

SSIE – Training days, March 24-25, 2023

Introduction to software licensing

Joël Dumoulin
Idiap Research Institute

Table of content

1. Software licensing
2. Open source software
3. License your software

Table of content

1. **Software licensing**
2. Open source software
3. License your software

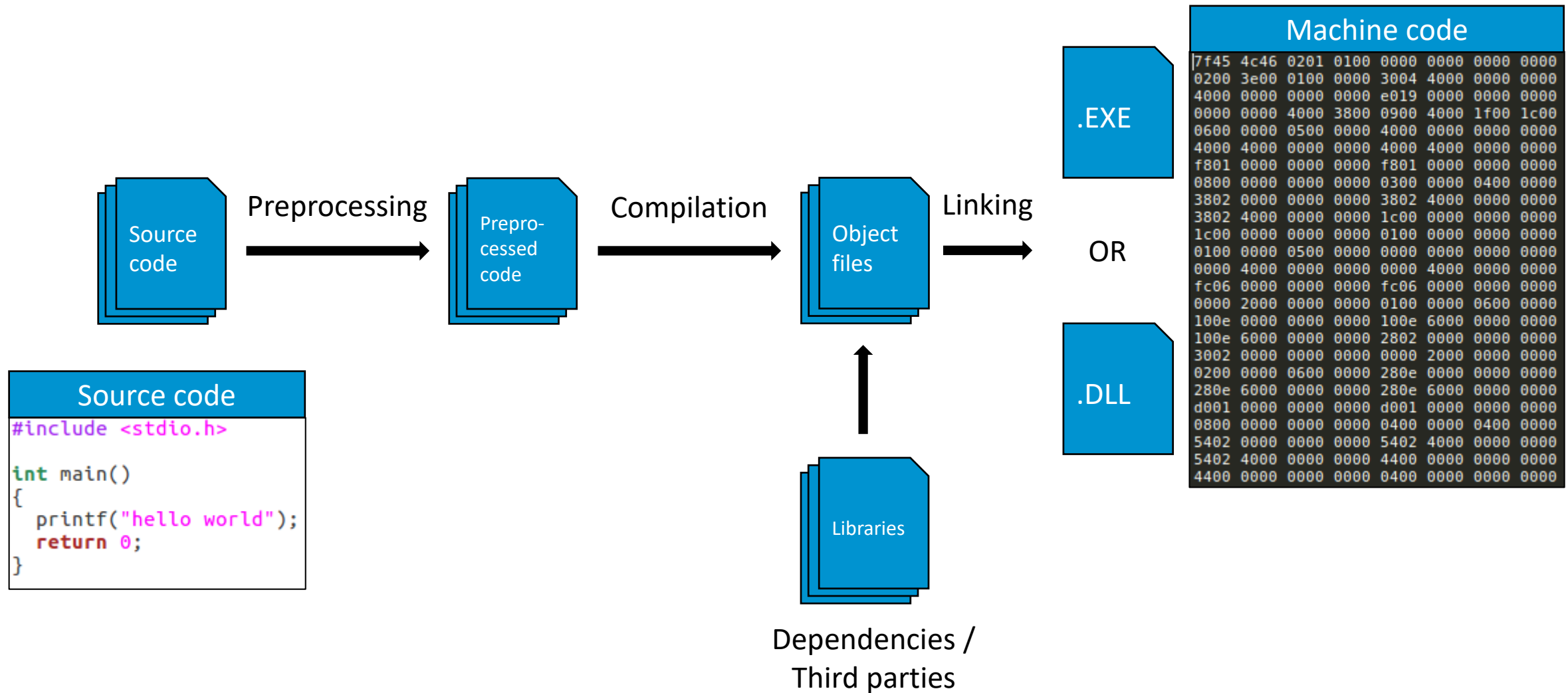
1. Software licensing

- Legal instrument governing the use and redistribution of a copyrighted software
- Software licenses examples: BSD, MIT, GPL
- Usually it is a simple text file (LICENSE.txt)
- **WARNING**
 - No explicit licence? → "All rights reserved" = No rights for you
 - Ex: code published on Github without a license file → you cannot even run it!

Copyright

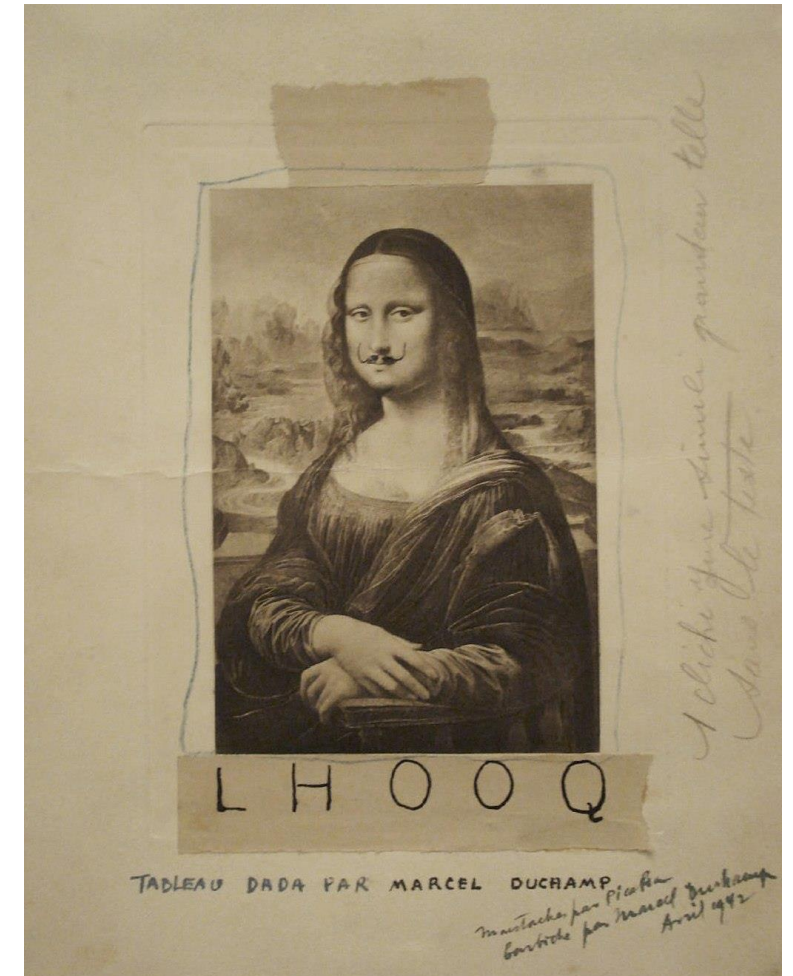
- Type of intellectual property
- Gives its owner the exclusive right to make copies of a creative work
- Allows the author to prohibit others from reproducing / adapting / distributing copies of his work
- Fair use
 - EU copyright laws allows some exceptions
 - Reproduction made by libraries, educational establishments, museums, which are non-commercial
 - For non-commercial research or private study

Source code



Derivative work

- Expressive creation that includes major copyrightable elements of an original work
- The transformation, modification or adaptation of the work must be substantial to be original
- Examples
 - Translations
 - Cinematic adaptations
 - Musical arrangements



Types of software licenses

Software licenses can be categorized in 5 types

- Public domain
- Permissive
- Protective
- Proprietary
- Non-commercial licenses

Types of software licenses (1/5)

Public domain

- The most permissive type of software license
- There is no copyright, as the author donates his/her work to the public domain
- Anyone can use, modify, distribute or even sale the software without any restrictions and without even mentioning the author

License examples	Software examples
<ul style="list-style-type: none">• Creative Commons CC0• Unlicense• BSD0	SQLite, ImageJ, BLAST, SHA-3

Types of software licenses (2/5)

Permissive

- The most popular type of license
- They contain minimal requirements about how the software can be modified or redistributed

License examples	Software examples
<ul style="list-style-type: none">• Apache 2.0• MIT• BSD-2, BSD-3	.NET Core, Django, React

Types of software licenses (3/5)

Protective

- Also known as reciprocal licenses or restrictive licenses
- These licenses allow you to modify the licensed code and distribute new works based on it, as long as you distribute any new works or adaptations under the same software license.

License examples	Software examples
<ul style="list-style-type: none">• GNU family of public licenses (GPL, LGPL, AGPL)• Mozilla Public License• Eclipse Public License	GIMP, MySQL, Notepad ++, Webkit engine, PyQt

Types of software licenses (4/5)

Proprietary

- The most restrictive of all
- All rights are reserved
- Generally used for proprietary software (closed source software) where the work may not be modified or redistributed

Software examples

Microsoft Windows, Skype, macOS, Adobe Photoshop

Types of software licenses (5/5)

Non-commercial licenses

- Freeware / Freemium / Shareware
 - The free here refers to the price
 - The source code is not made available

Types of software licenses (😊/5)

“Funny” licenses

- THE BEER-WARE LICENSE (Revision 42)
 - It requires that the notice be retained and suggests that if users like the software, they buy the author a beer.
- CDL (Chicken Dance License)
 - GPL-like, but as an alternative to distributing source code with binaries, it gives the distributor the option to post a video of employees doing the Chicken Dance to social media sites.
- WTFPL (Do What The F***K You Want Public License)
 - No comment!
- And many more!





Nobody uses these licenses? Have a look to the repository of the YOLO project ;)

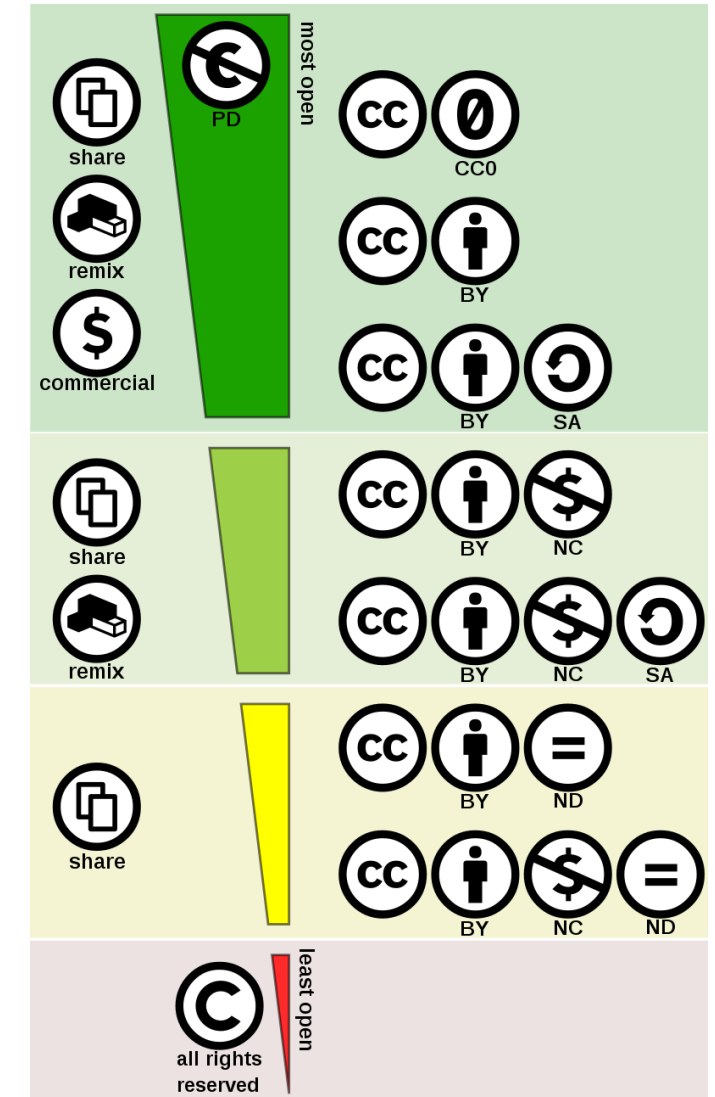
- <https://github.com/pjreddie/darknet>

EULA

- End User License Agreement
 - Used for closed source software: the user receive a software, not its code
 - Legal contract between a software developer or vendor and the user of the software.
 - Specifies in detail the rights and restrictions which apply to the use of the software.

- All work falling under copyright
 - Including books, plays, movies, music, articles, photographs, blogs, and websites.
 - **Data and databases**
 - But not recommended for software licensing
- Combinations of 4 conditions

Icon	Right
	Attribution (BY)
	Share-alike (SA)
	Non-commercial (NC)
	No Derivative Works (ND)



License: CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>)

Documentation

- The code documentation is usually covered by the same licensing as the code
- Other documentation such as user guides, tutorials, and reference manuals can have different licensing needs
- Examples
 - Open Publication License - <http://opencontent.org/openpub/>
 - GNU Free Documentation License (FDL) - <https://www.gnu.org/licenses/fdl-1.3.html>
 - Public Documentation License (PDL) - <https://www.openoffice.org/licenses/PDL.html>
 - FreeBSD Documentation License - <https://www.freebsd.org/copyright/freebsd-doc-license.html>
 - Creative Commons - <https://creativecommons.org/>

Table of content

1. **Software licensing**
2. Open source software
3. License your software

Table of content

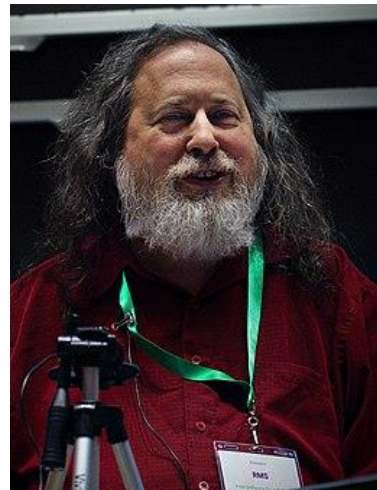
1. Software licensing
- 2. Open source software**
3. License your software

2. Open source software (OSS)

- Open source vs Closed source
- Open Source Software (!)= Free Software
- Free Software
 - Means software that respects users' freedom and community
 - Free → liberty, NOT price
 - Freedom to **run, copy, distribute, study, change** and **improve** the software
- F(L)OSS: Free (Libre) and Open Source Software
- More than 200 existing open source licenses!
 - Most popular open source licenses: BSD, MIT, GPL



- Described by Richard Stallman (president of the Free Software Foundation) in the GNU manifesto in 1985
- Copyleft principles
 - Freedom
 - **Freedom 0:** the freedom to use the work
 - **Freedom 1:** the freedom to study the work
 - **Freedom 2:** the freedom to copy and share the work with others
 - **Freedom 3:** the freedom to modify the work, and the freedom to distribute modified and therefore derivative works
 - Reciprocity
 - Derivative works must also be released under a compatible copyleft license
 - It is described as “viral” for this reason



Author: Ruben Rodriguez

Title: Richard Stallman, at LibrePlanet 2019

Source: <https://media.libreplanet.org/u/libreplanet/m/richard-stallman-at-libreplanet-2019-2113/>

License: CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>)

GPL



- GNU General Public License (GPL)
 - Created by Richard Stallman
 - **Strong** copy left
- Any software written based on any GPL component must be released (the whole source code!) as open source
 - What triggers the GPL is any call made to any code (usually a library) which is licensed under GPL
 - If you only use it internally (you don't distribute it), GPL is not triggered
- Companies are usually afraid of GPL code!
- You are not forced to publish the source code online (i.e., on Github)
- LGPL: GNU Lesser General Public License
 - **Weak** copy left

BSD Licenses

BSD

- Berkeley Software Distribution (BSD)
 - 4-clause license (the original BSD License)
 - 3-clause license (the 2nd version, the advertising clause was removed)
 - 2-clause license (“Simplified BSD License” or “FreeBSD License”, it omits the non-endorsement clause)
- Permissive free software licenses
- You can freely modify and distribute your code in the source or binary format as long as you retain a copy of
 - the copyright notice,
 - the list of conditions,
 - and the disclaimer

3-clause BSD License example

Copyright 2023 Idiap Research Institute

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

MIT License

- One of the most permissive free software licenses
- You can do whatever you want
- You only have to add a copy of the original MIT license and the copyright notice



MIT License example

Copyright 2023 Idiap Research Institute

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Multi-licensing

- Two or more different licenses
- Dual licensing
 - Usually refers to offering software under a proprietary license and an open source license
 - Examples
 - PySide2 (<https://pypi.org/project/PySide2/>)
 - Open Source (L-GPL/GPL) or Commercial license
 - MySQL (<https://www.mysql.com/>)
 - GPLv2 or Proprietary

Table of content

1. Software licensing
- 2. Open source software**
3. License your software

Table of content

1. Software licensing
2. Open source software
3. **License your software**

3. License your software

How to choose a license?

1. Ask yourself the right questions
2. Have a clear view of the dependencies you are using

How to apply the license?

1. Prepare the license
2. Add the license to the root of your project

Some questions to ask yourself

- Will my software be distributed or is it for internal use?
 - If for internal use only, no need to prepare a license
- What is the purpose of my software?
 - Open source vs closed source?
- How are the dependencies licensed?
 - Permissive vs Copyleft? Special license?
 - If you rely on dependencies, your license has to be compatible with the licenses of your dependencies
- Are you releasing data or models as well?
 - Release them separately and license them with an appropriate license (i.e., Creative Commons)

Distribution of your software

- List all your dependencies and check their licenses
 - What if you were not allowed to use some specific code in your project?
 - Try to contact the author for asking permission
 - What if one of your dependencies is GPL and you cannot license your project under GPL?
 - Rewrite the code or find another library / code which does the same, or
 - Remove the code!
- Add your copyright to headers
 - Not needed, but a good practice
- Add LICENSE.txt to the root

Third party code license

- Open source
 - Redistributing third party code
 - Make sure you comply with their license
 - Not redistributing third party code
 - Good practice: Add a section to your LICENSE.txt file to cite all third parties (Copyright + License type)
- Closed source
 - Create an “About” screen and add a link pointing to a page that lists all your third-party licenses

Datasets and AI

- Model trained on top of a dataset should be considered as a derivative work of this dataset
 - Model is subject to the limitations indicated in the license regarding derivative works
 - If a dataset is licensed as non-commercial / research only, the model will “certainly” have to be licensed in the same way
- On the other hand, the code itself is not impacted

What about ChatGPT / Midjourney / etc.?

- As the results are “generated”, they don’t already exist, so the results should not pose a problem of plagiarism!
 - Still, the situation is not clear from a legal point of view
 - Several complaints have been filed
- AI technologies can never be considered authors as only humans can be considered authors
 - It should not be possible to copyright content generated by an AI
 - But who owns the results?
 - ChatGPT → users, to the extent permitted by applicable law!
 - Midjourney → users, to the extent possible under current law!

Summary

- For any existing piece of software code you might want to use, you need a license!
 - Published on Github without a license file → no rights!
- Open source softwares
 - Permissive: BSD, MIT
 - Copyleft: GPL, LGPL, A-GPL
- Artificial intelligence
 - Model should be considered a derivative work

Quizz

- Quelle est la principale différence entre les Open Source Software et les Free Software ?
 - Vous pouvez vendre un Open Source Software, mais vous ne pouvez pas vendre un Free Software.

N'oubliez pas que le mot "free" dans free software fait référence à la liberté et non au prix.
 - Les Open Source Software sont des logiciels dont le code a été publié sur l'internet, tandis que les Free Software sont des Freemiums et des Freewares.

La publication d'un code source sur l'internet n'en fait pas un logiciel libre ! Et les Freemiums et les Freewares ne sont pas des logiciels libres mais des logiciels non commerciaux.
 - La principale différence est qu'ils ont des philosophies / idéologies différentes, mais à part cela, les Open Source Software et les Free Software sont très similaires car les deux termes se réfèrent essentiellement au même ensemble de licences et de logiciels.

Quizz

- Vous voulez que votre logiciel soit un FOSS. Comment le diffuser ?
 - Si vous souhaitez distribuer un logiciel open source, vous devez le rendre disponible sur Github ou toute autre plateforme similaire conçue pour partager le code d'un logiciel.

Non, ce qui fait d'un logiciel un logiciel open source, c'est sa licence et non la publication de son code. Sa licence doit être une licence Open Source (BSD, MIT, GPL, etc.).
 - Il n'est pas nécessaire de publier le code source de votre logiciel pour en faire un logiciel FOSS, mais vous devez lui accorder une licence Open Source.

Quizz

- J'ai téléchargé des images d'insectes sur Google image pour créer un dataset, et j'ai entraîné un modèle d'apprentissage profond sur cet ensemble de données pour faire de la classification d'insectes. Dans mon code, j'ai utilisé de nombreuses bibliothèques BSD, et je n'ai fait qu'un seul appel d'API à une bibliothèque publiée sous GPLv3. Sélectionnez les affirmations suivantes qui sont vraies :
 - J'ai le droit de publier mon projet (code source + modèles) sur Github sans fournir de licence.
 - Je dois accorder au modèle une licence compatible avec la licence que je choisirai pour distribuer mon ensemble de données, car le modèle est une œuvre dérivée de l'ensemble de données.
 - Comme je ne publie pas le code source de la bibliothèque GPLv3 (elle sera installée par les utilisateurs de mon logiciel), je peux mettre mon logiciel sous licence BSD.

C'est faux ! Même si je ne redistribue pas le code de cette tierce partie qui est sous licence GPLv3, ce qui déclenche la GPL est le moment où mon logiciel fera appel à cette bibliothèque GPLv3. Je dois donc publier mon logiciel sous une licence compatible avec la GPLv3.
 - Comme les images ne contiennent pas d'informations personnelles, je peux distribuer l'ensemble de données et les modèles avec la licence de mon choix.

Même si elles ne contiennent pas d'informations personnelles, chaque image est protégée par le droit d'auteur, et ce même si l'annotation "Copyright" n'a pas été ajoutée sur l'image, ce qui signifie que vous avez besoin d'une licence pour être autorisé à l'utiliser. Dans tous les cas, vous devez savoir quelle licence est associée à chaque image que vous utilisez avant de publier votre jeu de données, afin que vous puissiez choisir la licence pour votre jeu de données en conséquence.

Interesting resources / optional readings

- https://kapeli.com/cheat_sheets/Licenses.docset/Contents/Resources/Documents/index
 - Summary of what is Forbidden / Required / Permitted for most commonly used licenses
- <https://github.com/raimon49/pip-licenses>
 - Dump the software license list of Python packages installed with pip
- https://en.wikipedia.org/wiki/GNU_General_Public_License#Linking_and_derived_works
 - The tricky question of linking and derivative works for GPL
- https://en.wikipedia.org/wiki/Software_Package_Data_Exchange
 - File format used to document information on the software licenses

Sources

- <https://en.wikipedia.org/>
- <https://www.gnu.org/>
- <https://opensource.org/>
- <https://blog.graphqleditor.com/software-licensing-cheat-sheet/>
- https://kapeli.com/cheat_sheets/Licenses.docset/Contents/Resources/Documents/index
- <https://www.synopsys.com/blogs/software-security/software-licensing-decisions-consider-dual-licensing/>
- <https://resources.whitesourcesoftware.com/blog-whitesource/open-source-licenses-explained>

Thank you for your attention!