Joe Lollo
LIS 530
Professor Joe Tennis
2 February 2023

Quest 2: The Wayback Machine

Time Periods Selected for "Water" on Wikipedia:
October 2004
August 2008
February 2012

Categories Present (directly pasted from page):
2004:
Categories: Forms of water | Liquids | Beverages | Nutrition | Materials

2008:
Categories: Water | Natural resources | Oxygen compounds | Hydrogen compounds | Oxides | Hydrides | Liquids | Solvents | Beverages

2012:
Categories:

- Hydrogen compounds
- Triatomic molecules
- Oxygen compounds
- Solvents
- Inorganic solvents
- Liquids
- Oxides
- Beverages
- Water

Reflection:

One of the central tenets of the website Wikipedia, an open-access, comprehensive "free encyclopedia," is the site's category system, which aims to "[enable] users to browse sets of related pages" when reading a specific page on the site (Wikipedia Content Bible). These categories are added by users to best describe specific pages, and when clicked on (interacted

with), lead to a separate page containing an index of all the pages under that category alongside a short statement describing what the category is meant to encompass. I chose to use the Internet Archive's Wayback Machine to look at categorization of the "Water" page under three different time periods with a four-year gap, namely 2004 (the site's inception), 2008, and 2012, to observe the continuity and change over time in Wikipedia categories, which are documented in the below paragraphs.

I believe that, of the three category lists, the 2008 selection is the one that best encompasses the way that "water" is communicated on Wikipedia. When looking at both the 2008 Wikipedia entry and its latest revision in January 2023, the first sentence refers to "water" as "a natural inorganic compound with the chemical formula H2O." This sentence alone checks off five categories: "Hydrogen compounds," "Oxygen compounds," "Hydrides," "Oxides," and "Natural resources." Moving forward, the 2008[1] entry refers to water being a "natural liquid at standard temperature and pressure," and a "source of nutrients for human consumption," leading to its inclusion in the "Beverages" and "Liquids" categories. Moreover, water is also described in the 2008[2] page as "acting as a solvent in living organisms and solutions," justifying the inclusion of the "Solvents" category (Wikipedia). My opinion is that, while the 2012 categories are more specific, the 2008 categories match the page best, as both the page and categories both refer to its chemical composition and its status as a solvent and source of nutrients.

Users' browsing experience over time, at least when looking only at "water," seems to be fairly consistent, as there are very similar categories across the three entries. While the page's content is fairly consistent, there is a move towards more specific categories as the years of Wikipedia's hosting progress, which makes navigating to the category pages and to all of the other pages under that category much easier. Additionally, the composition of the "water" page over time, just one part of general improvements in Wikipedia's web design, shows how good web design and – categories are more clearly identifiable through a divider or box at the bottom, separating the links from the general page, in the 2008 and 2012 versions. In the 2004 page, it is much harder to find categories, as while I knew their location intuitively, they are not located in a place where new users would understand that they are categories. The 2008 and 2012 versions

[1] Although a version of this phrase still exists in the 2023 version, it uses different wording, and the 2008 entry is quoted here.

[2] See footnote #1, this is also a direct quotation of the 2008 entry on "water."

both have a divider or box separating the links to the categories from the general page, under a header saying “Categories,” showing a much easier way to access the categories.

“Liquids” and “Beverages” are both consistent across all three time periods, which shows that the common viewpoints on water being a liquid in a natural state, said multiple times in the Wikipedia page, and something that is drinkable and consumable by humans. The view of water as a compound of hydrogen and oxygen is not necessarily shown in the 2004 pages, but the 2008 and 2012 pages do show “Hydrogen compounds,” “Oxygen compounds,” “hydrides,” and “oxides,” showing a shift towards a more scientific and professional definition of “water.” The 2004 category “Materials” is a strange inclusion, especially because it is no longer on the 2008 and 2012 pages – this makes sense, however, since because the entire world is made of materials, this is a very broad category. The 2008 and 2012 versions have more specific categories than something as broad as “Materials,” which encompasses many things in the world besides water, which is why they had to nix some of the less specific categories – “Natural Resources” from 2008 is slightly more specific, and provides links to natural resources rather than a mix of natural and manmade ones, but it was removed from the page in the 2012 edition, showing a divergence from that classification. Wikipedia’s categories over time generally show a trend towards specificity, which shows that the website values accurate, comprehensive information on pages.

Wikipedia’s categories function like an inverted index or database, as they list and display terms that are connected by a broader category – since readers know about the defining characteristics of “water” after reading its page, they might be interested in learning more about other natural resources, other hydrogen compounds, or even other beverages to compare and contrast them. This is why I believe the specific categories are a stronger representation of the page and its connections to other content on Wikipedia, as while a broader category would contain more links, a more specific one would contain “better” links that are closer to what a user is searching for. The 2008 page has the strongest categories because of how it encompasses everything that the page describes it as being, and thus “water” is indexed in places where the pages most similar to it in subject are.