Data Policies of the International LOFAR Telescope (ILT)

Extension of the policies in the ILT Reference policy document

Version 2014-04-10
René Vermeulen

Sections 1,2,3, Appendix approved by ILT Board 2013-06-26
Section 4 approved by ILT Board 2014-04-10
Whenever a function of the PC is mentioned below, then for Director's Discretionary Time (DDT), the ILT Director takes the place of the PC.

1. Summary

- The default ILT data proprietary period for the science of the proposers is 1 year.
- All ILT observing and standard processing pipeline data remain property of the ILT. The ILT operates the LOFAR Archive; all ILT data (eventually) become public.
- The PC can change data rights and scope for each ILT project.
- Parallel/additional use of any data obtained in ILT projects requires prior approval.
- Connection and operation of add-on hardware or software requires prior approval.
2. Data Policies

2.1 The Long Term Archive (LTA)

ILT data are stored in the LOFAR Long-Term Archive (LTA), subject to technical readiness and capabilities (e.g., possibly in averaged form). The LTA is operated as part of the ILT. It has a single logical database and uniform interface, covering data stored in several physical locations (in mid-2013 these are in Amsterdam, Groningen, and Jülich).

Observatory staff decides on data flow and dwell times, and storage locations for specific datasets, applying operational considerations, but taking into account where appropriate and possible the convenience for processing and retrieval (e.g., for off-site post-processing) by science teams.

Any dataset stored in the LTA becomes publicly accessible on a pre-defined timescale, in accordance with the general policies on data ownership and any specific proprietary rights conferred by the PC.

2.2 Data ownership

2.2.1 Standard observed & processed data: ILT property

All standard data products resulting from ILT project observations, whether interferometric, beam-formed, single-station (data obtained directly on or through one or more ILT stations), or otherwise, remain the property of the ILT. All processed data products produced by any standard advertised LOFAR pipeline run on any ILT (post)processing machine (in mid-2013: either on CEP in Groningen or at the computing centres associated with the ILT in Amsterdam and Jülich) also remain the property of the ILT.

The PC grants (exclusive) data use rights for specific proprietary times to specific groups for specific science. By default the team to which observing resources are originally granted shall have the exclusive right to use the data within the confines of the science goals approved by the PC during a proprietary period of 1 year after observation. Details are given in the section on Proprietary Rights.

Science groups may not disseminate any such standard observed or processed data beyond their specific collaboration, neither within the proprietary time nor later. In particular, for such data they should avoid having any form of public “LOFAR data centre”, but should if desired refer (e.g. on their websites) to the LOFAR Long Term Archive.

2.2.2 Derived data: user property

Derived data products, on the other hand, obtained by a science group in further post-processing of ILT data beyond the standard pipeline products, or by processing of ILT data using non-standard software, or on non-ILT facilities, are the property of that group by default. By default, the group may choose to make such data public through its own means, or to keep it private, as it wishes. However, any resulting publication must acknowledge the dependence on ILT data or facilities, using a standard phrase given in the appendix.
However, science groups may propose to deliver specific derived data products to the LTA, and/or the PC may stipulate the delivery of specific derived data products by the group into the LTA on specific timescales, for example as a condition for granting observing time, proprietary access to standard data, or ILT processing time.

Derived data in the LTA will become public on timescales stipulated by the PC. However, deliveries of derived data products to the LTA will not impinge on the right of a science group to disseminate such data on their own as well. The onus of maintaining the integrity of the branch of the logical database for such derived data by default rests with the science group; however, in certain cases, ILT operations may agree to take on this responsibility for derived data delivered to the LTA.

2.2.3 Stand-alone observing: owner property

For any data originating from non-ILT observing, during stand-alone operation of a LOFAR station by its owner, that station owner decides on data ownership. However, any resulting publication must acknowledge the use of LOFAR equipment, using a standard phrase given in the appendix.

2.2.4 Possible data delivery to Stichting/CV LOFAR

ILT data may at any time be delivered by the ILT to the NL Stichting LOFAR or LOFAR-CV, for non-astronomical uses only, in partial payment for the contracted astronomical use of infrastructure.
3. Policies on proprietary rights

3.1 Default 1 year proprietary period

By default the team to which observing resources are originally granted shall have the exclusive right to use the data within the confines of the science goals approved by the PC during a proprietary period of 1 year after observation.

For data observed in a given Cycle, the proprietary period starts by default after the last data instalment of the full observing project in that Cycle has been made available to the PI (e.g., by Science Support notifying the PI that the data are ready to be retrieved from the LTA, possibly after an initial data processing run).

The proprietary time will by default apply to data observed in each Cycle separately, even for projects with Long-Term allocations.

Further data post-processing may happen at some later time on ILT facilities (e.g., CEP, Jülich, Grid computing), subject to allocation of processing resources by the PC. By default this does not extend the proprietary period on the original data already stored in the LTA.

3.2 Non-standard data use rights set by the PC

Any proposal may contain a motivated request to the PC for non-standard data use rights. The PC will grant extensions to the default 1 year data proprietary time only in exceptional circumstances.

The PC may decide, on its own initiative or prompted by a proposal, to stipulate different data proprietary rules, on reasonable grounds related to promoting the widest possible science use of LOFAR data, but taking due account of the interests of the proposing group.

The PC may in particular consider making the observed data in the LTA for larger projects public without proprietary time. The PC may also stipulate that data must be shared by different science projects in the same Cycle, or may allow a different science proposal submitted for a later Cycle to use and process data even while still in its proprietary period. More generally, the PC may make stipulations that delimit the scope or nature of observations and/or the use of data to a specific science topic, when these are clearly beneficial to optimise the overall science output from the ILT, or to create viable options for Open Skies and Reserved Access Shares next to each other.
4. “Private” and ILT use of stations and “add-on facilities”; co-use of copied data streams

4.1 Definitions and policies on “Stand-alone”, “Single-station”, and “sub-array” use

4.1.1 Sub-arrays

The LOFAR system has been designed in principle to allow the ILT stations to be partitioned into several groups, termed sub-arrays, each of one or more stations, assigned to different independent observing projects, running in parallel. Apart from the contractually agreed stand-alone periods of private use of an individual station by its owner (see Section 4.1.2), the full range of operational sub-array options is allocated via ILT proposals, leading to specific resource allocations by the ILT PC, and optimised parallel or sequential scheduling by ILT operations. This applies regardless of whether one or more stations are involved, and regardless of the observing mode used (correlation, beam-forming, or a “single-station mode” – as described in Section 4.1.3 – with standard and/or add-on facilities – as described in Section 4.2).

4.1.2 “Stand-alone” LOFAR station use by the owner

Station owners retain the right, by contract, to operate and use their station privately for at least a specified minimum fraction of time in so-called “stand-alone mode”. In this mode, operational control of a specific station is explicitly passed to the local owner, and handed back to ILT operations at the end of a pre-arranged time block. Use of a station in its stand-alone periods is fully outside of ILT control; no use is made of central infrastructure committed to the ILT. Stand-alone station use and data handling may in no way impact ILT operations, or resources dedicated to ILT use. However, owners of several stations may wish to collaborate on their stand-alone observations, with a view to joint storage, processing or analysis of their stand-alone data; ILT operations will upon request make a best-efforts attempt to schedule common stand-alone time blocks, also depending on ILT proposal pressure involving these stations. Local station owners must make private arrangements for any desired storage, transport, and processing of stand-alone data.

The owner decides on the use of the station during a stand-alone period. The owner may use the standard facilities of the LOFAR station including “add-on facilities” falling under an ILT “general-use agreement”, and any private “add-on facilities” (software and/or infrastructure) of the owner that have been approved for implementation at the station (see Section 4.3). Data from “stand-alone” operation periods are not entered into the Long Term Archive; their distribution and use rights remain at the discretion of the local station owner. However, any publication resulting from stand-alone observing must acknowledge the use of LOFAR equipment, using a standard phrase, as given in the appendix.

4.1.3 “Single-station” LOFAR station use through the ILT

By contrast to “stand-alone use”, the term “single-station modes” pertains to observations under central ILT control, in which the data are initially collected and processed on a per-

---

1 In December 2013, it is expected that the ILT use of parallel (i.e. simultaneous) sub-arrays for correlation, beam-formed modes, and “single-station modes” will increase following on from the installation of the COBALT online data processor. Specific types of parallel observing modes will successively be implemented. ILT operations staff will optimise the scheduling of the full complement of resources.
station basis, as opposed to either being cross-correlated between stations in real time (for interferometry), or being used online for (sub)array beam-forming. Single-station observing refers to all modes using any applicable standard LOFAR functionality as well as any of the add-on facilities available for a particular station, either private facilities, or facilities under an ILT general-use agreement (see Section 4.3), whenever such observing modes and facilities are allocated for use to anyone outside of agreed stand-alone periods.

As with any other ILT observations, single-station observations may involve a sub-array: either only one station, or several, or all of the LOFAR stations simultaneously. If there is more than one station participating, the stations may either all have an identical setup, or a purposely chosen range of set-ups. Offline, depending on the specific science goals, single-station data from stations in the sub-array may be combined in different ways during further processing and analysis.

The allocation of observing time and the science use of single-station data must occur through standard ILT proposals for the use of one or more stations and their resultant single-station data. The standard ILT policies on allocations and on data ownership and use apply; standard LOFAR data products obtained with single-station modes are placed in the Long Term Archive, as for any other ILT observations².

### 4.2 “Add-on facilities” to process, copy, or co-use LOFAR data streams

At some stations, privately developed software and equipment, present or planned³, makes it technically possible to access the local aggregated station data or even the individual receptor data, for local or off-site storage, and/or to carry out non-standard processing. Examples are the generation of non-standard derived data products, such as all-sky snapshots, pulsar timing data, (near-)real-time transient detections/triggers, etc. Some add-on facilities may require operation in dedicated observing time, but add-on equipment may also be designed to allow “commensal” data collection and/or processing during ILT observing, by making a local copy of the station data being observed for an ILT-allocated primary program.

Similarly, it is also technically feasible, using privately developed equipment, to make copies of primary ILT data at some intermediate stage of transport via connector or concentrator facilities² or at the central processor. And it is furthermore technically feasible

---

² In March 2014, single station modes and their standard data products remain largely under development; consequently, the Long Term Archive is not yet technically ready to store these data.
³ In March 2014, local station data co-use facilities in existence or under development include Artemis (originating in the UK, for data from one station, but already installed in other international stations; description by Karastergiou et al.) and AARTFAAC (originating in NL, for use of all receptor data on the Superterp, description by Wijers et al.). The Cosmic Ray KSP group has plans for co-use of data on the TBBs (description by Hörandel et al.). The use of the Nançay Superstation receptors in parallel to ILT observing will similarly require regulation (to be described by Tagger et al.).
⁴ In March 2014 the ILT is assisting in work to configure several GLOW station network switches so that their data streams can be transported via the “FZ Jülich concentrator node” to private facilities of the MPIfR in Bonn. This channel is intended for use during ILT time as well as during periods of stand-alone use, but will redirect rather than copy station data; its use will only be enabled when there is no data flow to the ILT central processing facilities in Groningen from the stations involved.
to obtain copies of primary ILT data at some intermediate stage of (central) processing. Such data copies could then be further transported and/or processed using dedicated local or offline equipment, also developed privately.

4.3 Proposals required for add-on facilities

Recognizing the role of add-on facilities in enabling additional science output using ILT resources, and recognizing that some specific add-on facilities were already conceived or constructed before relevant ILT regulations were drawn up, Section 4.3.1 gives general policies to ensure appropriate implementation and connection of any add-on facility. These general policies form the basis for specific initial agreements regulating the connection of each add-on facility, present or planned, to be drawn up between the ILT and the owners/developers of the add-on facility.

Section 4.3.2 gives the policies regulating subsequent specific science use of such add-on facilities during any ILT observations, or requiring the use of other ILT resources such as data storage or processing. There can be a desire for commensal use of ILT resources (i.e. co-use without impact to the primary program), or for dedicated use of ILT resources (such as allocated station observing time to use add-on facilities), or for a more mixed shared use of ILT resources (for instance, if ILT network or compute capacity needs to be split).

With regard to any add-on facility for which any use during ILT observations is desired, even if originally developed and used “privately”, the ILT can choose eventually to make it available, at least in part, to the full ILT user community through the regular ILT proposal mechanism. As detailed in Section 4.3.1, an initial period of “private use” may be agreed, lasting, as a guideline, no more than 2 years. As a goal, this will be followed by a (gradual) transition to a common-user facility, under an “ILT general-use agreement”. Any such decision will be made by the ILT after consultation with the owner, taking into account the technical and operational state of the facility, and subject to dedicating adequate ILT resources to bring the add-on facility into public operation. Proposal and allocation mechanisms for the actual use of “private” and “general user” add-on facilities for projects outside of stand-alone periods are specified in Section 4.3.2.

Some add-on facilities may only be intended for use during stand-alone private LOFAR station use by the owner (see Section 4.1.2). These private facilities will be accepted for connection and stand-alone use subject only to initial verification of basic safeguards for the integrity of the standard ILT equipment and of the ILT observations, as described in Section 4.3.1. No science proposals to the ILT PC (as described in section 4.3.2) are needed for private add-on facilities used solely in stand-alone time.

4.3.1 Initial proposals for implementation of an "add-on facility"

To establish or extend any facility to access (record and/or copy) and/or to process local station or receptor data, or data being transported to the central processing facility, or data at intermediate stages of processing, a one-time initial proposal is required. This will lead to a specific agreement, to be decided by the ILT Board, on the implementation and

---

5 For example, in March 2014 there is an incipient project DRAGNET (to be described by Hessels et al.) which aims to co-use/copy all data flowing through COBALT, for a beam-formed pulsar survey carried out on private processing and storage equipment located in Groningen.
connection of the “add-on facility”. The proposal should outline the overall science goals and describe the proposed technical setup. The ILT Director will verify that:
- There is an appropriate interface to ILT operations, and there are appropriate general safeguards on the integrity of ILT equipment.
- There are appropriate facilities for ILT staff to safeguard the proprietary nature of any ILT observations as and when stipulated by the PC.
- There is either a statement that the use will be restricted to “stand-alone” private periods, or there is an adequate policy on data rights and intended future ILT and/or public access.
- In case a “private use period” is requested in connection with intended future Cycle proposals for allocation of ILT dedicated or commensal resources (e.g., observing time) and data rights, then, before operations are initiated, these matters will be settled in the specific connection agreement. This will cover the length of the “private use” period, the nature of use during it, and the modalities foreseen for the transition to a (partial) common-user facility after it. Towards the end of the initially agreed “private use” period, a specific plan for the transition can be drawn up, with the arrangements foreseen by the ILT for the long-term operational use described in an “ILT general-use agreement”.

4.3.2 Per-cycle proposals for ILT use of “add-on facilities”

The use of an “add-on” facility restricted to “stand-alone” private time operation by the station owner does not require a (science) use proposal. For that, only an initial overall “add-on facility implementation proposal” (Section 4.3.1) is required.

a) Dedicated ILT allocations
Whenever any dedicated allocation of ILT resources or science data rights (such as observing time) is desired for operation of the “add-on facility”, this requires a regular proposal, in each Cycle, subject to competition, review, allocations, and stipulations under the normal ILT Reserved Access and PC allocation and science rights policies. This applies regardless of whether or not that facility is still in a “private use period”, restricted to proposals from the owner, or is under a “general use agreement”, where the broader community may propose.

b) Commensal ILT allocations
In case the use of the “add-on facility” is fully commensal to other allocated “primary” ILT observations (typically by generating a copy of sensor or station data), this does require a proposal in each Cycle, but the ILT PC will use these proposals only to verify that there will be no infringement on the science use rights conferred on the primary investigators of any specific ILT project.

Commensal proposals to use “add-on facilities” may either be for a very specific co-use science project, or could describe a category of co-use science projects, provided these can make use of the same ILT observing condition(s). The co-use proposal may specifically request to be allowed to be active in conjunction with a particular ILT observing project, or more generally may request to be allowed when a specific mode or setup is being used on the ILT (potentially covering several ILT projects), or may request admittance under even more generic conditions. In any case, the PC will stipulate during which specific ILT observations or in which specific ILT observing modes the proposed co-use may or may not take place (either by inclusion, or by exclusion, as is most easily defined).
When an add-on facility in “private use” is proposed for commensal ILT observing, the PC will not give a scientific evaluation or establish precedence between different competing co-use proposals, should these arise. Whenever co-use proposals are stipulated as non-conflicting and allowed by the PC, the decision on the actual use of the local facility remains at the discretion of the (local) owner having the “private use” rights. When multiple owners are involved, negotiation between them may have to be arranged by the co-use proposer(s).

However, when the “add-on facility” proposed for commensal ILT use is under a “general-use agreement”, the PC will give a full scientific evaluation, will establish precedence between different competing co-use proposals, should these arise, and will confer science data rights as for other regular proposals.
Appendix: Standard acknowledgement phrases

**Paper resulting from ILT data**

This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the ILT foundation under a joint scientific policy.

**Paper resulting from “private” co-use of ILT data**

This paper is based (in part) on data obtained from facilities of the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the ILT foundation under a joint scientific policy.

**Paper resulting from stand-alone use of LOFAR equipment**

This paper is based (in part) on results obtained with LOFAR equipment. LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON.

**Paper resulting from stand-alone use of ILT equipment**

This paper is based (in part) on results obtained with International LOFAR Telescope (ILT) equipment. LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the ILT foundation under a joint scientific policy.