Odd-ball idea for web site capture

Hi Y'all!

I'm a post-Master's grad student at U. of N. Texas, writing a paper on preservation metadata (especially PREMIS). I was struck by what I think's a great approach to capturing web sites, but I can't find anyone who is using it.

Browsing Library of Congress's (LC) archived websites available online, I stumbled on/over an Indian Catholic Bishops' site, which was collected as part of the worldwide online reaction to the choosing the new Pope. And it hit me: I have no clue what the *real* look and feel of the thing was, my perspective being so far removed from that of a Catholic bishop, or even a Catholic, in India. And it occured to me, the easiest way to get that info., and, in the process, capture pages linked to by the home page, would be:

1. Get one or two real life users willing to talk about how they use the site

2. Create videotapes of them searching the site, with Q & A along the way - the video would be the different screens that the user links to, the audio being the "can you tell me why you clicked on this link vs. that link," etc.

3. If the users would be wiling to, basically, click on each link, whether or not they's go their in real life, we'd capture a snapshot of the whole site, without agonizing about whether our crawler was getting access to everything needed to re-creat the screen, etc.

4. Call my thinking crass, but: if you've got good pics of all the pages, why worry about assembling a bunch of files and hoping you can re-assemble them later? Plus, the q&a supplementing the screen shots would provide the best possible sample of the real-life look-and-feel as experienced by the actual web site users achievable w/out much greater expense.

Apart from the usual objections:

 Obj: it would take too much time/resources (Ans.: better collect fewer sites more intelligently than tons of sites without context);

Obj: how do we pick the users? Ans.: common sense, plus, just abt. any use I'd think of is likely better than a librarian 1/2 a world away w/, for example, no understanding of Indian Catholicism

Obj: what software to use? A: I'm sure there are several options - let's just pick an Open Source one and optimize it for web site capture

Obj: But we won't have the source code, the images, all the different files that make up the site! Ans.: Think abt. it for awhile, and you may conclude, along with me, that capturing a few real sessions, inclding the captujre of all screens linked to, is worth a lot more than collecting files which may well turn out not to work together for the desired result in the future *anyway*

Would appreciate any comments/questions/examples of sites that do use this approach *GREATLY APPRECIATED*, esp. w/ my paper due tomorrow! [: (yikes - procrastination is my downfall!]

Ultimately, as a 25+ years in the field librarian, the current approach to archiving web sites - which might as well be called, "under no circumstance shall you consult a user" - simply feels like deja vue all over again :) (e.g.: libraries never used to - I hope it's changing - *ever* run online library system screens by real users for input - all we went on was our idea of how our users might react (and, on a sort of metaphysical level, what the info. resource itself "wanted" - ie, perfect description, even if that meant 5 years in the backlog).