

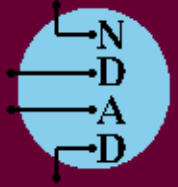
Web Archiving: an amateur's view

Kevin Ashley

ULCC/NDAD

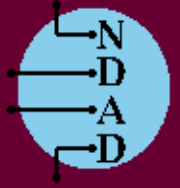
<http://ndad.ulcc.ac.uk/>

Where we are



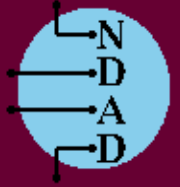
- ◆ We run our own web sites
- ◆ We host other people's web sites
- ◆ We have to archive some government sites
- ◆ We'd like to be able to archive lots of web sites
- ◆ We're not happy about any of it

Our own sites



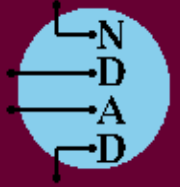
- ◆ Mixture of static, database-driven and exotic content
- ◆ There really shouldn't be a problem...
- ◆ ... but the will has to be there

Other sites we host



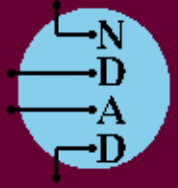
- ◆ Facility provided since 1994
- ◆ Updated via in-house CMS
- ◆ Designed for easy error recovery
- ◆ It gives potential infinite regress
- ◆ People were itchy about copyright

Government sites



- ◆ The owners do try to help...
- ◆ ... but a little knowledge is a dangerous thing
- ◆ PDF and screendumps aren't the answer, are they?
- ◆ Need to think what the purpose is, and what are the 'significant properties'

Experiments elsewhere



- ◆ Pilot with another London institution
- ◆ Central site as well as departmental
- ◆ Capture is easy (luckily)
- ◆ But what next ?
- ◆ Paradox - it should be easier from the inside, but it actually requires more work