

Wellcome Library/JISC Web archiving feasibility study

Robert Kiley & Neil Beagrie

Overview of presentation

- Wellcome & JISC are co-funding a feasibility study into Web archiving
- This presentation will focus on the need to archive the medical Web -
 - ⇒ although the JISC/Wellcome study will be broader than this.
- Examine the range of medical Web resources that need to be archived
- Discuss how the Wellcome Library and the JISC are working together to find a solution to web archiving.

Archiving the medical Web: the need

Dear Librarian,

Future generations will need to know about medicine today, not just for antiquarian reasons but often for the sake of pressing scientific, medical and epidemiological inquiry. Those researches will largely be dependent on what you do now.

Future researchers will bless us or blame us depending upon how wisely we act now.

(Roy Porter - HILJ, 2001 18:138-138)

What are we in danger of losing

- Unless we start archiving the Web we are in danger of losing
 - ⇒ General medical Web sites
 - ⇒ Online databases
 - ⇒ Discussion lists
- Library's mission is to "preserve the record of medicine past and present"
 - ⇒ We cannot ignore the digital stuff - just because it is more difficult.

Range of resources - Web sites

- Estimates in 1999 suggested that there were over 100,000 medical web sites.
- These sites represent a mix of the good, bad and the downright ugly.
- Problem - how do we decide which ones to archive? Should it be based on
 - ⇒ a specific domain (.uk?)
 - ⇒ what we can do (perhaps ignoring resources that contain multimedia file formats such as Flash and Real?)
 - ⇒ on the quality of the site?
 - ⇒ a mix of these?

- [Disclaimer](#)
- [The basis of the discoveries](#)
- [The promise](#)
- [Cancer can now be cured](#)
- [You may not have the time](#)
- [The real cause of Cancer](#)
- [Abstract/Summary](#)
- [The herbal parasite cleanse](#)
- [Zapping](#)
- [Order the book](#)
- [Order the parasite herbs](#)

The Cause

For many years we have all believed that cancer is different from other diseases. We believed that cancer behaves like a fire, in that you can't stop it once it has started. Therefore, you have to cut it out or radiate it to death or chemically destroy every cancerous cell in the body since it can never become normal again. **NOTHING COULD BE MORE WRONG!** And we have believed that cancers of different types such as leukemia or breast cancer have different causes. **wrong again!**

In this book you will see that all cancers are alike. They are all caused by a parasite. A single parasite! It is the *human intestinal fluke*. And if you kill this parasite, the cancer stops immediately. The tissue becomes normal again. In order to get cancer, you must have this parasite.

How can the human intestinal fluke cause cancer? This parasite typically lives in the intestine where it might do little harm, causing only colitis, Crohn's disease, or irritable bowel syndrome, or perhaps nothing at all. But if it invades a different organ, like the uterus or the kidneys or liver, it does a great deal of harm. **If it establishes itself in the liver, it causes cancer!** It only establishes itself in the liver in some people. These people have *isopropyl alcohol* in their bodies.

All cancer patients have both isopropyl alcohol and the intestinal fluke in their livers. The

info@drclark.net

Last changes: 01 February 2002

(c)1999-2002 by Dr. Clark Research Association and David P. Amrein

[home](#)



Range of resources - Discussion lists

- Are the next Watson and Crick communicating via some public forum, such as *behaviour-genetics*, or *gene-imprinting*?
- Equally, future historians interested in attitudes of health professionals to stories such as the BSE/CJD, or “MMR scare” would find rich pickings if the *gp-uk* list was preserved.
 - ⇒ Typically, these lists archive their postings for about 5 years...and then they vanish. Early postings about BSE/CJD, for example, are already inaccessible.

Range of resources - Online databases

- Online databases - only published electronically.
 - ⇒ May be a key source in claims of medical negligence in the future
- E-bnf - an archive of past issues would be useful for determining how fashions in drug therapy changed
- Clinicaltrials.gov - again, if archived this resource would provide a useful perspective on research priorities and how these changed over time.

Solving the problem?

- Wellcome Library and JISC are joint-funding a feasibility study to:
 - ⇒ provide an analysis of existing Web archiving arrangements and determine to what extent they address the needs of the UK research/DE/HE communities.
 - ⇒ look at a controlled sample of medical web and JISC project web sites
 - ⇒ And, if the needs are not being addressed the consultancy should suggest ways in which this could be met.

Key issues to be addressed

- Collection policy - which collection approach best addresses the needs of our communities
- Copyright issues - can we harvest “sensitive” information etc
- Costs - setting up a Web archive
- Tools - how can we harvest the Web
- Potential partners - may allow responsibilities, development risks and costs to be shared.
- Skills - developing and maintaining the archive
- Implementation plan - so we can get on and do it!

Timetable

- Contract for consultancy awarded to end of March 02
- Draft report August 2002 - to be circulated to the Advisory Group (includes BL, NLA, and hopefully LoC and NeLH)
- Completed - October 2002
- Dissemination - to the community.
 - ⇒ Web archiving must be a shared activity. No one can do it on their own.
- implementation thereafter