



Definition of an "event" in the context of net archiving

The netarchive.dk has three strategies for collecting internet materials, covered by the Danish Legal Deposit Law:

1. Snapshot harvesting (all relevant domains) - quarterly
2. Selective harvesting (appr. 80 domains) – frequency from one hour to 1 month
3. Event harvesting

As the purpose of the harvesting activities is to assure a complete collection of Danish internet materials in the net archive, the above-mentioned strategies are strategies for harvesting not for creating special collections. Consequently, the materials gathered through event harvesting will not become separate collections within the archive. The archive will, however, keep a list of events harvested as part of collection documentation.

The following guidelines deal with Event harvesting.

Definition

An event is one which

- *Creates debate among the population and is expected to be of importance to Danish history or have an impact on the development of Danish society*
- *Causes the appearance of new websites devoted to the event*
- *Is dealt with extensively on existing websites*

Harvesting an event can be a brief extraordinary harvest (that is outside the scheduled harvests) during a short period, or it can be a longer-lasting harvest with a frequency that is in- or decreased according to the activities on the websites. An event may cover a few or many websites.

Certain events, such as elections, are predictable. Most events are not and – in fact – do not appear as an event until it has started, that is, when new websites devoted to the event have appeared, when the event has created debate, and when the event is dealt with on existing websites. In order to catch such an event as early as possible, it may be necessary to begin harvesting websites that appear to be part of a developing event. If this turns out not to be the case, the harvest will be stopped. The same happens if the activity of the websites is such that the content will be caught by the other two harvesting strategies.

In order to assure that unpredictable events get properly collected it should be considered to do a separate snapshot harvest once a month exclusively on new domains that have been fed into the system since the previous complete snapshot, including new domains from the DK-Hostmaster list.

At present, the resources of the netarchive.dk allow for only two major events a year. Minor events can be captured as part of the daily work with finding and harvesting all “Danish material published in electronic communication networks” as the law requires us to do.

Predictable events

Are events that are planned and whose beginning and end is known in advance:

Type	Begins	Ends	Harvesting
National elections	When the websites of the political parties contain campaign materials	When other harvest strategies are sufficient for getting the materials	Always
Local and regional elections	When the websites of the political parties contain campaign materials	When other harvest strategies are sufficient for getting the materials	Always
Other events (e.g. royal events, sports events ...)	When the event causes activity on the Internet	When other harvest strategies are sufficient for getting the materials	Not all events. Decision is taken at each event

Unpredictable events

Are natural disasters, outbreak of war, terrorist attack, political crises (at home or abroad) and certain personal events (scandals involving politicians e.g.)