Why archive?

- **Technical, financial reasons – 7 years**
- **Patent enquiry – 14 yrs design, 20 years for industrial etc**
- **Prior art – say 20-30 years**
- **Copyright – depends on Mickey Mouse…**
  - **Say minimum 50 years, can easily run to 150 years….**
- **Historical interest**
  - **Analysis or archaeology? - who knows**

The problem

❖ **At present we are collecting hay**

❖ **And we are looking for a needle…**

The problem

❖ **Soon we are collecting hay**

❖ **And we are looking for straw…..**

History….

- ❖ **Newton had books, diaries and manuscripts. Finite and simpl(ish) links**
- ❖ **Bohr had telex, fax, typewriters, photographs, early wp etc**
- ❖ **Berners Lee and the semantic web – how can you show blogs, pseudonyms, aliases, obfuscated email (to avoid spam etc), logs of what he looked at, who he spoke to, let alone the volumes of all multimedia data he looked at…..?**

Information overload….

- ❖ **Can use visual searching at 10s of millions of searches per second with current technology**
- ❖ **Generating image signature takes c. 1 second on i7 though….**
- ❖ **So for hundreds of millions of images?**
- ❖ **And for linkages….?**

IPR and data protection

❖ **Increased use of and need for identity and information protection**

❖ **DRM, encryption, certification, usage limits, proprietary hardware etc**

❖ **Major libraries have arrangements with publishers to ensure they have a copy of a book**

❖ **However for born digital – inundation…..**

❖ **And for protected data??**

❖ **Also Web 2 and targeted information means data is generated on the fly depending on your arrival path!**