ATL-SOFT-2006-003

Review of ATLAS Software Documentation

(February 8-9, 2006)

Norman McCubbin[1], Hans von der Schmitt, Silvia Schuh, Uta Stoesslein, Rick St Denis, Siva S.02/

Last Revision: 2006-05-02

Executive Summary

The ATLAS documentation has improved in scope and quality as a result of several, related stimuli during 2004/05:

- the need to produce the Computing TDR;
- reviews of the major software domains;
- a growing number of end-user physicists;
- dedicated effort being allocated to prioritized aspects of the documentation.

The panel considers this overall trend to be highly encouraging and commends it. The panel is optimistic that by building on this recent momentum ATLAS documentation can reach a satisfactory state for the start of data-taking.

We note that to maintain this positive momentum as we move into data-taking requires a clear understanding of scope and priorities over the next ~2 years. Furthermore a continuing maintenance effort will be required; otherwise the quality of the documentation will decay.

For each topic (General Issues, Workbook, New Pages, Twiki, HyperNews, Doxygen) the report gives Findings of Fact, Observations, and Recommendations. The major recommendations are given below[2], but the full text needs to be read to understand their proper context:

General-1: Oversight of ATLAS documentation as a whole should be part of the mandate of a member of the ATLAS top-level management.

General-3: For the software documentation the SIT should produce a resource-loaded, prioritised list of tasks to cover the period up to and including first data-taking.

---

[1] Contact: N.McCubbin@rl.ac.uk
[2] Some subsidiary recommendations are omitted in the Executive Summary, which is why the numbering is not strictly consecutive.

General-4: A Documentation Co-ordinator should be appointed with responsibility for managing the ATLAS software documentation.

General-5: Each major software domain should appoint a documentation 'contact' to assist the Documentation Co-ordinator, and the general documentation effort.

Workbook-1: The ATLAS Computing Workbook initiative should be continued and supported as category 1) and layer 1) documentation. This will require an editor for the foreseeable future.

Workbook-3: A Workbook covering more advanced topics for the growing number of more experienced users should be prepared in good time for data-taking.

Workbook-4: The Workbook approach should be considered for other areas, e.g. Physics analysis. (see also recommendation Twiki-3)

New Pages-1: The conversion of top-level pages into the 'new' style should be completed, and the top-page co-ordinator should be supported in her efforts to achieve this.

New Pages-2: Ongoing effort should be provided to ensure these top-level pages are maintained, with the active co-operation of the documentation contacts in the various domains.

Twiki-1: A Twiki 'minder' (probably the same person as the Documentation Co-ordinator) should be appointed to monitor and control the Twiki pages.

Twiki-2: The sitemaps for the Twiki and the top-level pages should be harmonised where possible and appropriate.

Twiki-3: Twiki pages should be gathered into other 'workbooks', where appropriate. The tools developed for the Computing Workbook will be of use. (See also recommendations Workbook-3 and Workbook-4)

HyperNews-3: We recommend consideration of a mechanism to implement restricted HyperNews access to a forum, which could be useful for dealing with e.g. early drafts of papers and reports.

HyperNews-4: We recommend that, if feasible, the clear layout of threads in bulletin boards be incorporated into HyperNews, so that threads are more easily distinguished from each other.

Doxygen-1: We recommend that the Documentation Co-ordinator organize system-by-system examination of the documentation aspects that have come out of the recent software reviews.

Doxygen-2: We recommend that the various areas of failure, particularly where robustness of the build is threatened, be examined.

1) Introduction

A two-day review of ATLAS software documentation was held at the instigation of the SIT Co-ordinator, Fred Luehring, at CERN on February 8th and 9th, 2006. The membership of the review panel, charge to the review, agenda, and presentations are available from http://agenda.cern.ch/fullAgenda.php?ida=a06492. (For ease of reference the membership and charge are given at Appendix 1.) Input was requested from the ATLAS software community prior to the review, and at the open session of the review in the morning of February 8th. Use of a HyperNews forum (https://hypernews.cern.ch/HyperNews/Atlas/get/docReviewFeb06.html) to post comments and information was encouraged.

Software documentation in its entirety spans differing user groups and aspects of the ATLAS experiment: documentation for beginners, documentation for experts, documentation for physics analysis, data-quality monitoring, etc. Furthermore, there are the distinct, but related issues of its presentation (primarily through web pages), and the tools used to produce and maintain it. The charge to the review was to concentrate on the documentation for beginners, primarily the Workbook, the top-level web pages, and the tools, i.e. the areas the SIT considers to be its primary responsibility. In this report we do, however, flag some wider considerations, which seem of particular importance.

After a set of presentations in the morning of February 8th, the panel prepared a set of questions and presented them to the documentation team late in the afternoon. The team's answers were presented to the panel and discussed in the morning of Feb 9th. The questions were broken down into the areas: 'General Issues', 'Workbook', 'New Web Pages', and 'Tools', and these topics are presented in that order in the following sections of this report. For each major topic there are sub-sections covering Findings of Fact, Observations, and Recommendations.

The major recommendations are repeated in the Executive Summary.

2) General Issues

2.1 General Issues: Findings of Fact

Software documentation is of fundamental importance to an enterprise like ATLAS, but must compete for resources (primarily individuals' effort) against many other important, perhaps more pressing demands. It is evident that good documentation can make a direct contribution to obtaining results in the most efficient way. Hence, the provision of good documentation must be viewed as a high-priority task for the collaboration.

The ATLAS software documentation represents the efforts of many people over several years. For example, CMT and Athena have been documented since their inception.

We record here (and do not repeat elsewhere in this document) the panel's appreciation of the work done by all who have contributed to the software documentation, through writing primary documentation and/or through organising and editing it. We refrain from giving names for the simple, if inadequate reason that it is highly probable that we would omit people who have made important contributions.

The panel also wishes to record its thanks to those who contributed directly to the review by making presentations, answering questions, or by submitting comments.

The documentation has improved in scope and quality as a result of several, related stimuli during 2004/05:
- the need to produce the Computing TDR (http://cern.ch/atlas-proj-computing-tdr/Html/Computing-TDR.htm);
- reviews of the major software domains (https://twiki.cern.ch/twiki/bin/view/Atlas/ReviewsPlans2005);
- a growing number of end-user physicists;
- dedicated effort being allocated to prioritized aspects of the documentation.

The panel considers this overall trend to be highly encouraging and commends it. The panel is optimistic that by building on this recent momentum ATLAS documentation can reach a satisfactory state for the start of data-taking.

We note that to maintain this positive momentum as we move into data-taking requires a clear understanding of scope and priorities over the next ~2 years. Furthermore a continuing maintenance effort will be required; otherwise the quality of the documentation will decay.

2.2 General Issues: Observations

The panel invited Fred Luehring (FL) to give his views on the related topics of the scope of the documentation project, the short- and medium-term priorities, and the effort required.

We agree with the general thrust of FL's reply, that, from the perspective of the SIT, there are three categories of documentation:
1) Documentation for which the writing and maintenance are entirely the responsibility of the SIT (e.g. top-level pages, Workbook);
2) Documentation that is co-ordinated by the SIT, and SIT feels responsible for ensuring that (up-to-date) documentation exists, but it is written by others. (e.g. documentation for Athena, Reconstruction)
3) Documentation for which SIT provides infrastructure (tools, web templates etc), and possibly advice, but is not responsible for its production. This category might include Data Quality Monitoring guides, physics analysis "how to" and procedures, and detector documentation.

The corollary is that someone other than the SIT needs to assume responsibility for the important "category-3" documentation. The panel considers that "oversight" of ATLAS documentation *as a whole* should be an explicit part of the mandate of someone at the highest level of ATLAS management, since the need for

documentation spans much of ATLAS, not just software and computing. (See General Issues: Recommendation 1) By "oversight" we mean the responsibility of seeing that the documentation gets produced, not of producing it.

Furthermore, the panel considers that this member of ATLAS management and the SIT co-ordinator (and possibly others, as they think fit) should define the ATLAS documentation project, as a whole, so that it is clear who is responsible for what, what additional management structure is needed, if any, and how documentation fits with related topics. (See General Issues: Recommendation 2) The panel's reasoning behind this recommendation is that "documentation" has, a priori, broad scope and ill-defined boundaries, and hence a definition of responsibilities and interfaces is needed to help ensure that all documentation is produced as required. (A small example that illustrates the haphazard nature of the organisation of the ATLAS documentation as a whole: there is a Documentation tab on the top ATLAS page, http://atlas.web.cern.ch/Atlas/index.html. Is it intended that this should be the point of entry to *all* ATLAS documentation? If not, is its scope made clear?)

The above considerations lie outside the strict charge to the review, but the panel felt that the issues they raise should not be ignored.

Focussing now on the those aspects of documentation for which the SIT has a direct responsibility (category-1 and -2 above), the panel notes that there are ~1.25-1.5 fte's of effort currently working on Workbook, new web pages, and general web-page support. To assess future requirement, the panel recommends that SIT produce a resource-loaded, prioritised list of tasks for the next ~2 years. (See General Issues: Recommendation 3) The panel expects the effort required over the next two years is likely to be at least as large as the current effort.

The panel agrees with FL that one of the people providing this SIT-based effort should be appointed as Documentation Co-ordinator, with responsibility for managing the ATLAS computing documentation[3], reviewing it, checking that it is current and accurate, and cajoling documentation from those areas where he/she considers it to be inadequate in some respect. One of the main functions of the Documentation Co-ordinator is to be the Twiki "minder", as discussed in section 5.2. (See General Issues: Recommendation 4 and Twiki Recommendation 1)

We mention here (and discuss further in later sections) the two other documentation tasks that will require a significant amount of effort for the foreseeable future: a co-ordinator for the top-level pages ("top-page co-ordinator"), and an editor for the Computing Workbook. We note that these documentation tasks would count towards M&O Computing Category A and this should help to fill the posts.

In order to assist in the production and maintenance of 'category-2' documentation, the panel recommends that each major domain should appoint a documentation 'contact'. (See General Issues: Recommendation 5)

---

[3] By 'computing documentation' we mean documentation for computing and software topics that do, or could, come under the Computing Twiki https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasComputing ,
i.e. not Detector, Physics,…

The three categories of documentation discussed above differ according to who is responsible for producing and maintaining the documentation. These categories correlate to a considerable extent with different 'layers' of documentation, which differ according to their target audience:

1) Documentation of broad scope that will therefore be used by a large number of people. The user community will expect this documentation to be accurate, verified, and up-to-date; i.e. it can be trusted. Some form of 'change control' will be needed. Care must be taken that the writing is clear, and the 'look and feel' of the pages should conform to general schema. Current examples are the Workbook and the top-level web pages, but general documentation that falls into this layer will grow with time; for example strategies and procedures for physics analysis.

2) Documentation for a narrower area, likely to be of primary interest to a small group of experts. The requirements on accuracy and being up-to-date are similar to layer 1), but the need for 'change control' and conforming in style are less rigorous.

3) Evolving documentation, where ideas are being exchanged and information is changing quickly. Areas of the Twiki and HyperNews are the environments in which this happens. Valuable information is being generated here, and may be 'harvested' into layer 1 or 2) once a certain degree of stability is achieved.

Finally, the panel notes that the hardware implications of the documentation should not be overlooked: the value of good-quality documentation is dissipated if it is not available essentially 24/7. Accordingly, we recommend that the service requirements with respect to disk space, machine Input/Output rates, and necessity for mirror backup be evaluated and potential hardware issues be assessed; see General Issues: Recommendation 6.

2.3 General Issues: Recommendations

General-1: Oversight of ATLAS documentation as a whole should be part of the mandate of a member of the ATLAS top-level management.

General-2: The overall ATLAS documentation project should be defined in terms of scope, responsibilities, and management structure. This definition should be done by the top-level person identified in General-1 and the SIT co-ordinator.

General-3: For the software documentation the SIT should produce a resource-loaded, prioritised list of tasks to cover the period up to and including first data-taking.

General-4: A Documentation Co-ordinator should be appointed with responsibility for managing the ATLAS software documentation.

General-5: Each major software domain should appoint a documentation 'contact' to assist the Documentation Co-ordinator, and the general documentation effort.

General-6: At an appropriate time the SIT should assess the hardware implications of the documentation.

3) Workbook

3.1 Workbook: Findings of Fact

In his presentation to the review panel, the Workbook editor, Steve Lloyd (QMUL), said that the Workbook is intended to provide:
1) an introduction to the Atlas computing environment;
2) a consistent set of instructions to allow a newcomer to ATLAS to generate, simulate, reconstruct, and carry out basic analysis using Athena;
3) a reference manual for analysis-related material.

The Workbook was first released in September 2005. The current version (3 Feb 2006) corresponds to software release 11.0.3. It gives a brief overview of the Atlas experiment, explains how the off-line software is organized, what are the basic packages used, how to set up the Athena environment, and gives detailed instructions on how to get the Monte Carlo full chain for physics analysis running on different platforms (lxplus, batch queues, Grid). The Workbook also contains a glossary of some acronyms and abbreviations in common use in Atlas. Thus the Workbook already fulfils most of 1) and 2) above.

It is the intention that the Workbook should evolve continually to meet the changing circumstances, e.g. cover access to real data once the LHC starts.

The Workbook is implemented as a collection of Twiki pages, which are gathered from a large number of contributors and sources. These are then edited into a coherent set of pages, with uniform "look and feel", by the editor. It is also possible to convert the entire workbook, or sub-sections of it, to PDF.

3.2 Workbook: Observations

The Workbook has proved extremely useful, particularly for newcomers to the ATLAS software. It should be continued and supported. In the notation of the previous section (General Issues), it is both category 1) (SIT supported and produced) and layer 1) (of broad scope and general usage).

It is clear that bringing documentation of this type together into a single set of pages, combined with the fact the editor was himself a newcomer who tested out all (or most) of the documentation, has added significant value. We believe the Workbook will continue to need an active editor for the foreseeable future.

The panel is optimistic that the Workbook has achieved "critical momentum", so that those who are able to provide information or documentation that is currently missing (and there is some) will feel motivated to do so. Clearly the Workbook editor should have the full and active support of the proposed Documentation Co-ordinator to get gaps in the documentation filled.

It would be particularly useful, for both beginners and others, to have an up-to-date list of contacts for various computing topics gathered on to one page, and included in the Workbook.

Although the ATLAS software is becoming more stable, it is still the case that some newcomer-level instructions change from (major) release to (major) release, and it seems prudent to assume that this will continue for some time yet. The panel and Workbook editor agree that it would probably make the Workbook awkward to read and too cumbersome if *all* previous releases that might be of interest were covered in the current version. The preferred solution, we believe, is to maintain separate (archived) versions of the Workbook for earlier releases.

The workbook concept could of course be applied to 'harvest' other sets of documentation into a coherent set, and there is a probable requirement to cover more 'advanced' cases as an increasing number of users become more experienced. However, the panel felt that the current Workbook should not expand to try to cover "everything" to prevent it becoming too long; other topics (e.g. topics for the experienced user, analysis-related topics or detector topics) might be better handled by a separate workbook. The issue of one or several workbooks becomes almost moot in terms of web pages, but does affect the length of the printed version and the workload for an editor. In this spirit we suggest that the current Workbook should be referred to as the ATLAS Computing Workbook.

3.3 Workbook: Recommendations

Workbook-1: The ATLAS Computing Workbook initiative should be continued and supported as category 1) and layer 1) documentation. This will require an editor for the foreseeable future.

Workbook-2: We recommend that the ATLAS Computing Workbook for earlier releases be maintained as separate 'editions', and it is not attempted to cover all releases in the latest edition.

Workbook-3: A Workbook covering more advanced topics for the growing number of more experienced users should be prepared in good time for data-taking.

Workbook-4: The Workbook approach should be considered for other areas, e.g. Physics analysis. (see also recommendation Twiki-3)

4) New Web Pages

4.1 New Web pages: Findings of Fact

The topmost web pages for ATLAS Offline Computing are being checked for content, obsolete information is being removed (archived), and fresh pages are being produced, or old ones updated, to a new, common design. The work is being co-ordinated and carried out by Traudl Hansl-Kozanecka, the current "top-page co-ordinator".

To date the task is more than half completed, but significant sets of web pages still need 'treatment'.

Considerable effort has been put into the design and implementation of the new pages to give the pages a common 'look and feel', and to ease maintenance. Key features are:

(1) Separation of style and contents: A common cascading style sheet (CSS) is used which offers (optional) navigational elements: a sidebar with a choice of sets of links to related pages, and a header with links to the site map and to often-used pages, including the Workbook. Using the .css provided, the contributor does not have to concern him/herself with any other 'style' tags.
(2) The use of the <div> tag is encouraged, in preference to frames and tables. This helps to reduce problems with different browsers.
(3) URL-addressing is relative to a base tag, which helps keep the pages easily relocatable.
(4) Use of php is encouraged, though not mandatory, and instructions are available.

4.2 New Web pages: Observations

The areas to be covered by the new web pages are, almost by definition, 'category 1/layer1'. Once set up, these top-level pages would not be expected to change very often, although the links pointed to by them may change more frequently[4].

The cleaning up and removal of obsolete material is as important as the generation of the new pages[5], and this maintenance must of course continue. Even a small fraction of such high-profile pages becoming out of date affects the credibility of *all* the pages.

The approach followed seems appropriate for this set of pages, i.e. common design, use of "best practice" (e.g. HTML compliance, use of css, php) to write the pages, and the content is reviewed. The pages still "to be done" (see the site map http://cern.ch/atlas-computing/webAssets/sitemap/sitemapColumnFrag.php) should be done.

The css approach seems to work well for the new pages, and contributes to the common "look and feel", which is an important factor for these pages. Indeed several of the features of the design and implementation should be recommended for other sets of web pages.

What other information should be migrated into the 'top-level' set is a matter for case-by-case judgement. The top-level pages of course can, and do, point to Twiki pages among others, and migration from Twiki to top-level may be appropriate for mature, important pages. This may be significant for e.g. long-term maintenance, but should be essentially seamless for the user.

---

[4] These top-level pages are loosely analogous to base classes in C++; they may not do a whole lot, but they matter.
[5] There *is* something worse than no documentation, and that's incorrect documentation!

4.3 New Web pages: Recommendations

New Pages-1: The conversion of top-level pages into the 'new' style should be completed, and the top-page co-ordinator should be supported in her efforts to achieve this.

New Pages-2: Ongoing effort should be provided to ensure these top-level pages are maintained, with the active co-operation of the documentation contacts in the various domains.

Tools

We turn now and for the rest of this report to various tools that support the documentation in differing ways. Good tools facilitate the production and maintenance of the documentation, and the panel was asked explicitly to consider the Twiki, HyperNews, and Doxygen. All three of these products are in current use in ATLAS, so the discussion of the tool and the ATLAS use of it are inevitably intertwined.

5) Twiki

5.1 Twiki: Findings of Fact

The Twiki is an open-source product[6] supported by CERN and is in wide use by many groups at CERN. The one in use is one of many variants of Wiki. The links on Twiki pages consist of "bumpywords", which are strings such as "TrackingPage" and hence by nature are flat and unstructured.

The Twiki is a powerful and democratic documentation tool. It is easy to create and edit Twiki pages with a simple formatting. Within ATLAS it has exploded in popularity with over 1800 links, most of which have been modified in the six months since September 2005. It is a huge resource in ATLAS. It has nice tools for making summaries and searching. It has also been the tool used to implement the ATLAS Computing Workbook.

The ATLAS Twiki is aimed at all ATLAS members, with different sets of pages for major topic areas. This review focussed on the Computing pages (https://uimon.cern.ch/twiki/bin/view/Atlas/AtlasComputing), though some of the comments apply more widely.

The Wiki philosophy combines easy and open access with the remarkable notion that anyone registered with the Twiki can edit anything, not just their own contributions.

---

[6] http://twiki.org/

It is therefore by its nature chaotic and it is not entirely clear how best to organize it.[7] It does provide for automatic revision control so that incorrect modifications to a page that are can be rolled back. (This feature has been used in ATLAS and does work.) It also provides the capability to exclude malicious users[8]. It is accessible via secure login through a web browser where ATLAS users can use their lxplus logins.

As part of his work on the Workbook, Steve Lloyd has written software that examines the Twiki links and has found there are many dead pages. He has also observed that there are many out-of-date pages with approximately 240 orphaned pages, i.e. they have no links from within the Twiki. So far the contents have been monitored on a volunteer basis by Ed Moyse and a few students, who may send mail or fix things if there is a repair needed. All Twiki pages contain the name of the author and the last update date. Some contain a "Reviewed by" line at the bottom of the page with the date of the last review.

This flat hierarchy enforced by the Twiki bumpyword syntax leads to problems of uniqueness, causing names to get very long as unique prefixes such as "Workbook" or "RStD" are used. A version of the Twiki offering more structure is in beta test at CERN.

There is no formal migration to new versions of the Twiki, and no experience of doing so since its implementation in the last year.

5.2 Twiki: Observations

There is a (fuzzy) boundary between Twiki and "static" pages: Twiki is particularly appropriate for collaborative writing of documentation and gathering of information, with the expectation of fairly rapid evolution and change, at least initially; static web pages are appropriate for information that will be valid for some time, and that has been checked carefully for accuracy etc. Thus the ATLAS home page is not, and should not be, a page that can be freely edited by anyone in ATLAS, but a description on how to understand, for example, some new algorithm on beta release is a good candidate for a Twiki page. It may also be beneficial to organise a set of mature, well-validated Twiki pages into a coherent set ('book'), as has been done for the Computing Workbook, which is just a set of Twiki pages.

We emphasise again that the most important issue for the end-user is the validity and trustworthiness of the documentation, not the underlying tools used to produce it. We return to this point below.

Editing Twiki information consists of operating with a cursor in a text box, using a very simple HTML-flavour syntax. (For display, the Twiki software renders the text into proper HTML, on the fly.) For anything more sophisticated, it is required first to prepare the text in an external editor and then to cut and paste the results into the Twiki editor box. The simple Twiki syntax requires that spaces, carriage returns and

---

[7] Ward Cunningham, who came up with the wiki idea, describes it as "The simplest online database that could possibly work." It is interesting to note that the authors of the King James Bible, which has been called "the greatest work in prose ever written in English", used something of the wiki approach, though the technological implementation in the early 17th century was of course different.

[8] One would hope that this would be unnecessary in a community like ATLAS, but the wiki design and concept makes it vulnerable. Vandalism of the Wikipedia is an increasing concern.

special characters are significant. The Twiki also offers a simple upload feature for external files and figures with a size limit of 10MB. Normal HTML can be imported into Twiki text, and will display properly (almost always) when viewed as a Twiki page, and therefore HTML features (eg. access control) and the powerful suite of HTML editors and tools can be used to prepare Twiki pages.

In order to view the page one has to exit the editor and either examine the saved page or look at it in a previewer. This is more awkward than editing a standard HTML page in an editor and pressing reload in a browser to check the results. The version of the Twiki that is in beta test at CERN also offers a WYSIWYG, but there is no feedback at present on the utility of this editor.

The fact that there are a number of different flavours of Twiki leads to some difficulty for users who are working with various Twikis in various contexts (teaching, local University communication, other experiments).

There is nobody formally assigned to support the ATLAS implementation, although CERN does support the deployment of the basic Twiki software. The lack of a need to update the Twiki in the last year implies a low maintenance load.

The Twiki has proved very popular, and useful. The ability to just "put things there" is proving a very valuable way of tapping into the community's collective knowledge. The Twiki is here to stay in ATLAS.

Nevertheless there is a tension between the intrinsic "anything goes" philosophy of the Twiki, and the obvious requirement by ATLAS that documentation should be trustworthy. Without *some* monitoring and control of the Twiki pages, we fear they will over time degenerate into a large mass of pages, undifferentiated between the good and the bad. The effect of wrong, typically out-of-date, information is amplified by a large factor throughout ATLAS and large amounts of effort could be wasted. This is the standard problem of documentation, and there are no "silver-bullet" solutions. However, we consider that it is clearly part of the role of the Documentation Co-ordinator (see General issues) to oversee the Twiki pages, a function we refer to as the Twiki "minder".

The minder should validate the Twiki pages, and periodically cull (presumably archive) those that are no longer of use. Clearly the minder cannot be expected to do this by personally checking them all in detail. All pages must have "Maintainer", "Last significant update" and a "Reviewed <date> by..", and these can be used to indicate potentially out-of-date pages. The documentation 'contacts' (see recommendation General-5) would also assist the minder in this task, and help to ensure the quality of the pages. The minder may wish to solicit effort to develop tools that assist him or her in this task. (e.g. Steve Lloyd has offered to provide a link checker.) Such tools might include a hit-count summary, scanning of dead links, and monitoring of depth of links.

The minder should consider whether, and if so how, to split the Twiki in some way between "managed" and "unmanaged" pages, with the implication that managed pages have a higher "trustworthiness and validity" rating, and "managed" pages might for example be validated and culled more frequently. As noted in the section on the

New Web Pages, migration of key, relatively stable information into the pool of top-level pages may also be appropriate.

Since the top-level pages will certainly point into the Twiki, 'harmonisation' where possible of the structure of the ATLAS top-level pages and Twiki is recommended. (Compare   https://uimon.cern.ch/twiki/bin/view/Atlas/AtlasComputingSiteMap   with http://cern.ch/atlas-computing/webAssets/sitemap/sitemapColumnFrag.php)

This minder should also create and maintain the list of malicious users, though we would hope this is not an onerous task. The minder would also be responsible for developing a plan for deployment of new versions of the Twiki, and working out a safe migration that keeps the Twiki live 24x7x365. The minder would need to be able to evaluate new beta versions and do testing to ensure that the new version is compatible with the previous one.

5.3 Twiki: Recommendations

Twiki-1: A Twiki 'minder' (probably the same person as the Documentation Co-ordinator) should be appointed to monitor and control the Twiki pages, as explained above.

Twiki-2: The sitemaps for the Twiki and the top-level pages should be harmonised where possible and appropriate.

Twiki-3: Twiki pages should be gathered into other 'workbooks', where appropriate. The tools developed for the Computing Workbook will be of use. (See also recommendations Workbook-3 and Workbook-4)

6) HyperNews

6.1 HyperNews: Findings of Fact

The SLAC HyperNews is "a discussion management system providing both the functionality of mailing list systems as well as the web interface of web forums. Discussions are organized into forums and, depending on the configuration, each forum may have an email gateway." Depending on the access settings, the user may need to give a password in order to access the forum lists.

SLAC HyperNews has been developed by SLAC from the original HyperNews, and has been used extensively by BaBar, STAR and H1, e.g. accumulating up to ~250k postings for BaBar.

HyperNews replaces email for discussion threads, allowing a user to browse postings in a convenient way without the overhead of each user having to sort through or organise his or her mailbox messages. It has the potential to replace the use of mailing

lists for certain purposes (e.g. to announcement of meetings), and some experiments have done this, but we have not had the time to consider the implications of this in detail. It is not designed to replace the global information finding on the Offline Computing webpages, Twikis and documents in CDS, which are meant for already established information, rather than discussions.

6.2 HyperNews: Observations

HyperNews allows easy threading, archiving of postings, and searches of the full archive. The ATLAS HyperNews is an interface to the Simba mailing lists where a forum *NAME* corresponds to the mailing list hn-atlas-*NAME*@cern.ch. Direct references to other material (e.g web pages) are implemented via urls. Contrary to most web-based bulletin boards, it is possible to point directly to an individual posting within a thread.

Currently, it is not possible within ATLAS to restrict a HyperNews forum to a sub-set of people, which would allow restricted discussions and tracking of exchanges for e.g. analysis topics, committee work, paper drafts etc.[9]

It is also not possible to attach material to a posting on HyperNews, although allowing attachments is under active consideration by the developers. Links (urls) can be given, but must be given as the full url, unless the HTML format is used in which case an alias can be given using the usual HTML <A > </A> tag.

HyperNews is proving very popular. Since December 2005, HyperNews has been used at CERN by CMS (about 40 forums[10] with ~300 users) and ATLAS (about 30 forums). There is IT support for HyperNews, and SLAC HyperNews is a "particle-physics product", which allows close interaction with and ATLAS influence on the developers. We also note that a SLAC group is about to join ATLAS, which is likely to be beneficial for our use of HyperNews.

A user can bookmark a link to see recent HyperNews postings, selecting the number of days to look back, and which forums to search.

The panel concludes that HyperNews has a number of attractive features and the potential to develop into a powerful tool for the collaboration.

Experience from other experiments suggests that a certain amount of maintenance and 'housekeeping' of HyperNews will be required, at the level of ~5% FTE.

6.3 HyperNews: Recommendations

HyperNews-1: We recommend that the issue of long-term support for HyperNews, currently provided by SLAC, IT and Peter Elmer (CMS), be addressed.

---

[9] It is of course for ATLAS, not this review panel, to decide on policy in this area, i.e. how open or restricted such discussions should be.

[10] We abjure the pedantic 'fora', as does HyperNews.

HyperNews-2: We recommend a quick decision on the document-attachment policy – if and how it should be implemented.

HyperNews-3: We recommend consideration of a mechanism to implement restricted HyperNews access to a forum, which could be useful for dealing with e.g. early drafts of papers and reports.

HyperNews-4: We recommend that, if feasible, the clear layout of threads in bulletin boards be incorporated into HyperNews, so that threads are more easily distinguished from each other.

7) Doxygen

7.1 Doxygen: Findings of Fact

Doxygen[11] is a freeware-based automatic documentation system best suited for application programmers. It is advertised as working for C++, C, Java, Objective-C, Python, IDL (Corba and Microsoft flavors) and to some extent PHP, C#, and D. It automatically provides UML diagrams and the method call sequences. It is structured around the concept of packages. A simple syntax added to comment statements in code allow for more in-depth documentation. Doxygen also supports Latex syntax in comment statements in the source code, an option that is particularly useful for documenting mathematics. Finally, it is possible to include diagrams in the source code comments.

The author, Dimitri van Heesch, maintains Doxygen, and is quite responsive. Doxygen is used by many experiments and in industry as well. A list can be viewed on the Doxygen site. ATLAS has to maintain the configuration file and the scripts that generate automatically the documentation. ATLAS also has to do the set-up to connect to external packages, once these packages use Doxygen for their documentation. A main page using few Doxygen tags together with selected HTML tags provides overall structure for the web pages for the documented package.

The use of Doxygen was integrated into the ATLAS Software environment and CMT by Christian Arnault, and this will continue to be supported by the CMT team. It can be used by including a standard make target, *doxygen*. Documentation is generated with one directory 'html' per package. The documentation generated for any package is automatically linked to the documentation of other related packages. Access to the generated documentation is available through a Web service: http://cern.ch/atlas-php/docs/docs.php, where one can select the project or release and browse the documentation in the hierarchy of packages. It is possible to link to external packages although this is not done automatically. A template for the main page exists, although much of the ATLAS code already implemented Doxygen before the standard main page was introduced. The documentation is not automatically distributed into a release.

---

[11] http://www.stack.nl/~dimitri/doxygen/

7.2 Doxygen: Observations

Doxygen provides a very useful service to those users who are in effect application programmers and, depending on the dedication of the developer providing the comments, to users who wish to know the intention and implementation of algorithms. It fills the gap between the general user who wishes to drill down to some specific method that may be causing problems, and would use LXR to do this, and the programmer who is an expert at a few packages and uses ViewCVS to examine updates on a line by line basis. This middle ground is one in which the application programmer wishes to know what methods are available and what the structure of the code is. It therefore offers a valuable tool to ATLAS.

All code can benefit from Doxygen documentation. Code descriptions will always be extracted and are useful. By adding C++ comments, the documentation becomes even more useful. Examples can be seen at https://uimon.cern.ch/twiki/bin/view/Atlas/DoxygenDocumentation. In addition to inserting comments, one can use special Doxygen key words to add structure and more detail, e.g. author in the header or parameters of a function. Dependency diagrams are produced automatically.

The main pages and so-called "related pages" (all pages in doc/*.txt) are a general introduction to a package, whilst the detailed information (graphs and text) is for developers or anyone who wants to understand the code in more detail. A standard main page for each package would seem to be desirable, and templates exist, although it would be a large effort to try to standardize all the already existing code.[12] Such a decision to re-work existing software would have to be taken by software management.

Because the author of Doxygen has been responsive, updates and releases have not been a problem.

Some identically named packages in the software release confuse Doxygen and cause the links to be incorrectly formed. It is also possible if one makes a mistake in Latex syntax to bring down the nightly rebuild as the Latex processor simply returns to the cursor.

Python is supported by Doxygen, but this has not been tried out in ATLAS, and the question of how best to document Python files needs to be addressed.

The degree to which Doxygen has been used to extend the documentation beyond the basic default and the use by the packages varies. For example, about 90% of the EDM is documented using Doxygen, whereas usage by tools is known to be far lower. However, we have not examined systematically its use across the packages, and it is not clear if there is a good match between the carefully documented packages and those for which users most need it.

In GLAST this patchy use of Doxygen has been very successfully overcome by systematic code and documentation reviews. Good examples and careful tutoring of

---

[12] As of April 2006, ~20% of packages have a main page. (private communication Traudl Hansl-Kozanecka.)

individual groups, plus clear statements on recommended style proved very successful. We have observed that if more than a minimal amount of documentation (for a package) is obtained from Doxygen, then the code is in fact reasonably commented. Thus the load on a programmer who follows good programming practices is minimal: good code will be well commented and hence Doxygen will automatically pick up those comments and generate good documentation.

We believe it would be beneficial for the Documentation Co-ordinator to organise a light-weight assessment of the documentation issues that arose from last year's set of reviews. The exercise should take note of and be based on the successful experience of GLAST. We emphasize that the aim of this exercise is to enforce good commenting and hence good coding practice.

We have observed that it would be convenient if the code that one reaches within Doxygen were marked up and connected to LXR; however, this seems not to be technically feasible.

7.3 Doxygen: recommendations

Doxygen-1: We recommend that the Documentation Co-ordinator organize system-by-system examination of the documentation aspects that have come out of the recent software reviews.

Doxygen-2: We recommend that the various areas of failure, particularly where robustness of the build is threatened, be examined. (e.g. it must be possible to run Latex in batch/quiet mode.)

8 LXR

We briefly mention that LXR is available in ATLAS and documents code without any action on the part of the developer. Experience shows that this offers an invaluable tool.

We recommend that support for this be included in the tasks for which the proposed documentation manager is responsible and that the FTE load and agreement by support be pursued by the documentation manager.

Appendix 1

a) Membership of the Review Panel

Rick St Denis (Glasgow)
Norman McCubbin (RAL; Co-Chair)
Hans von der Schmitt (MPI Munich; Co-Chair)
Silvia Schuh (CERN)
Uta Stoesslein (DESY)
Siva Subramania (Indiana)

b) Charge to the ATLAS Software Documentation Review on February 8-9

The review is intended to cover all areas of documentation for the Atlas offline software project including (but not limited to):

- The new Atlas Offline Software web pages
- The Atlas Workbook
- The Atlas Computing Twiki
- The use of Doxygen by Atlas
- The possible use of HyperNews by Atlas

The review is intended to be a "light-weight" with the final report ready (or nearly ready) at the end of the review. The reviewers are expected to evaluate critically the program of documentation that the SIT has in place and comment on the following points:

- Is the current documentation suitable for use by the world-wide Atlas collaboration?
- The effort level: are there enough people involved?
- The time frame: will the documentation be ready in time for commissioning?
- The coherence of the documentation: are there any gaps or inconsistencies?
- Is the documentation suitable for related activities such as training and user support?
- Does documentation seem suitable for Atlas users of all types
- (users, developers, beginners, experts, etc.)?
- Is there sufficient control of the documentation (e.g. the Twiki)?
- Does it seem reasonable to transfer some or all mailing lists to the HyperNews System?
- Are there any overlooked areas of documentation?
- What should be the long-term plan for maintaining the documentation?

The reviewers are not expected to provide in-depth comments on the content of the more technical documentation, which would normally be of interest to developers and experts. They should comment more extensively on the quality of the user-level and beginner-level documentation.

The intended audience of the review report is the SIT leadership, the Computing Project management, Atlas Physics Project management, and the top-level Atlas Management.