# Norwegian Meteorological Institute

**MET**report

# User expectations for quality control of observations

Summary of user interviews performed for the Confident project

Anette Lauen Borg, Ketil Tunheim, Nina Elisabeth Larsgård

| Title | Date |
|---|---|
| User expectations for quality control of observations | 2022-03-17 |
| **Section** | **Report no.** |
| Division for observation quality and data processing | No. 6/2022 |
| **Author(s)** | **Classification** |
| Anette Lauen Borg, Ketil Tunheim, Nina Elisabeth Larsgård | ● **Free**    ○ **Restricted** |

**Abstract**

This report contains a summary of the user interviews conducted in preparation for the Confident project, which aims to create a new scalable and modular observation quality control system at MET Norway. We interviewed representatives of both internal and external user groups, using an interview template containing open questions about the users' expectations, requirements and wishes for a new system. In order to transform the information collected during the user interviews into a useful tool for the project planning phase, we made prioritised lists of user requirements for the new quality control system. We also grouped and summarised other user input to identify their potential impact on the project.

**Keywords**

Quality control, observations, user interviews, user requirements, metadata

_____

_____

**Disiplinary signature**

_____

_____

**Responsible signature**

# Abstract

This report contains a summary of the user interviews conducted in preparation for the Confident project, which aims to create a new scalable and modular observation quality control system at MET Norway. We interviewed representatives of both internal and external user groups, using an interview template containing open questions about the users' expectations, requirements and wishes for a new system. In order to transform the information collected during the user interviews into a useful tool for the project planning phase, we made prioritised lists of user requirements for the new quality control system. We also grouped and summarised other user input to identify their potential impact on the project.

The results indicate some broad trends, which can be summarised as follows:
- Most observation data users assume that a basic quality control has taken place.
- Few (API) users know that requests for observation data should contain a definition of quality level. They will unknowingly receive observations of unknown quality and model values along with quality controlled data.
- The users want observations that are "good enough for my use". This may be a better starting point for defining quality than an absolute scale of numbers.

- Quality is not the only factor to many users. Other priorities are time, quantity and good coverage (in time and geography).
- The availability of metadata and data provenance influences the data useability.
- Many users emphasised the need for user friendly documentation and tutorials.

# Table of contents

# 1 Introduction

This report was written as part of a deliverable for the Confident project, an internal project at MET Norway to create a new scalable and modular observation quality control system. The purpose of this text is to document the process and results of the user interviews done as part of the project preparation phase. Present and future users of quality controlled observation data were interviewed to map their requirements, wishes and ideas as well as their expectations of the new system.

The intended audiences of this report are the Confident project staff, other MET Norway staff and anyone else interested in the planning phase of a new meteorological observation quality control system.

# 2 About the Confident project

*The CONFIDENT project aims to create a new scalable and modular observation quality control system at MET Norway. The success of the project depends on how well we meet the user needs for the quality control system. The possible users range from weather forecasters to employees managing metadata to external partners who both share and request data, and each of these have their own specific needs. This report covers the part of the CONFIDENT project that collects and prioritises the different user needs, however the inclusion of a specific need in this report does not guarantee that the project will cover it in the final product. The goal is to make sure that the system we build does not block the possibility for the need to be met at a later time.*

## 2.1 Background

The domain of weather observations is changing rapidly, and the current MET Norway quality control (QC) system is not sufficiently scalable. It is beyond the scope of the old

quality system (Kvalobs) to adjust for a future with increased amounts of input data and more heterogeneous in situ observations.

The CONFIDENT project was set up to build a system that is scalable and ready to take on the increased spatial and temporal resolutions and to make more information, especially regarding data quality, available and understandable for the end users.

The new QC system will be closely interconnected with our new data storage system. The data storage system (ODA) will allow for implementation of new quality flags, and for new ways of filtering data using quality information. A new QC system in combination with ODA has the potential to better meet the needs of MET Norway and the needs of the observation data user groups, while increasing the end-user knowledge and appreciation of observation QC.

## 2.2 Project description

The CONFIDENT project aims to create a new unified, scalable and modular QC system that, together with the new scalable database storage system ODA, will form the basis of a new data flow for weather observation data. The project begins at the start of 2022 and the last deliverables are expected at the end of 2024.

The five project objectives are also the basis for the five work packages related to the project. From the project plan:

**O1 Communication:** Communicate with users to identify their metadata and data needs. Provide support to users of the new QC system. Collaborate with internal and external partners and projects.

**O2 Metadata**: Adjust metadata definitions based on guidelines for unified data management and users needs. Increase automatic metadata generation. Create a quality driven hierarchy of weather stations based on metadata.

**O3 QC flagging**: Track which quality controls have been run on incoming data, flag the data in the ODA database, and provide visualizations of the QC process.

**O4 QC process control:** Develop a scalable overarching system that can perform the quality control workflow on different types of data. Create a catalogue system that defines which tests will be run on which data.

**O5 QC code:** Develop new, and adapt existing, code to build effective and autonomous quality control modules which require limited human input. Utilize the potential of spatial statistics in the quality control system.

## 2.3   The purpose of this report within Confident

This report is part of the first deliverable (D1.1) for work package 1 (WP1, covering O1 above), whose main objective is to keep the communication flow going between the end users and the project group. WP1 can be divided into two parts:

1. Collect and identify the user requirements that need to be met by the project.

2. Inform and help the users get what they need from the project result.

Deliverable D1.1 is defined in the Confident project plan as "a technical specification of interdependencies in time, space and quality for the quality control and data flow management".

For a quality control system to function optimally, there are some (user) requirements that must be met. These requirements will often be dependent on each other, or limit each other, along several axes such as time, space and quality. An example: there is a dependency between the requirements "quality controlled observation data must be made available to users within X minutes of arrival" and "this quality test depends on additional observation data to arrive". If we plan the new system with no knowledge of this dependency, the latter requirement can limit our ability to deliver the first. The Confident project needs an overview of such dependencies to plan the structure of the new system. The idea behind deliverable D1.1 is for WP1 to collect user requirements

and put together specifications that will subsequently be used by the other work packages to plan their parts of the system. In D1.1 the specifications come in the form of prioritised lists of requirements gathered from existing and potential users of quality controlled observation data, and from our external partners. The report you are reading complements those lists, by summarising  the work process behind the lists and giving a fuller insight into all the user inputs collected by WP1.

The lists and this report are meant as input for the other work packages, to help plan the architecture and software solutions, and also to prioritise their work. Later they may also  be used by WP1 when we develop the information needed by the users to start interacting with the new system at the end of the project.

## 2.4   Limitations

Confident WP1 overlaps somewhat with WP2 (metadata). For this report, the impact of this overlap is that a list of partners and projects important to the Confident project collected partly by WP1 will not be included here, but rather in a WP2 deliverable at a later date.

This report describes interviews with users of quality controlled observation data. A few user groups were left out of this round of interviews on purpose, as we concluded that the Confident project may contact them at a later time (see Chapter 3.2)

# 3 The Confident user interviews

*How did we get the information we needed? That information was held by a diverse number of users and partners. We did not know who was using quality information today, how they were using it, or what their individual needs were. We did know how to locate these users, and so that is where we started.*

*In this Chapter, we will describe our plan for understanding users' requirements, the user groups we located and interviewed, and the method we used to understand their requirements.*

## 3.1 Why user interviews

The Confident project is complex, and there are many different user needs involved. This complexity and variety makes it hard for us to predict what user information we are looking for. To account for that unknown information, we chose to do user interviews with several open questions. Through this dialogue with the users, we could

discover and investigate subjects and needs that we did not foresee. That is why we chose user interviews rather than a survey.

The intent of the interviews was to discover the complete user needs, not in the context of today's way of working, but phrased in a way independent of current software, user interfaces and procedures. With these core needs defined, we hope it will be easier to design solutions that meet them in the most effective and efficient manner.

The main scope of the interviews was the usage of quality controlled data and quality control information relating to in-situ observational data. This includes direct interaction with the quality control system, but also interaction with any part of the data pipeline, from data ingestion to graphical user interfaces. This is because any point along that pipeline may be affected by the quality control system. If a user's needs cannot be met by a QC system alone, we still want to know about those needs to ensure we build a QC system that can support it.

## 3.2  Who was interviewed

There are many different direct and indirect users of quality information. We wanted feedback from as many user groups as possible, covering a wide variety of use cases. To locate candidates for interviews, we contacted most divisions at MET Norway, as well as external partners and users who we know are using our data today.

Among MET Norway's departments and divisions, we interviewed representatives for the following (Norwegian acronyms in parentheses):

- Development Centre for Weather Forecasting (SUV)
- Forecasting Department (VDIV)
- Observation and Climate Department (OBSKLIM)
    - Division for Climate Services (KLT)
    - Division for Maintenance of Observation Systems (OD)
    - Division for Observation Quality and Data Processing (OKD)
- Research and Development Department (FOU):
    - Division for Model and Climate Analysis (MK)
    - Division for Ocean and Ice (HI)
    - Division for Oceanography and Marine Meteorology (OM)
- Information Technology Department (ITDIV)

- Division for Geodevelopment (GEOUTV)

The following MET Norway roles and topics were also covered through our interviews:

- Klimavakten
- Human Quality Control (HQC)
- Numerical Weather Prediction (NWP) Verification
- Research and development

Interviews were also held with users working at the following external actors (Norwegian acronyms in parentheses):

- yr.no
- Norwegian Public Roads Administration (SVV)
- Norwegian Water Resources and Energy Directorate (NVE)
- Norwegian Geotechnical Institute (NGI)
- BarentsWatch
- University of Oslo Department of Geosciences

These groups, and the individual contacts within them, were chosen from a larger list of potential contacts. From this initial list, we chose users covering a large spectrum of the different use cases.

We held a total of 30 user interviews. Of these, 15 were held in October–December 2021 and 15 were held in January–February 2022.

There are some groups that we did not contact who are still important users of quality information. The reason we have not covered all of the actual user groups, is because different groups' use cases often overlap. Our aim was that as many as possible ways of using quality control information, and the related needs, should be covered by the interviews that we performed. If we realise that key user groups have been omitted, additional interviews may be held.

For the current deliverable, the main focus has been on direct users of quality information, and the developers of systems that use that information. As the technical work packages (WP3, WP4, WP5) start up, we aim to facilitate communication between the work package members and relevant colleagues concerning technical solutions and implementation.

## 3.3 The interview templates

The templates used for the interviews can be found in the Appendices linked below:

- [Interview template, English version](#)
- [Interview template, Norwegian version](#)

In the interviews, we wanted to avoid the context of the current quality control system. To achieve this, the questions focused primarily on the use of data, quality information and related metadata. Furthermore, several questions were quite open and we used these to ask follow-up questions. The aim of these follow-up questions was to identify and phrase succinctly the core needs that the user has.

The questions in the template can be summarised as follows:

- What data do you use, and what characterises good and inadequate data?
- What metadata do you use?
- What expectations do you have for quality control of data?
- How do you use quality information?
- Are you doing quality control yourself?

For user groups in MET Norway's Information Technology department, some of the questions in the regular template were less relevant. We made a modified template for these interviews, focusing more on the needs of automated systems and user interfaces:

- [Interview template for Information Technology department](#)

For the same reason, we made a modified template for MET Norway's Division for Maintenance and Observation Systems. The users interviewed from this division do not work extensively with the data itself, so this template focuses more on setup, calibration and troubleshooting of station hardware:

- [Interview template for Division for Maintenance and Observation Systems](#)

## 3.4  How we did the interviews

In most interviews, we invited one interviewee. On several occasions, we invited two interviewees together. When we invited two, they were part of the same team or otherwise worked closely together and used data in very similar ways. Due to COVID-19 restrictions, many of the interviews were held digitally, with a few held physically. All but one of the interviews were held in Norwegian.

We attended each interview with one interviewer and one scribe. During the interviews held in 2021, all three project members attended the interviews, with the third being an observer. In 2022, we only attended with an interviewer and a scribe. The roles were not set but rotated between each meeting.

Before starting the interviews, we gave an introduction to the interviewees. This included a short description of the Confident project; a mention of the aim and format of the interview; a disclaimer that we cannot guarantee that all ideas and feedback are followed up; and an interest in scheduling a follow-up meeting later in the project if need be, for the purpose of testing prototypes.

The duration of the interviews were initially set to 60 minutes. Seeing that we could use some more time, we scheduled 90 minutes for each interview starting in 2022.

# 4 The interview outcomes

*In order to transform the large amount of information gained through the user interviews into a deliverable useful for the Confident project, we had to analyse the minutes and notes. This Chapter contains a description of the analysis process and the summarised results.*

## 4.1 Interview analysis

The minutes from the interviews were studied with two goals in mind: 1) create the lists of user requirements, wishes and ideas described in Chapter 2.3 and 2) summarise the user expectations and other input for this report.

During the interviews, the users described the expected outcomes of the new quality control system in the form of expectations to the observation data and metadata after quality control. Summarised, the observation data were expected to have the following properties (translated from Norwegian):

- Observation data should be quality controlled in a well documented way
- The observations I receive must be of a quality that is "good enough for my use"
- The observation quality control and metadata must follow the standards defined by the World Meteorological Organization (WMO)
- Large and obvious measurement errors should be removed and/or flagged

- The observations should be made by well documented, well maintained meteorological stations of a known quality
- An observation time series should not contain too many holes or sudden jumps, implying minimum instrument downtime and few changes in the station setup
- The observation data should have their provenance (see below) attached
- Genuine extreme measurement values must not be flagged as errors in a time series, implying a flexible valid range
- The observations made by one instrument should match the observations made by instruments at neighbouring stations, and by other types of instruments covering the same area
- Observations should be made available as soon as possible after measurement
- Measurements must have a "good enough" time resolution for the user's needs
- The geographical coverage must be "good enough" for the user's needs
- It should be easy for the user to understand and use our meteorological observations, implying a user friendly format and documentation

The users' expectations of data, metadata and the system in general were expressed in full through a long list (in Norwegian) of all the user input (requirements, wishes and ideas) reported during the interviews. This list can easily be converted into user stories, to be used as e.g. work package boundary objects during their planning phases.

The next step was to turn the list of user input into a tool useful for the planning phase of the Confident project. The main priority was to sort the requirements, wishes and ideas in ways that would allow the Confident WPs to plan the QC system structure and to prioritise their work. In the interview template we included questions that were meant to distinguish expectations describing actual user requirements from those less necessary to the user (wishes and ideas, but still good to have). This allowed us to sort user input according to whether or not it was a user requirement. We were also able to sort the list by which inputs would directly affect the project, and which were more tangential. Both these ways of sorting were useful, but we saw the need to first sort the long list into shorter lists corresponding to related project tasks. We tried to sort according to which work package(s) the user input would affect, but we found that too often a piece of user input would clearly affect more than one work package. We finally

decided that the most useful approach would be to instead divide the user input into broad topical categories relevant to the project. These categories were developed through a categorisation exercise described below.

## 4.2   Categorisation exercise

We made an exercise to get started on the process of dividing the list of user input into more manageable categories. The main goal for this exercise was to group the pieces of user input that had something in common. In this way we could perhaps get some indication on which tasks that need to be solved together, or if there are common themes that may recommend a specific prioritisation or technical solution.

The interviews were condensed to a list of user needs or wants in the form of sentences. For the purpose of this exercise, this list was reduced further by taking out some needs that were very single use specific, or of a very similar nature to some of the user needs that were kept. This reduced list of 50 sentences was used in the exercise.

The whole project group was invited to participate in the exercise, and they were divided into two groups. Both groups got a stack of 50 cards which each had one of the user needs from the reduced list. Both groups got the same stack of cards.

Task 1: Sort the user needs that belong together into groups. Rules for the sorting exercise:

- The categories are not to be decided before the sorting begins
- Do not start "solving the need" on the cards, just sort as quickly as possible.
- Keep the number of categories to less than 10, preferably less than 5 categories.
- The card should be marked if it seemed particularly important.

Task 2: Name the groups of sorted cards. These are the final categories.

By not naming the categories before the sorting process started we hoped to avoid "getting stuck" in how today's system works, and perhaps find some new connections that can be used later in the project. The results from the exercise can be found in Appendix 1: Results from the sorting exercise

**Post-exercise sorting:**

After the group exercise was completed, WP1 made their own version of the categorisation by combining  the two groups' results  into five "common categories":

- Metadata [Metadata]
- Overvåkning og systemutvikling [Monitoring and system development]
- Brukerinteraksjon [User interaction]
- Grunnleggende tester [Basic tests]
- Avledede tester (inkludert ikke-sanntid) [Derived tests]

We then sorted all the user input into these final categories.

This exercise gave us some new thoughts about which user needs should have a higher priority, but it was also helpful in sorting which needs belong to particular work packages or perhaps belong to more than one work package.

## 4.3    Summary of the user input by category

The process of sorting allowed us to see clear patterns and associations, and so the content of the user input belonging to each category can be broadly summarised as follows:

**Category: Metadata**

Basic information about the meteorological measurements and their source that the users require or wish for

Requests for provenance of each point of measurement (what has been done to this measurement up to now? Which quality tests have been run on it and has it been corrected in any way?)

Input on how we should describe the quality of the observation data

Requests for documentation of the quality control system, and of the products made by the system

## Category: Monitoring and system development

How to monitor the quality control system and send automatic notifications of errors in order to keep a regular data flow and avoid downtime

Requests for real time observation data access

List of tools for the users to monitor the data flow through the quality control system of observation data relevant to them

Some future system demands that will need further planning and development

## Category: User interaction

User outcome expectations when searching for quality controlled observation data

Ideas on which parameters will be useful when searching for observation data

How the users would like metadata, observations and quality presented

## Category: Basic tests

A list of self contained quality control tests

## Category: Derived tests

A list of quality control tests where additional information, observations back in time or observations from other stations/instruments are needed

Input from users that basic and derived tests should be run separately, allowing users who need real time data to request observations as soon as they have cleared the basic quality control

## 4.4　Other user feedback, input and discussions

As well as questions directly probing user expectations and requirements, the interview template contained questions that would expand our knowledge of how and why observation data are used. During the interviews we would also receive information, feedback or questions that were not framed as user expectations. In this chapter we will try to present some of this information and the possible implications for the Confident project.

A number of users do not have a concise definition of what level of quality they need in order to use the observation data. They also cannot provide detailed information on which quality tests they would like to have included in a quality control system. What they want is data that is "good enough for my use". When searching for observations they would like to receive only the data that are fit for their use, although they would also like to be notified if there is data of lesser quality available. We also notice that about 20% of the users interviewed never look at quality information at all. Additionally, users ask for examples of acceptable areas of use for our observation time series, based on their quality (e.g. "this time series is of good enough quality for modelling use but not for climate research"). **Implication for the Confident project**: Should areas of use ("these data are good enough to be used for X") be our starting point for defining quality, instead of an absolute scale of numbers? The users like today's simplicity of a single number showing the data quality, but this number can be calculated differently in the future.

Most users external to MET Norway trust the observations they receive through our APIs, without setting a quality level in their request or doing their own QC. **Implication for the Confident project**: Should there be an automatic default quality setting when requests for data do not contain quality levels? Today the users will receive everything in the database, including observations of unknown quality and weather model data.

Many users perform their own quality control by manually looking at numbers, or plots of numbers. Often the reason for this manual control is that they do not know enough

about the automatic quality control. A simple count suggests that just under a third of our interview subjects know nothing about the current quality control system. Another third knows a bare minimum. **Implication for the Confident project**: Can we cut down on some of the time people spend on manual control by making the future quality control documentation easy to read and readily available?

The quality of the station and the instrument doing the measurement is often seen as more important than the quality of a single point of measurement. **Implication for the Confident project**: Can we integrate the station and instrument quality (including station placement and representativity) into an overall quality flag?

Quality is only one of our users' priorities when it comes to observation data. Other priorities are (in random order):

- Time (they need real time data)
- Quantity (they need as much data as possible)
- No downtime or other holes in observation time series
- Good geographical coverage
- Observation data from more than one type of instrument in the same geographical area

The focus on observation time series without holes or jumps is mainly due to users running data through statistical tools. Their object is to make homogenised time series, look for extreme values or do other climate research. A user also pointed out that there is a psychological effect when we share data externally; time series with holes look less reliable than complete time series.

**Implication for the Confident project**: Make sure to focus on more than quality alone when designing the new quality control system.

Aggregation of observations from one time resolution to another, e.g. into hourly values from sub-hourly values, raises additional questions relevant for the Confident project:

- Which part of the data flow system should do the aggregation?
- Should we store all aggregated parameters, or can we do some aggregation on demand?
- Is our documentation of the aggregated parameters good enough?

- Do we keep clear and consistent metadata when we make aggregated parameters? If so, are they available to the users?
- Do we run quality control on aggregated values?

Historically, some but not all aggregation has been performed by the quality control system. The new data storage system (ODA) has extensive in-built capacity for time series products such as aggregation and calculation.

**Implication for the Confident project**: At least some of these questions need to be answered in the system planning process. This process needs to involve the ODA project, where these questions have already been raised and a lot of the work is already implemented.

Documentation and tutorials: Users need many examples in different programming languages, and the examples should also include explanatory comments written into the code. If quality flags are not mentioned in the code examples, they will not be used. **Implication for the Confident project**: Examples are important and need to be well planned. Quality levels should always be included in all observation data request examples. We should keep in mind that these examples are usually the basis for all later code and data requests made by the users! The examples should cover several areas of observation data use (research, monitoring, verification of models, etc). They should be tested on users before final publication, and ideally be updated regularly.

During the interviews we asked whether the users do any quality control of their own and if they could share any code they have made. The following could be of interest to the Confident project:

- MeteoIO: From https://gitlabext.wsl.ch/snow-models/meteoio "The MeteoIO library aims at making data access easy and safe for numerical simulations in environmental sciences requiring general meteorological data." The program is made by the WSL Institute for Snow and Avalanche Research. SLF is part of the Swiss Federal Research Institute WSL.

- Wavy: https://wavyopen.readthedocs.io/en/latest/index.html made by Patrik Bohlinger at MET Norway. This is an open source package (under development) with the aim of easing the task of wave model validation.

# 5 Appendices

## 5.1 Appendix 1: Results from the sorting exercise

**Results from the exercise:**

The two groups categorised the user needs differently(as expected) but we can see some familiarities. These are the categories the two groups came up with for the list of 50 user needs:

Group 1:

- Test
- Stasjonsmetadata
- Avledet kvalitetsinformasjon/ekstra
- Anomalier
- Historisk QC
- Brukervennlighet (Ease of use)

Group 2:

- Brukergrensesnitt

- Avanserte (romlige) tester
- Tilbakevirkende kraft
- Gammel kvalitetsinfo
- Langtidsevaluering
- Bindinger mellom tidsserie
- I tråd med tiden
- Utvikling+evaluering av systemet
- Overvåkning/driftssituasjon
- Metadata (delt i enkle og avanserte metadata)

What the different groups put in each of these categories can be found in the lists of prioritised user input, as mentioned in Chapter 4.2.

## 5.2 Appendix 2: Interview templates

## Interview template, English version

**Can you give us an overview of your work tasks? Extra: Are you involved in any (external) projects?**

1. **What are "good data" to you (observations)?** *[Hvorfor-spørsmål]*

2. **Which MET observation datasets do you or your systems use?**

   *If relevant: External and internal datasets?*

3. **Under what circumstances are (observation) data not good enough for you to use?** *[Hvorfor-spørsmål]*

4. **What are metadata? Which metadata are absolutely necessary to your data use?**

5. **Do you submit any data into our systems and/or do you download/extract data?**
   *[Hvorfor-spørsmål]*

   a. *Submitting:*
      i. *What metadata are you submitting?*

         *Could you submit more metadata?*

         *Do you prefer to submit metadata together with the data, or separate?*

      ii. *What do you expect will happen to the data after you submit them?*

   b. *Download/extract data:*
      i. *What should be done to the data before you receive them? What are your expectations?*

      ii. *Are there any data you do not find when searching, that you know exist?*

6. **What is your definition of "quality controlled data"?**

7.  **What do you know about the quality control of observations done by MET today?**

    a.  *Nothing:*
        i.  *If you knew more about quality information, would you use it?*

    b.  *Some knowledge:*
        i.  *Where and how do you get quality information?*

    c.  *Advanced user:*
        i.  *What quality information do you use?*

        ii.  *Is there any quality information you think is missing today?*

        iii.  *Where and how do you get the quality information?*

8.  **In what way would you like quality information to be presented to you?**
    *[Hvorfor-spørsmål]*

9.  **Would you use quality information to filter your data searches, or would you rather have default quality settings when you request data?**

    a.  *Default:*
        i.  *What should be done to the data before you receive them? What are your expectations?*

        ii.     *If your data request returns nothing because the data that exist do not match the default quality settings, is that ok?*

   b.  *Ellers:*
        i.     *What quality information would you use to filter the data in a search?*

        ii.     *Do you need to know why data has received a specific quality code or flag?*

## 10. Do you or your systems make any choices based on quality information?

   a.  *Yes:*
        i.     *Which errors HAVE TO be flagged or removed for you to use the observations? Alternatively, which quality tests MUST be run? (Absolute minimum)*

        ii.     *Which errors would you like to have flagged or removed?*

        iii.   *Do you prioritize any other parameters than quality? (Would you wait for better quality or do you need the data as soon as possible? Would you be ok with lower quality if you get a large volume of data?)*

## 11. Do you use any MET observation data that you know are not quality controlled?

## 12. Do you run your own quality control on our MET observation data?
*[Hvorfor-spørsmål]*

      a. *yes:*
          i. *Could your code/software be useful to others?*

      b. *No, but:*
          i. *Do you use any other data that has gone through external quality control*?

**13. Is there any step or process in your use of our data that could be made easier when we renew our systems?** *(Noe i den stilen, for å få vite hva vi bør ta hensyn til dersom noen skal lage et verktøy som løser utfordringen deres) [Hvorfor-spørsmål]*

**14. Any comments, questions or suggestions?**

## Interview template, Norwegian version

**Hva driver du med?**

**1. Hva er gode observasjonsdata?** *[Hvorfor-spørsmål]*

**2. Hvilke observasjons-datasett er du/ditt system borti?**
      a. *Hvis mange:*
          i. *Kan du dele inn i kategorier/typer?*

b. *Eksterne/interne datasett?*

3. **Når ville dataene ikke være gode nok for deg til å bruke?** *[Hvorfor-spørsmål]*

4. **Hva er metadata for deg? Hvilke metadata er nødvendig for deg/dine systemer?**

5. **Sender du data inn i systemene våre, og/eller henter du data ut?**
   *[Hvorfor-spørsmål]*

   a. *Sender inn data:*
      i. *Hvilke metadata sender du inn i dag?*

         *Hvilke metadata kunne du sendt inn?*

         *Er det enklest å sende dem sammen med data, eller separat på noen måte (koble seg på eksterne metadatakilder, feks)?*

      ii. *Hva forventer du at gjøres med disse dataene?*

   b. *Henter ut data:*
      i. *Hva forventer du er blitt gjort med disse dataene?*

      ii. *Savner du noen data du vet eksisterer når du søker etter data i dag?*

6. **Hva legger du i uttrykket «kvalitetskontrollerte data»?**

7. **Hva vet du om hva vi gjør av kvalitetssikring i dag?**

   a. *Vet ikke:*
      i. *Ville du brukt QC-data hvis du visste om dem?*

   b. *Vet masse:*
      i. *Hvor henter du kvalitetsinformasjon?*

   c. *Avansert bruker:*
      i. *Hvilke kvalitetsinformasjon bruker du (ofte)?*

      ii. *Er det noen type kvalitetsinformasjon du savner i dag?*

      iii. *Henter du det via frost og/eller KDVH eller andre steder?*

8. **På hvilken måte vil du ha kvalitet presentert?** *[Hvorfor-spørsmål]*

9. **Bruker du kvalitetsinformasjonen når du søker etter/bruker observasjoner, eller vil du bare ha ferdige default-innstillinger?**

   a. *Default:*
      i. *Hva forventer du er blitt gjort med disse dataene?*

      ii. *Kan du godta at det ikke dukker opp noe data, eller mindre enn forventet, dersom du har en default-innstilling som ikke matcher din kvalitetsforventning?*

    b. *Ellers:*

        i. *Hvilken kvalitetsinformasjon vil du bruke?*

        ii. *Trenger du å vite hvorfor data har fått det kvalitetsinformasjonen det har fått?*

## 10. Gjør du eller dine systemer noen valg ut fra kvalitetsinformasjon?

    a. *Ja:*

        i. *Hvilke feil MÅ være markert/fjernet for at du skal kunne bruke observasjonsdata? / Hvilke tester må være kjørt? (Absolutt minimum)*

        ii. *Hvilke feil ønsker du at skal være markert/fjernet?*

        iii. *Er det noen krav du må veie opp mot kvalitet? (Hvis du får bedre kvalitet eller mer data om du venter en time, en dag… vil det være bedre for deg? Er det ok med litt lavere kvalitet på data hvis du får store mengder data, f.eks. Netatmo?)*

## 11. Har du tilgang til observasjonsdata som ikke blir offisielt kvalitetssikret i dag?

## 12. Gjør du noe kvalitetskontroll selv? *[Hvorfor-spørsmål]*

    a. *Ja:*

        i. *Er det noe andre også kunne dratt nytte av?*

    b. *Nei, men:*

        i. *Benytter du kvalitetssikring som er gjort eksternt/av andre enn deg?*

**13. Har du en utfordring som et nytt system kanskje kan hjelpe deg med?**
*(Noe i den stilen, for å få vite hva vi bør ta hensyn til dersom noen skal lage et verktøy som løser utfordringen deres) [Hvorfor-spørsmål]*

**14. Har du noen kommentarer eller spørsmål eller forslag?**

## Interview template for Information Technology department

**Hva driver du med?**

**1. Hvilke observasjons-datasett er du/ditt system borti?**

    a. *Hvis mange:*
        i. *Kan du dele inn i kategorier/typer?*

    b. *Eksterne/interne datasett?*

**2. Når ville dataene ikke være mulige for deg/ditt system å bruke?**
*[Hvorfor-spørsmål]*
I hvilke tilfeller får du ikke brukt data? (Formater, fil/arraystørrelse, hvor data ligger lagret, ..)

3. **Hva er metadata for deg? Hvilke metadata er nødvendig for deg/dine systemer?**

4. **Når du/dine systemer håndterer observasjonsdata, hva forventer du er blitt gjort med disse dataene?**

5. **Savner du noen data du vet eksisterer når du søker etter data i dag?**

6. **Hva legger du i uttrykket «kvalitetskontrollerte data»?**

   a. *Hvilke kvalitetsinformasjon bruker du (ofte)?*

   b. *Hvor henter du kvalitetsinformasjon?*

   c. *Er det noen type kvalitetsinformasjon du savner i dag?*

   d. *Gjør du eller dine systemer noen valg ut fra kvalitetsinformasjon?*

   e. *Er det noen krav du må veie opp mot kvalitet? (Hvis du får bedre kvalitet eller mer data om du venter en time, en dag… vil det være bedre for deg? Er det ok med litt lavere kvalitet på data hvis du får store mengder data, f.eks. Netatmo?)*

7. **Hva vet du om systemet som gjør kvalitetssikring av observasjoner i dag?**

   a. *Berører dine systemer dette?*

8. **På hvilken måte vil du ha kvalitet presentert?** *[Hvorfor-spørsmål]*

9. **Hvordan ville du helst hatt observasjonsdata med tilhørende kvalitet levert/åpnet for henting?**

10. **Har du/dine systemer tilgang til observasjonsdata som ikke blir offisielt kvalitetssikret i dag?**

11. **Gjør du/dine systemer noe kvalitetskontroll selv?** *[Hvorfor-spørsmål]*

    a. *Ja:*
        i. *Er det noe andre også kunne dratt nytte av?*

    b. *Nei, men:*
        i. *Benytter du kvalitetssikring som er gjort eksternt/av andre enn deg?*

12. **Hvordan ville du satt opp ditt drømme-kvalitetskontrollsystem (IT-messig)? Er det noen løsninger vi absolutt bør/ikke bør velge?**

    a. *Er det noe ved eksisterende systemer vi bør beholde?*

    b. *Er det noe ved eksisterende systemer vi ikke bør beholde?*

13. **Er det noen formater eller standarder i ITs retningslinjer vi bør vite om?**

**14. Er det noen prosjekter (pågående eller framtidige) på IT vi bør vite om?**

**15. Har du en utfordring som et nytt system kanskje kan hjelpe deg med?**
*(Noe i den stilen, for å få vite hva vi bør ta hensyn til dersom noen skal lage et verktøy som løser utfordringen deres) [Hvorfor-spørsmål]*

**16. Har du noen kommentarer eller spørsmål eller forslag?**

## Interview template for Division for Maintenance and Observation Systems

**Hva driver du med?**

**1. Hva er gode observasjonsdata?** *[Hvorfor-spørsmål]*

**2. Hvilke observasjons-datasett/instrument-typer er du/ditt system borti?**

    *a. Eksterne/interne datasett?*

**3. I hvilke situasjoner ser du på observasjonsdata?**

    *a. Hvilke systemer bruker du for å se på data? Og i hvilke situasjoner?*

    *b. …eller får du noen andre til å se på/analysere data for deg? I så fall, i hvilke situasjoner og systemer?*

c. Hva forventer du er blitt gjort med disse observasjonsdataene, i de ulike systemene?

d. Savner du noen data du vet eksisterer når du søker etter data i dag?

e. Ser du noe kvalitetsinformasjon i disse systemene? For instrumenter og for data.

f. Ville du hatt (mer) kvalitetsinformasjon tilgjengelig i disse systemene?

g. I så fall, hvordan skulle det vært presentert?

4. **Hvordan vet du hvilke feil som skal løses, og hvilke av disse som skal prioriteres?**

5. **Når er observasjonene ikke gode nok?** *[Hvorfor-spørsmål]*

a. Ser du feil i observasjonene som ikke andre har oppdaget? Hvordan?

b. Hvordan ville det vært best for deg å melde om og dokumentere disse feilene (til sluttbrukerne)?

6. **Hva er metadata for deg? Hvilke metadata er nødvendig for deg/dine systemer?**

a. Bruker du status-informasjon fra instrumentene/loggerene?

b. *Kan vi gjøre bedre bruk av denne informasjonen i systemene våre?*

c. *Hvilke metadata sender du inn i dag?*

d. *Hvilke metadata kunne du sendt inn?*

e. *Er det enklest å sende dem sammen med data, eller separat på noen måte (koble seg på eksterne metadatakilder, feks)?*

7. **Hva legger du i uttrykket «kvalitetskontrollerte data»?**

8. **Hva vet du om hva vi gjør av kvalitetssikring i dag?**
   (Bruker du output eller flagg fra noen del av kvalitetssikringssystemet?)

9. **Har du tilgang til observasjonsdata som ikke blir offisielt kvalitetssikret i dag?**

10. **Gjør du noe kvalitetskontroll selv?** *[Hvorfor-spørsmål]*

    a. *Ja:*
       i. *Er det noe andre også kunne dratt nytte av?*

    b. *Nei, men:*
       i. *Benytter du kvalitetssikring som er gjort eksternt/av andre enn deg?*

11. **Har du en utfordring som et nytt system kanskje kan hjelpe deg med?**
    *(Noe i den stilen, for å få vite hva vi bør ta hensyn til dersom noen skal lage et*

*verktøy som løser utfordringen deres) [Hvorfor-spørsmål]*

**12. Har du noen kommentarer eller spørsmål eller forslag?**

  *a. Spørre om Inspeksjonsrapport-prosjektet, og om det er noen andre systemer eller prosjekter vi bør ta hensyn til!*