

SDQA Scope Issues/Actions/Decisions? for DC3b

The SDQA sub-system for DC3a primarily emphasized automated evaluation of metrics, which is most applicable in an operations environment with short- (visit level) cadence. The implementation of the DQA Ratings (i.e., the instantiation of metrics with computed values, thresholds, etc.) is less well tuned to the contemplative assessment of a release production, however, and certain issues were identified in that context:

- Constructing SDQA ratings in pipelines requires a higher level of effort than the alternative of merely persisting metadata, which pipelines will be doing anyway.
- SDQA ratings as implemented is a process-heavy endeavor: the system is insufficiently flexible for adding new metrics, changing threshold values, or for performing QA on quantities that we did not have the foresight to compute in a production.
- There is a strong need to visualize the results not only of metrics, but also of the data themselves. This is partially addressed with the SQuAT tool.
- SDQA ratings cannot currently store arrays, such as vectors or distributions.
- Users have a strong need to compute and store derived quantities (that were not persisted in the DB) during the course of QA analysis.
- It is unclear how to compare the results of a QA analysis from one production run to another.

As a result of multiple follow-on meetings a few decisions have been reached, and others have been identified, for DC3b:

- SDQA ratings will be de-emphasized (or possibly, not used) in DC3b.
- SDQA will focus on enabling dynamic, curiosity driven QA assessment. This is well matched to the perceived needs of science team analysis at the end of DC3b.
 - ◆ Develop example user programs to document access to Science DB for power-users.
 - ◆ Emphasize use of 3rd party visualization tools for DC3b, including (SQuAT, VisVO, TOPCAT, etc.) if possible.
 - ◆ Build into design the concept of a local User data store (i.e., a Sand Box) for intermediate results, which the user can access just as easily as the Science or Facility DBs
 - ◆ This anticipates what we think will be a natural evolution of metrics & code back to the production as experience grows.
- Explore the possibility of running user programs within afw framework, and also of running pipeline components as user programs.
- The scope and depth of the support for DC3b QA analysis will be focussed on the most critically needed assessments, based on experience with DC3a QA results.
- Support for comparison of QA results across production runs is also critical; it may be that this requirement will be satisfied by cross-matching SciDB catalogs, but the requirements need to be fleshed out.

There are some action items on the way forward:

- Additional use cases (and/or additional detail on UCs in-hand) will be solicited from the apps teams, science teams, and project science/technical staff to illuminate the most pressing QA analysis patterns
- Dick will glean additional metrics from contributed use cases and construct an "official" list for DC3b. [This is well along, but need to finish incorporating input from teams received just prior to SDQA meeting in June.]
- SDQA team will coordinate with Apps teams to ensure that desired metrics are included in "official" list, that they are actually computed during production runs, and (working w/DAWG) are persisted in the DB.

- SDQA team will coordinate with SAT on needed architecture changes. See cartoon of high-level ideas: [wiki:DC3BSDQA:DQA_Arch.pdf?](http://wiki.DC3BSDQA:DQA_Arch.pdf?)
- Tim will contribute some Python scripts he has used to support his DC3a QA activities, and will solicit same from others who were performing QA. These will be posted on the Trac in an appropriate place as a start on power-user documentation.
- All use cases that were developed for SDQA will be posted to the Trac. [Mostly done, except for two that were contributed recently. The set would benefit from some copy editing.]
- Finish analyzing requirements for DC3b, update design in EA model, code, fix, etc....
- Research needed on certain technical areas, including broadcasting data among applications, support for user-defined local data stores, survey of data visualization tools in the community...

DC3b Scope decisions

[The following was written during the initial DC3b planning meeting on 19-20 May 2009.]

- No firm decisions made. The most significant conclusion was that we need additional "face time" to really sort out what we want to do with the SDQA Team for DC3b and the future. This has been scheduled for June 11 at IPAC as an all-day meeting. Suzy and Jeff to coordinate logistics. We did make a list of possible directions/changes for DC3b:
 - ◆ Possible Tasks for SDQA in DC 3b captured during Breakout Session on May 19
 - ◇ Tasks originally scoped by SDQA for DC3b
 - Statistics/aggregation/thresholding and persistence of higher level metrics
 - Build analysis support tools for display of SDQA data, i.e. the GUI Tool SQuAT which is based on the existing PTF QA tool prototype which Russ demoed
 - ◇ Tasks which are consistent with the existing philosophy of SDQA but which were not in the original DC3b baseline
 - Generate more metrics? coordinate with other tasks/experts including science collaborations and calibration
 - Design QA rating equivalent structure for data derived from images, including Data Release QA esp. source/object level QA
 - Modify persistence method (Metadata vs ratings structure)
 - Modify way thresholding is implemented
 - ◇ Tasks which deviate from current direction
 - Build other PV type analysis support tools
 - Stop doing ?SDQA? and do pipeline validation

Issues

The main issue seems to be that IPAC set out to develop a system to explore how we will handle operational Quality Assessment of science data during the survey. We felt this was a fairly difficult problem which will likely require innovative approaches to distilling the data down to a level at which a human or small number thereof can assess if we are collecting "good enough" data or not. On the other hand, the DC3a team is feeling the lack of tools to evaluate how the processing is working on small amounts of data right now with the aim of validating/tuning the processing (what we would tend to call Pipeline validation) the two things have overlap, but not complete overlap.

Actions The meeting to be scheduled June 11 is the main action. This is to Suzy/myself and Jeff.

[Jeff's Notes fro SDQA management meeting at IPAC](#)

DC3B Task List for SDQA (7/29/09)

Implement SDQA tool functionality to support DC3B SDQA goals (Laher)

Modify persistence method to use metadata (Laher)

Implement threshold comparison (Laher)

Identify SDQA metric(s) for each pipeline stage (stage owners)