Proposal for Handling Old Pages

PROBLEM: Tiki.org has pages that have old, out-of-date information and pages that were created as test pages or for other reasons by users and have content that isn't useful for the site. These pages may lead to unwanted results in web searches or give site visitors a bad impression of Tiki.

PROPOSED SOLUTION : Put pages that shouldn't be visible to search engines or anonymous visitors in an Admin-only category, and put pages that are old but might have "historical value" in a category visible only to Registered users.

(Although it's worth looking at other ideas, this is proposed rather than "starting fresh" with tiki.org wiki pages, which would seem to require more data storage resources (another Tiki installation?), person-power, and time to sort many pages to repopulate tiki.org back to a proper level fairly quickly. And the site would lose the impression of stability and continuity that comes from seeing content created 18 years ago and steadily updated, etc. Also, starting fresh would mean requested old pages would produce a 404 error, whereas with the categorization solution, the pages would display an explanation for why the content isn't shown, via the category alert box.)

It's proposed that...

1. Pages that are basically without current value are categorized as "Retired" and are given Admin-only access (since we have the policy to not destroy data).
2. Pages that have some merit but we don't want search engines or the anonymous public to access them are categorized as "Archived" and have Registered-only access.
3. Pages that have a mix of currently-valuable content and out-of-date content don't need to be categorized — they can just be edited to delete the old content (which still exists in the history views). It might be good to have a standard alert box for such pages saying something like "Check the page history for older versions of this page," or maybe just one such notice somewhere to cover all pages. If the person analyzing the page sees that there is old, unwanted content but doesn't have the technical knowledge to edit it, the page could be put in "Archived" at least temporarily and there could be a "Technical Editing Help Needed" page listing such pages that other people could check and hopefully edit the page in question.

Analyzing pages

The idea is to not spend a lot of time (actually, to spend almost no amount of time) examining and analyzing pages to try to sort out valuable content from not-valuable within a page — just put the page in one of those two categories if the content isn't relevant to currently supported Tiki versions or otherwise isn't of interest regarding the purposes of tiki.org.

Process details

- If it's difficult to decide between putting a page in "Retired" or in "Archived", perhaps put it in

"Retired".

- Pages that have old content, not accurate or helpful to readers regarding currently supported Tiki versions, but have some merit, would go into "Archived".
  - (This is subjective for sure — maybe some patterns/guidelines would become apparently after doing some pages.)
- There would be an alert box in the pagetop module zone, module visibility set to the category so would appear on every Retired or Archived page, that the content is old and maybe of questionable relevance and is being kept for historical purposes.
- With this approach, the situation could gradually be improved if people who are interested spend a little time from time to time.
- Suggested approach: go through tiki-listpages.php sorted by hits to check pages that have gotten relatively more hits but shouldn't be shown now.

Assumptions regarding site performance

This solution can be done with the tools we have now, but assumes that the old/bad pages aren't a drain on resources as they just idling in the database. If they are a resource drain, to the extent that it's a problem we need to fix, then this categorization wouldn't be the solution and something else should be done such as actually removing the pages from the database and storing them somewhere else, but then more resources are needed to do that and it's hard to see how it could be done incrementally the way simple page check and categorization can be.

Page redirects

Until search indexes are refreshed, the categorized pages may still get hits, resulting in a permission-denied error. Should the pages redirect to a similar content page? Maybe this isn't workable because the page may not have a similar page to redirect to that's a satisfying result for the search user. Also, because the page will have the pagetop module alert box explanation, the search user will know the reason for the permission denied. (Maybe the "Permission denied" message could be updated and toned down if it sounds too severe.)

Monitoring new pages

New pages that monitoring needs to look out for aren't created very often, so there isn't a lot to do regarding those; and they're visible in the recent-pages module — it's mostly the old pages that need to be dealt with.