How Good is your Content Quality?

Martin White, Managing Director, Intranet Focus Ltd

martin.white@intranetfocus.com

I wrote on this important theme in issue 3-4 of eLucidate last year but I would like to revisit the issues I raised in order to articulate and expand on a much wider range of topics.

1. Content quality management

Having not worked for a law firm for the last fifteen years I am now working for two of the world's leading firms at the same time. As you might expect there are both similarities and differences, and both are quite fascinating projects. However there is a common problem that both firms face, and that is that the content quality of their intranet content is below the level that would be appropriate for any firm, especially so in a major law firm. If you look at the individual content items on the intranets it might be quite hard to spot low quality content. However it becomes painfully obvious in search results, where tens if not hundreds of items are supposedly listed in decreasing order of relevance.

The other firm, which is in the process of a migration to SharePoint 2013, has set a target of only migrating high-quality content and then maintaining this level of high quality in the future. That is a very laudable objective but it begs the question about how content quality can be measured. Some years ago a multi-national engineering company commissioned me to develop a set of content quality guidelines and I have been using them, occasionally somewhat modified, ever since.

2. Four key principles

In no particular order there are four key principles of information quality.

a. Purpose

Every piece of information on an intranet should be there for a purpose. Before publishing, the content owner (who may not be the publisher) should consider the potential value of the content they are publishing and write it in a way that will help an individual employee or a group of employees to use and share the content.

b. Discovery

Intranet users will find the information they need by browsing through the site navigation, by carrying out a search or by setting up a profile to alert them to information as soon as it is published. The role of a content publisher is not just to create information but also to ensure that the people who will benefit from it can find it.

c. Trust
Every piece of information on the intranet will be used in some way to make a decision that could affect the reputation of the firm. Every user of information has to be able to trust it implicitly or if they have any doubts they are able to check with the content owner.

d. Ownership
It is the responsibility of all employees to ensure that content that they own is maintained in a way that it meets the purpose for which it was published, that it is findable and that staff can trust that it is accurate and valid.

3. Standards and guidelines

A standard is an absolute. Either an element of content meets the standard and can be published, or it does not meet the standard and cannot be published. A standard must be carefully defined so that there can be no dispute that the standard has, or has not been met. A guideline is a recommendation. It can be ignored and there is no sanction that can be applied.

As an example, a title such as:

Presentation at the 2014 Berlin workshop – New opportunities at sea…

…might meet the standard for a title but does not take account of a guideline on titles which might suggest that in fact a better title might be:

Expansion of marine contract capabilities in Singapore in 2015-2016

The standard title is not incorrect but neither is it at all useful as a means of helping a user decide on the potential value of the content item.

Successful content governance models have:

- A minimum number of standards which are rigorously enforced
- A set of guidelines which are illustrated with examples of good and poor practice
- Certified training procedures that ensure that publishers have a justification and a reward for using guidelines

As far as possible standards should be device-independent but guidelines should take into account the differences between desktop, tablet and smartphone devices.

4. A minimum acceptable quality standard

From the principles above there are three minimum standards for content

- It has a title
- It has a date of publication
- It has a content owner

A good quality title is important because we use titles as a filter, on a page, in a navigational list or a search results page. Search software also is biased towards a title. Perhaps counter-intuitively a title should not contain a version number. This is because only the latest version should be on an intranet! However there will need to be both a standard and a guideline. The standard states that a title must be present. An associated guideline sets out some characteristics of a good title. Individual publishers may have different views on what constitutes a good title, which is why the standard only relates to its presence or otherwise, as these cannot be disputed.

If you want to see an example of what happens when title management is missing just undertake a search of the Ofcom website.

The date should indicate the date on which the content owner passed the content as being fit for publication. In effect before this date the content did not exist. It is not the date of publication or the date of migration. In due course there should be a review date but for now the content owner certifies that a) the content is valid and definitive and b) the content owner will ensure that the content is updated and republished should it become invalid. At the same time the earlier version should be removed from the active server. For an example of an unusual approach to date management run a search on Syngenta and you will see that all the results carry the date of the last working day, a result of a server management issue.

Unlike web pages the name of the content owner is itself a very important indication of quality. The name signifies that the content owner is putting their reputation on the line for the quality of the content. The user, if they wish, is able to check the credentials of the owner and also contact the owner for additional information. If the owner is not with the firm then it is the role of the manager of the owner to appoint a new owner. No item of content can be owned collectively by a department or by 'Admin'.

To strengthen the definition the owner must have a corporate email, a corporate internal telephone number and/or be listed in the staff directory.

With this information a user is able to assess the content in three steps

1. Does the title suggest that this is at least close to the information I am looking for?
2. Does the date indicate that this is reasonably current information, something that I can also verify with the owner if needed?
3. Does the owner of the content have the authority (in expertise terms) to publish this content, which the user can always verify through the people directory?

5. Content quality guidelines

In the table below are brief descriptions of parameters that could define content.

Topic	Summary of scope
Audience	Style and language-level should be appropriate to the intended readership
Authors	Defining the authorship of content
Copyright	The copyright owner of 3rd party content should be identified
Date formats	Dates should be unambiguous
File types	Guidance on the applicability of web page, Word, pdf etc. formats
Images	Guidance on the use of images
Language	Guidance on 'corporate' English
Links	Ownership of content also entails ownership of links, and links should also be managed with care
Metadata	Metadata schema descriptions. This is a massive topic in its own right, especially when taxonomy management is considered.
Mobile content	Content likely to be used extensively on mobile devices should be written in an appropriate format
Names	How employee names should be presented
Owners	The owner of the content may not be the author
Page-level identity	Any individual pages of content should be able to be uniquely identified from metadata on each page
Protective marking	Security management for content
Record	When and how should content be a declared record
Related content	Ensuring that related content is identified so that content can be placed in context
Re-use	Where content is re-used the origin of the content should be declared
Review	The review period of content
Scanning and structure	Presenting content, especially on a web page so that it can be scanned in an F-pattern
Style sheet	Extent of linkage to corporate brand guidelines
Titles	Ensuring that titles are consistent and informative
Translated content	Linking to original versions of translated content
Validity	Any validity issues – i.e. for use in US only
Version control	Consistent version control designations

The challenge with deciding how comprehensive to make a set of guidelines is that there has to be balance between meeting these guidelines and not putting too great a workload on content publishers who are almost certainly carrying out the publishing work in their spare time. Another aspect to consider is how different categories of content (e.g. videos and PowerPoint presentations) may need to variants on these guidelines. It is difficult to be definitive about which guidelines are important, and which may need to be expressed as both guidelines and standards.

The benefits on fundability are shown in the results from the Findwise Enterprise Search and Findability Survey 2014. The charts below show a significant improvement in search performance where metadata is managed well and again where there is a taxonomy.
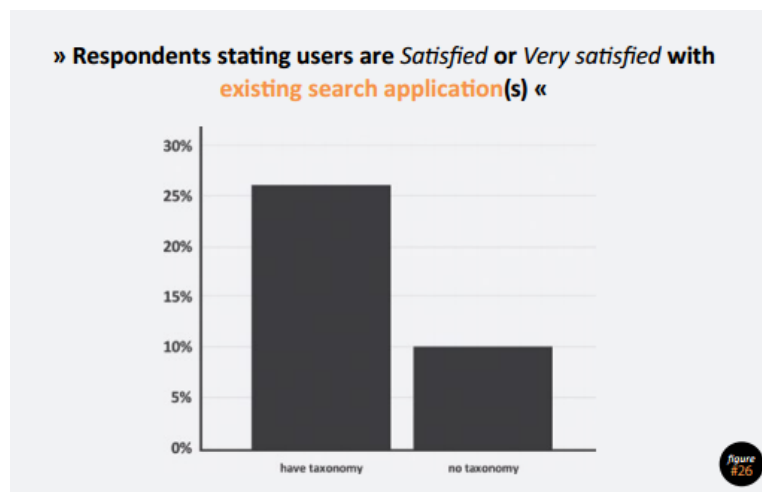
6. Managing content curation
The process of content curation (which includes publishing, revising, reviewing and deletion) is far more complex that most organisations comprehend. There is a superb paper by Stephen Dale entitled Content Curation: The Future of Relevance in *Business Information Review*, 2014, Vol. 31(4) 199–205.

In the abstract the author comments:

"We're all creating content, as originators or commentators, which is then shared and re-shared many times over. The resulting cascade of information requires new content organization and consumption techniques, and the disciplines, competencies and skills of content curation are now critical. This article argues that effective content curation requires real-time technology and tools used by knowledge domain experts who can interpret and add insight to content."

This paper is essential reading for any web or intranet content manager.

7. Training content publishers and owners

Content publishers need to be trained in how to contribute high-quality content, and I have found it useful to set up at least a two-level certification covering 'routine' content contribution (uploading Office, HTML and PDF files) and 'advanced' contribution for (e.g.) images and videos. However it is not enough to train the publishers. Content owners also need to realise that they have a crucial role to play in understanding the complexities of content contribution and being able to work in partnership with publishers.

8. Staff directories

The quality of the information in staff directories can be very variable. In the case of one of the law firms only the lawyers in the firm have profiles. To find information on 2,300 business support staff requires a search through LinkedIn. To me it seems that rarely is any thought given to the information that should be presented in a staff directory. In the case of the other law firm the same information is given in the profiles on the website as in the internal directory even though this information is used by a different group of people for a different purpose.

Even agreeing on how a name should be presented is not an easy problem. Take a look at this briefing paper from Basis Technologies (Site registration required).

For further reading see:

- Creating Intranet Content
- Organising Digital Information for Others
- SharePoint Content Authoring Quick Guide
- Quality-Biased Ranking of Web Documents
- Dublin Core Metadata Initiative
- Website Product Management

If you would like more details of the content quality guidelines mentioned above please email me at *martin.white@intranetfocus.com*.