Background

• In the “early days”, whois was used by network operators (ISPs, web-hosters, infrastructure providers) to identify domain holders – in case of email problems, DOS attacks, etc.

• The data collected in whois was rudimentary – registrant, administrative, technical and billing contacts
Background (continued)

- Data elements historically collected in whois:
  - **Registrant** (though no telephone number or email address)
  - **Administrative Contact** (usually web-hoster or ISP)
  - **Technical Contact** (usually ISP)
  - **Billing Contact** (sometimes registrant, sometimes web-hoster, depended on retail model used to register domain)
In 1999, ICANN required registrars to provide both web-based and Port 43 access to whois
   - Not all registrars have an operational whois – the requirement is detailed in the RAA
   - Most of top 10 gTLD registrars limit queries per IP address – maintain that this protects customers & system overload

Bulk Whois access also required
   - Registrars must make available their list of registrants (who do not “opt-out”) for a fee “not to exceed US $10,000”
   - Of the top 10 TLD registrars, only a fraction offer bulk whois with contracts that conform to the ICANN contractual requirements (most include provisions that are not “in the spirit” of the contracts)
• In 2001, ICANN required registries to provide both web-based and Port 43 whois
  – Registries with “thin” data – VCNR and PIR (at start) only show minimal data
    • PIR is converting from “thin” to “thick” throughout 2003
  – Registries with “thick” data show full registrant record – same as registrars
  – GNR (.name) was able to change ICANN contracts to offer minimal data (to protect their registrants – individuals)
The Data

- Registrars manage the registrant (customer) relationship
  - Collect domain name record details at registration
  - Manage the updating of the data
  - Furnish it to “thick” registries
  - Provide access via Whois

- Registries simply provide access via Whois
  - if .com/.net, then “thin” display
  - if .org (in transition), .info, .biz, new sponsored gTLDs, then “thick” display
Issues

• Many interests now want to keep whois status quo:
  – IP community – to protect marks
  – Law Enforcement – to help investigations
  – Network Operators – in case of attacks

• Many interests want to restrict whois data:
  – Privacy Advocates – protect PII
  – European Union - protect PII and data export
  – Registrars – retain customers/market share
  – Registries – PII concerns, not to violate EU restrictions
Some Misconceptions

• Spam is only based on domain registrations
  – If you post to public mailing lists, “bots” harvest your email address
  – If you purchase anything on-line and the retailer does not have a policy against spam, your email address will be sold

• Bulk Whois is the major source of email addresses for spammers
  – Since so few registrars offer Bulk Whois files, spammers use other methods to harvest email addresses
Some Realities

• The registry zone files are used to identify changes to each TLD
  – The best source for spammers & others
• Via contractual requirements, registrars and registries must make data freely available to anyone
  – Enables data mining/harvesting of Whois records
For Further Discussion

- Allow zone files to be restricted to legitimate uses
  - Necessary for IP, ISPs, network operators, etc.
  - Today, registries must comply with any request – cannot decline a request for free zone file access

- Limit whois access for legitimate purposes
  - Most queries can be addressed by “domain available” or “domain not available” response
  - Requires education to community
  - Requires interested parties to define “legitimate purposes”

- Engage community in refining the purposes for whois in this changed world and environment
  - Much has changed since “the early days”, registrar competition in 1999, registry competition in 2000
Whois Workshop

Introduction to Registry/Registrar Issues

Contact: bbeckwith@pir.org

June 23, 2003

Serving the Public Interest