ACAT 2013, IHEP, Beijing, China
Round Table Discussion, Sunday, 19 May 2013, 16:30-18:00
<u>Conveners</u>: G. Heinrich (MPI Munich), Jerome Lauret (BNL), Federico Carminati (CERN)
<u>Discussion Leaders</u>: Track 1: (Gang Chen, IHEP Beijing), Track 2: n.n., Track 3: Tord Riemann (DESY)

Open-source, knowledge sharing and scientific collaboration

ACAT is with us since more than 20 years now.
HEP = High Energy Physics is yet the heart of ACAT.
HEP became more and more a kind of industry.   To some extent even in theoretical physics.
The contacts of cooperating scientists get more anonymously. This trend is supported by the internet.
At the same time, the competition became more complex.

Knowledge sharing in early times:
  – you gave a Fortran code to somebody – or you did not.
  – Or alternatively: The CERNLIB „model"; its software was nearly free.

Knowledge sharing today:
  – many ways of distribution
  – more anonymously

A collection of uncommented statements

A1 – an author says:
When I create software, I want to get cited for its use.
I a way which I define.
Sometimes by applying a GPL-type licence, sometimes by references in articles, etc.

A2 – an author or an institution says:
Our software has not to be touched by the user.
Because I guaranty for its high standards.
Because it is a standard candle [etalon] for others, and it was created in order to be so.
So, please link my software to yours, or refuse from usage.
Or, please write your interface to my package as a whole.

U1 – a user or an institution says:
I need open-source software.   And I can tell you what that is.
Because I have to adapt and develop your software for my purposes, in favor of the scientific progress.
If you do not want misuse, you cannot publish your software.   This is the only safe way. *I go this way.*

U2 - a user or an institution says:
Any software in the internet with anonymous download is open-source software.

U3 - a user or an institution says:
There are no commonly accepted rules for the use and citation of open-source software.
Of course, I need not cite the software when using it.
And I can do what I like, even if the authors claim to have a „licence".

Few intro statements

- *Academic research* includes as basic element the ***equilibrium of competition and cooperation***
  Cooperation, knowledge sharing = the ***sum of collaborative work + direct exchanges +***
publications
- ACAT topics are ***truly international***, in all respects.
- Let us discuss exclusively about ***academic research***.

And only about *academic software*.
Just to be definite.
This would exclude e.g. any reference to commercial software.

- We live in an ***international community***.

As a consequence, national law, national licences, institutional regulations are not valid automatically.
See: http://en.wikipedia.org/wiki/Berne_Convention_for_the_Protection_of_Literary_and_Artistic_Works

- We folks at ACAT work on ***long-term projects***, often ***in teams***, sometimes in huge teams with ***changing compositions***.

- *Academic researchers depend in many respects on the recognition of their contributions to scientific progress.*

We have to fight for
  - project money
  - money for PhD and postdoc positions
  - permanent positions
  - resources, e.g. clusters of comuters for large-scale calculations
  - professorships, better professorships, Nobel prizes
  - the recognition of our contributions to the scientific progress [non-monetary hopes]

etc. etc. etc.

- So, since the Renaissance epoche, the ***tradition of citation of the work of others*** [works = creations] became more and more an essential part of scientific ethics in basic research.

Either because the own work relies on the work of others, or because the work of others is directly used.
This is beyond commercial arguings, but not beyond of material interests.

- The equilibrium of ***Competition and Cooperation*** gets disbalanced e.g. when researchers use the work of others without quotation.

Quotation is a key element of establishing the ***equilibrium of Competition and Cooperation.***

- In practice, there are ***additional expectations of the creators of scientific work to their users***.

Accepted by society, as for „creators of work" in general.
Remember the ***regulations in the general internet*** for photographs, videos, music etc. etc.
One has to care about the „conditions of use" when downloading anything.
Maybe in form of licences = standardized „conditions of use".
- For software, there may be certain very specific regulations.

E.g. in German law this applies.
But there are not so many persons who assume that ***software has „no genuine scientific content"***, i.e. is not a result of creative work.
Although, in scientific practice it may happen that software is considered of „minor relevance" when

2

accounting scientific ingenuity and scientific progress.

- Because national law often cannot be applied in practice, it is of importance that researchers feel an *ethical need* to *respect the „conditions of work"*, formulated by the authors of software.

- *Part of these „conditions of work" are licences.*

Examples [please search for details by yourself]:

- *Gnu public licence* = *GPL* and the derivatives, like e.g. the *lesserGPL*. → **See below.**
  Is often used and/or recommended.
  But: seems not appropriate for academic software because does not expect proper citation.
- *Creative Commons Licence* = *CC* with derivatives. → **See below.**
  Seems to be OK, although often not recommended for software.
- *„Computer Physics Communications"* software deposit licence → **See below**.
  Was in use for decades by e.g. GEANT, MINUIT, ZFITTER, FF, etc.
  Now the users of the CPC software deposit have a choice of licence.
- POWHEG BOX: Example of guidelines which you may find short, simple and useful
  http://powhegbox.mib.infn.it/
  Licence: See http://www.montecarlonet.org/index.php?p=Publications/Guidelines
  From HEPFORGE: http://www.hepforge.org/projects

Few remarks to be thought over:
It is obvious that there is *not the one approriate licence model* for all academic software projects and their purposes.
A decision on a licence is taken together *by the authors of the software* **and others, but not by the user:**
Have in mind:

- There may be [several] *employers* [in different countries].
- Sometimes a project is *supported by some organization* (e.g. in Germany DFG – Deutsche Forschungsgemeinschaft) and undergoes demands from that, e.g. to make the results publicly available in appropriate form.
- Sometimes a project is made *in a collaboration like ATLAS or CMS*; then there might be stringent reasons NOT to make it public – due to Competition.
- Sometimes software is made in an institution which does not mind at all.

- In the view of some experts of law, these licences are *„so-called licences"*, i.e. no true licences.

This might be true because national law might have certain opinions on what is a licence.
But: Not to respect these licences is not a good concept.
Bcause the licences in use in basic research are, in the best case, a *commonly accepted frame of use of work of others*.
They are, in practice, the rules of the game.
If somebody recommends not to respect licences formulated for scientific, non-commercial, academic softeware, he/she should then explain what else is the basis of an „agreed use of the software".

- Few ACAT organizers found a round table discussion of the rules of using software just not needed, because the lawyers of their institutes care about that.

Counter argument:
Legality of use of software by third parties is not a case for lawyers.
By no means: *We are poor.   We have no time.*

- Certainly, there happen quite different situations.
- Software created in a huge experimental collaboration like ATLAS or CMS often underlies agreements of the parties. So, ideally there are well-defined partners and they certainly respect the rules.
-
-

- Published academic theory software. If it is made available for others – often it is not -, often the „others" are not pre-defined. Then one has to hope as author, that the users respect the interests of the authors. In the sense defined above.
- Unpublished academic theory software. **Like BlackHat.** Here is no problem. But do we want this? Remember: Cooperation...
- *If researchers use software of others, but do not publish their own softweare, do not quote the use of the software of others – do we accept this as a honest model of academic research?*
- Is the distribution of software in form of executables to selected circles of users preferrable? Or is it ethically forbidden?

- *Researchers in academic research are paid by society.* So there are expectations to be fulfilled:
- Honesty
- Seeking for the truth
- Making achievements publicly available

etc.

- Sometimes the *research institutions develop the idea that they are the true holders* of the scientific achievements and can - more or less free of any boundary condition – decide what has to happen with a „work", e.g. with a certain software.

Evidently, there are arguments „PRO".
But there are also arguments „CONTRA":
- Assume that a software is created by a postdoc with a scientific carrier in mind, hired for a short term. Would he/she like to have a certain amount of control? Should he/she have this this? In a collaboration? Or in general?
- Assume a software is created by researchers of N_1 Countries, N_2 institutions, over N_3 years, financed by N_4 funding agencies, published in N_5 journals/archives/etc., in changing compositions, etc. What is here the role of institutions? It might not be distinguished.
- Assume the software is supported over a longer period. By whom? An institution??? This may happen.

Independent of the copyright of institutions, however this might be defined, there is the question:
Do software authors have any copyright at all?
E.g. the right to be quoted?

- It is evident that any regulation, which is hindering the arbitrary use of software, might also hinder scientific progress.

But, on a longer term, would the absence of copyright regulations destroy the equilibrium of Competition and Cooperation? And thus, on a longer term, scientific progress?

- We did not introduce to the intro so far the terms: *Dialogue and Respect*.

It is fatal if we feel a lack of the two.
Because these two help to establish the ethical stability of basic, academic research.

- *Players in the world of rules* of using the work of others, e.g. software, are:
- countries with their national ethical and law systems
- universities and research labs
- experimental collaborations
- but also: *single researchers* who create works, i.e.: we, the participants at ACAT.

The latter sometimes are not considered as „players" because they are usually hired.
But research relies on the acceptance of responsibility by the creators of works.
Responsibility in any respect, including e.g. user support and further developments, but also the care about the fate of the creation.

- It is not so crazy if ***creators of inventions do not tolerate arbitrary use of their work***, e.g. the uncontrolled modifications.

Often modifications are needed by users, and are even supported by the authors.
But if a work serves as an etalon, e.g., it is reasonable to forbid uncontrolled modifications.
In the interest of the community.
Further, nearly arbitrary use then is nevertheless possible, but has to follow few (well-known) rules.

- There are three different categories of rights to be taken into account:
– Need of citation
– Licences and conditions of use
– Copyright

We did not discuss:

- open-source software   versus
- source-open software   or
- public software (might be executables or n-tuples or source-open software)

We did not mention:

- **Plagiarism:**
  ≈ ***Present inventions of others as if there were my own.***

- **Reproducibility:**
  Physics is a natural science and relies on reproducibility.
  Confidential software, when used for discoveries like the Higgs, makes physics a non-scieitific adventure.

What is important?

1. We authors are players in the game.

2. Conditions of use (licences) are formulated by authors, not by users.
   Make them explicit!

3. Publish ASAP = As Source-open As Possible.

4. Institutions should have compliance statements.
   And not only definitions of plagiarism and lists of sanctions.

Creative Commons licence

English:
http://creativecommons.org/licenses/by-sa/3.0/

Chinese:
http://creativecommons.org/licenses/by-sa/3.0/deed.zh_TW

The Licenses:

Attribution   CC BY

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

Attribution-ShareAlike   CC BY-SA

This license lets others remix, tweak, and build upon your work even for commercial purposes, *as long as they credit you and license their new creations under the identical terms.* This license is often compared to "copyleft" free and open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.

Attribution-NoDerivs   CC BY-ND

This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

Attribution-NonCommercial   CC BY-NC

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

Attribution-NonCommercial-ShareAlike   CC BY-NC-SA

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.

Attribution-NonCommercial-NoDerivs   CC BY-NC-ND

This license is the most restrictive of our six main licenses, only allowing others to download your works and share them with others as long as they credit you, but they can't change them in any way or use them commercially.

License chooser:
http://creativecommons.org/choose/

Computer Physics Communications Program Library

ScienceDirect

Programs in Physics & Physical Chemistry

Introduction

With the aid of a grant from the UK Science Research Council the Computer Physics Communications International Program Library was established at The Queen's University of Belfast in 1969 by Professor P G Burke CBE FRS. The Program Library was financially self-supporting but non-profit making. In 1996 the CPC Program Library became an integral part of the Elsevier Science journal Computer Physics Communications.

The Library's function is the storage and dissemination of refereed computer programs in physics and physical chemistry,whose detailed descriptions have been published in the journal Computer Physics Communications.

In addition, the Program Library also provides:

- access via the Internet to the Library programs, and to the corresponding full-text articles, for all members of institutes with a subscription to Computer Physics Communications;
- an annual subscription scheme to individual scientists who require programs relevant to their research but who do not have access through a subscription to the Journal;
- a free Science Direct Alert service;

The Library Contents

The Program Library currently holds over 2200 refereed computer programs which have been contributed by scientists from all over the world. As such it represents a major repository of computational knowledge and technique. These programs range in size from under one thousand to tens of thousands of lines of code. The contributed programs have been coded in a variety of programming languages including Algol, Ada, C, C++, Lisp, Mathematica, Maple, Occam and Pascal: however, the vast majority of contributions are

programmed in Fortran.

Each Library program is concisely described by a [Program Summary](). These are indexed under the twenty-three headings ranging from Astrophysics to Statistical Physics. An information retrieval system enables the [Subject Index]() and [Author Index]() to be explored and the Program Summaries to be [viewed]() and searched [[full search]() | [quick search]()]. A list of the most [recently published programs]() is also available.

The copyright and all other rights to each program in the Program Library remains with the program's author. By submitting a program to the Program Library the author gives the Library permission to distribute it to all those eligible to receive it under the Program Library's distribution service. All programs in the program library are covered by the [Disclaimer and non-profit use Licence Agreement](). Where a program is covered by a more restrictive Licence Agreement the conditions will be stated in the corresponding Program Summary printed in the CPC journal and distributed with the program source. The Disclaimer and Licence Agreement should be consulted before acquiring any program distributed by the Program Library. receive it under the Program Library's distribution service. All programs in the program library are covered by the [Disclaimer and non-profit use Licence Agreement](). Where a program is covered by a more restrictive Licence Agreement the conditions will be stated in the corresponding Program Summary printed in the CPC journal and distributed with the program source. The Disclaimer and Licence Agreement should be consulted before acquiring any program distributed by the Program Library.

CPC Licence:
Disclaimer and Non-profit use Licence Agreement

The CPC non-profit use licence agreement is an agreement between the author(s) of a program distributed by the CPC Program Library and the person who acquires it. By acquiring the program the person is agreeing to be bound by the terms of this agreement.

1. **This licence entitles** the licensee (one person) and the licensee's research group **to obtain a copy of the source or executable code** and to use the acquired program for academic or non-profit use within a research group; or, it entitles the licensee (one company, organisation or computing centre) to install the program and allow access to the executable code to members of the licensee's organisation for academic or non-profit use. *No user or site will re-distribute the source code or executable code to a third party in original or modified form without the written permission of the author.*

2. **Publications** which result from using the acquired program **will reference the article** in Computer Physics Communications which describes the program.

3. This licence does not permit any commercial (profit-making or proprietary) use or re-licensing or re-distributions. Persons interested in for-profit use

8

should contact the author.

4. To the extent permissible under applicable laws, no responsibility is assumed and is hereby disclaimed by Elsevier for any injury and/or damage to persons or ...