PERSON objects in the RIPE Database
The Issue

- There are 2+ million referenced PERSON objects, with 8+ million total objects
- 23% of RIPE Database is personal data
- There are also about 104k ROLE objects, some may contain personal data
Justification for storing personal data

- T&C define one purpose of Database as facilitating contact for operational issues
- Address policies require contact details for resource holders and some end users of public address space
- Address policies acknowledge privacy concerns
Who are these people?

- About 1.5m have a single reference from INET(6)NUM objects
  - these may be tech contacts
  - or could be customers with assigned address space

- Half a million have same name/phone, probably duplicates
There has been a steady creation rate for about 15 years.
Email addresses

- There are about 100k unique email addresses in various notification attributes
- These may also be a mix of personal and corporate data
From the analysis I’ve done so far it is not possible to identify how much of this personal data can be justified.

Maybe 90% is justified, or perhaps only 20%.

With general EU privacy laws and GDPR in particular, does this give (someone?) a legal problem?

Who is accountable for this personal data being in the database?

If the justification for this data being in the database is questioned, who will take the responsibility for answering that question?
What is a contact?

- Original Database design intended PERSON objects to be mainly referenced from ROLE objects, all other objects to reference the ROLE objects
- This wasn't enforced by the software
- Historically most users ignored ROLE objects in favour of PERSON objects
- MNTNER objects cannot be created without a PERSON object
- No where is it documented that PERSON objects should contain business data
What is a contact?

- "abuse-c" was specifically designed to enforce the use of ROLE objects
- ROLE objects no longer need to reference PERSON objects
- Why can't most contacts be roles?
- Should the software (finally) enforce that?
- What information is needed from a contact?
- Can we justify storing personal data for a contact
  - if no personal data is needed?
How are these 2m PERSON objects linked to organisations?

No apologies for the dryness of these numbers

They need to be seen as they are quite surprising
Organisations with 100k+ PERSON objects

- 3 Telekom companies totaling 856k
- Top organisation has 613k PERSON objects
  - 5 x what next 2 organisations have
  - 10 x what most other Telekom organisations have
- Why?
Organisations with 5k to 50k PERSON objects

- After the top 3 there is a drop to around 50k

- 43 organisations totalling 704k PERSON objects
Who are these organisations?

- Summarising the organisations with 5k+:
  - 46 organisations
    - 35 telekom
    - 7 service providers
    - 4 others
    - Totalling 1.56m PERSON objects
  - 110 organisations between 1k and 5k, totalling 230k PERSON objects
  - Everyone else totalling 233k PERSON objects
What next?

- We still don't know if these people are contacts
- We still don't know if these PERSON objects contain personal or business data
- Contact details not changed in 20 years (address, fax-no?)
Questions?