4. Acknowledgement and Citing Rules. If you like to cite the Software or when you publish or distribute results that were fully or partially based on the research data obtained from the [XXX], please cite the following references papers: [XXX]

5. Contact. If you have any questions about the use of [XXX], please contact us at [XXXX@XXXX.XX].

[*Optional – only for AGPL software*]
6. Terms of Use. The access and use of [XXX] governed by the AGPL license is subject to the terms and conditions set forth in a separate document titled "Terms of Use" located at [*URL or file path*]. The Terms of Use are separate from and in addition to this AGPL license. By accessing and using [XXX], you agree to comply with the Terms of Use in addition to the terms of this AGPL License.

[You can copy/paste the plaintext version of the applicable software licence].

Snippet. You should also place a copyright notice in every portion of your source code package so that it can remain easily identifiable even if it is separated from the original package. You can insert the following statement at the top of your files:

---

* This [*source code / file*] is subject to the terms and conditions defined in file "Licence" [or] "Readme", which is part of this source code package. See the ["Licence.txt" [or] "Readme"] file distributed with this work for additional information regarding copyright ownership. You may not use this file except in compliance with the licence. A copy of the licence is available at [https://XXXX]

Acknowledgment. You may also indicate how to cite your software tool in research papers. If needed, you can refer to the following article which provides guidance on how to cite software tools: https://f1000research.com/articles/9-1257. In some situation, it might also be interesting to publish your software tools on public repositories so they can have a Digital Object Identifier and include it in your software files.

Third party software. If third party software has been incorporated into your own software, you should also include a notice listing all third party materials re-used and their respective licence.

Terms of Use. In case your software is used to provide computational services and generates Outputs, you should include terms of use in the root directory of your repository or in a designated folder that address specific issues that are not covered by the AGPL licence. You can create a folder called "docs" to store your "Terms_of_Use" file. This ensures that users can easily find and review the terms before using your software. You can either copy and paste the template provided in Annex C into your terms of use root file, or provide a link to access it.