



A tree full of leaves

Description Logic and data documentation

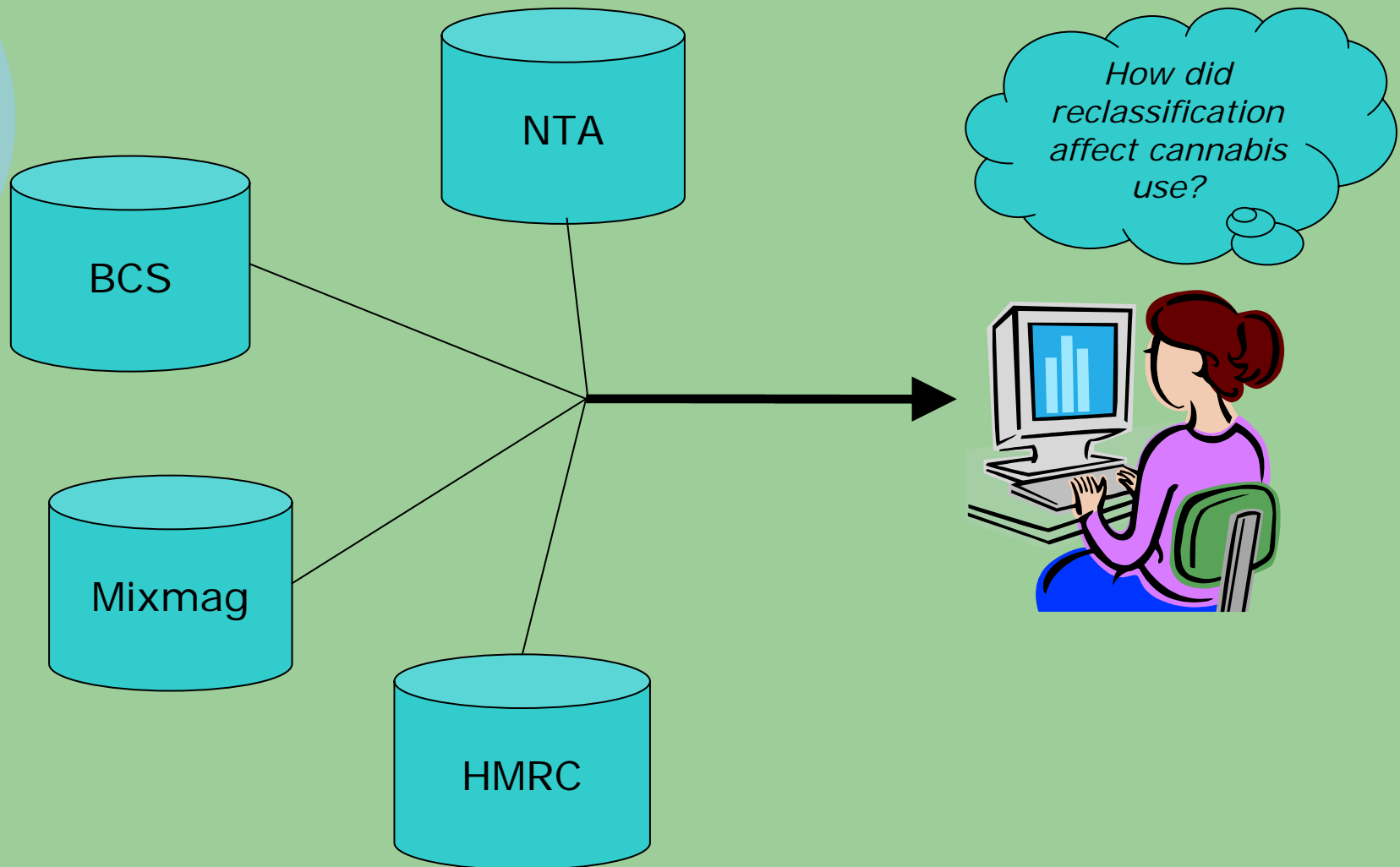
Phil Edwards, Judith Aldridge and Karen Clarke

**Second International Conference on e-Social Science
June 2006**

How the Grids stack up

- the Computation Grid
 - ...supports
- the Data Grid
 - ...underpins
- the Semantic Grid
 - ...makes possible
- the ***Knowledge Grid***

Your Grid on drugs



Problems!

- Clear numbers, fuzzy data
 - 10% of *exactly who* said *exactly what*?
 - Is my 10% the same as your 10%?
 - (Clue: Probably not)
- Clear data, cloudy concepts
 - *Imprecise*
 - *Contested*
 - *Mutable*

Prospects: cloudy

○ Imprecise

- Definitions often vague
- Sometimes by design ('anti-social behaviour')

○ Contested

- Multiple definitions of key terms
- Different users need different definitions ('drugs')

○ Mutable

- Concepts change as society changes
- Problem of time-lag ('Head of household')

Cloudbusting

- *“The ones who drove within half an hour of smoking cannabis **must** have been under the influence...”*
- *“The British Crime Survey asks about sniffing glue and Mixmag asks about alcohol and tobacco, but those aren’t really **drugs**...”*
- *“Class B drug use dropped in 2005, but that’s only because of the reclassification so it didn’t **really**...”*
- You don’t want to do this!

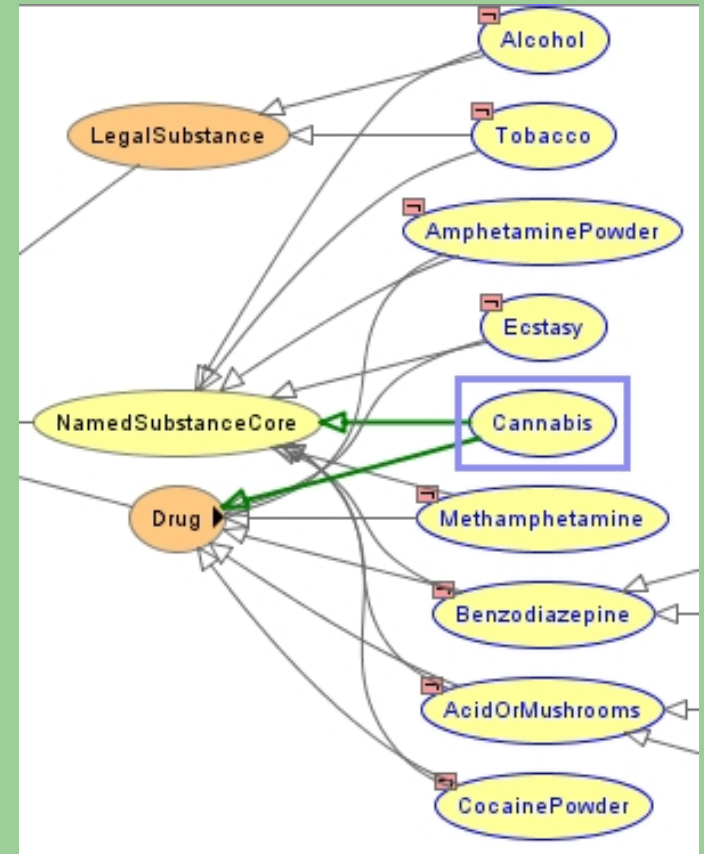
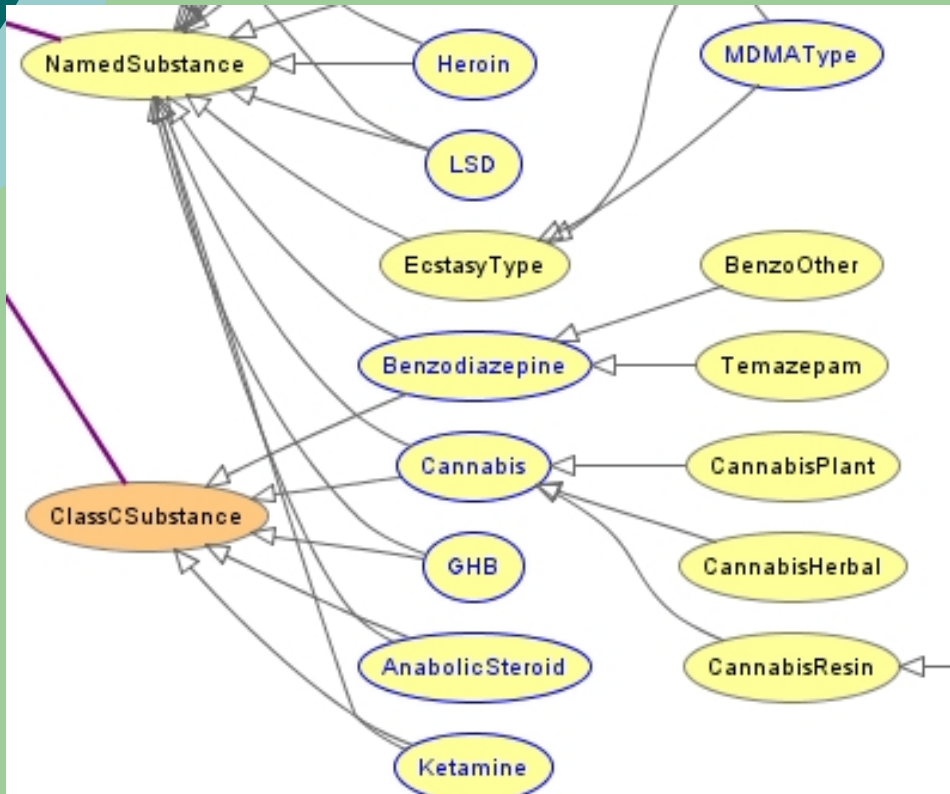
Cloudbuilding (1): a bare tree

- SOCIAL ISSUES
 - PROBLEMS
 - ABUSE
 - ALCOHOL ABUSE
 - CHILD ABUSE
 - DRUG ABUSE
 - DRUG-RELATED CRIME
 - DRUG-RELATED HARM
 - DRUG-RELATED DEATH
 - ...
 - ELDER ABUSE
 - SPOUSAL ABUSE
 - ...

Cloudbuilding (2): a pile of leaves

- Seizures by HM Customs and Excise, calendar year 2001
 - Class B drugs seized
 - Amphetamines (kg)
 - Cannabis (herbal) (kg)
 - Cannabis plants (number of plants)
 - Cannabis (resin) (kg)
- Cannabis
 - Ever used
 - Used in the last year
 - Age when first used
 - Used in last month (0-30 days)
- Has the reclassification of cannabis
 - Changed the amount you use (use more/use less/no change)
 - Changed the amount you buy at one time (larger/smaller/no change)
 - Changed where you smoke cannabis (more public/less public/no change)

Cloudbuilding (3): two trees



Cloud logic: the uses of OWL-DL

- Conceptual hierarchy
 - Cannabis is a type of Drug
- Horizontal links (properties)
 - Cannabis Use [is use of] Cannabis
- Multiple hierarchies
 - Cannabis is a type of Class C Substance
- Multiple ontologies
 - Cannabis [corresponds to] Dope

Summary

- Social science data is built on inherently fuzzy concepts
 - Imprecise, contested, mutable
- Imposing a unifying structure is not satisfactory
 - Nor is doing without structure
- Wanted: logical documentation of illogical concepts
 - Enter OWL

Any more questions?

