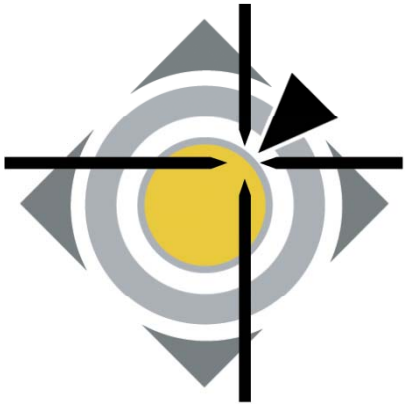
IMF 2008

Workshop-Day

Mannheim, Germany – 25.08.2008



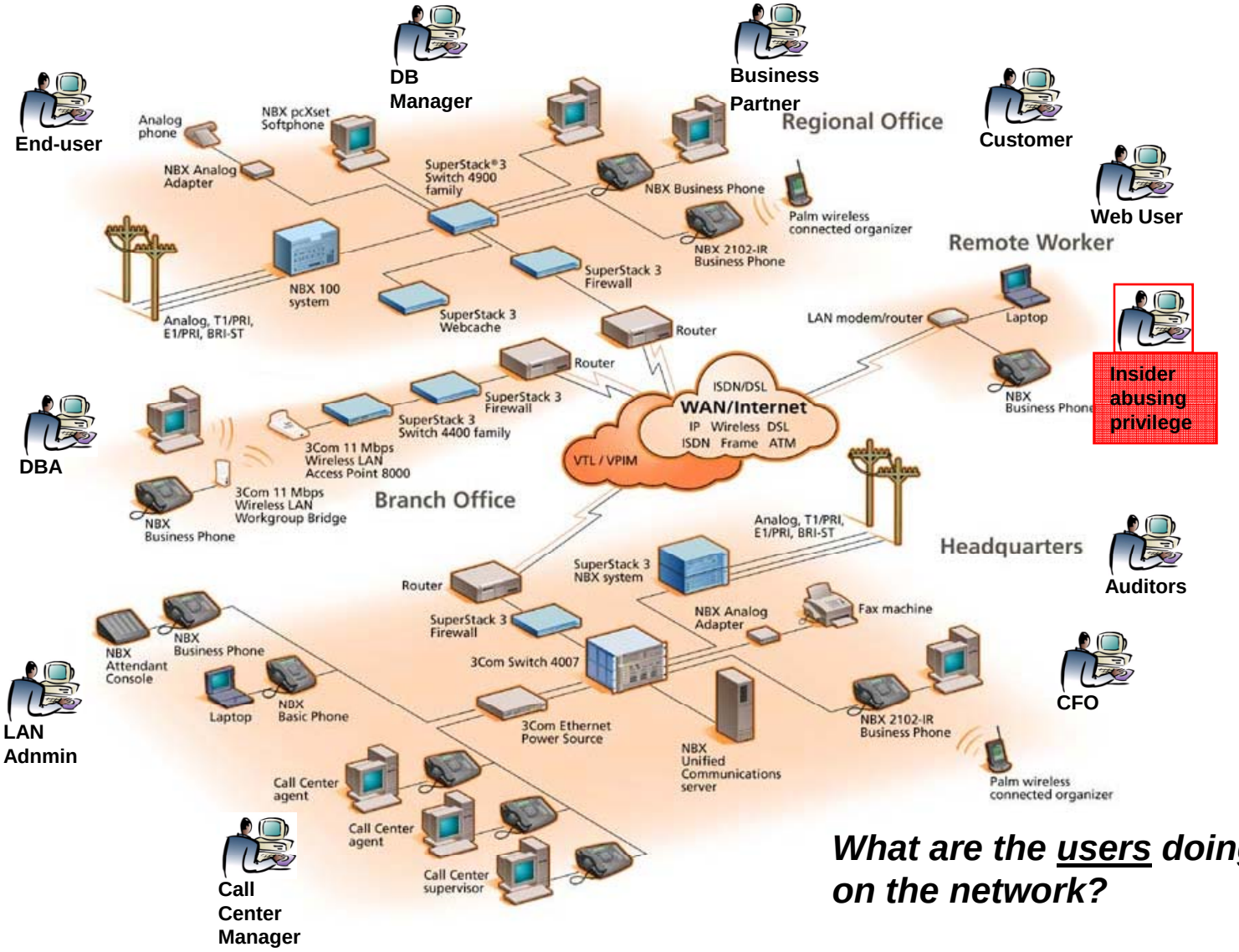
Best Practices (Internet) Auditing

Andreas Rohr

Agenda

- **Audit in General**
  - Why Auditing
  - Audit Sources
  - Dealing with the amount of audit data
- **Internet-Audit (case study)**
  - Private internet usage
  - Statistics
  - Calculate internet usage time consumption

Why Auditing?



What are the <u>users</u> doing on the network?

Why Auditing?

- **Motivations to audit can be driven by:**
  - Requirements, stated by an (international) law: **Compliance** due to legal constraints (BASEL II, SOX, …)
  - **Policies**, that can not be technically enforced
  - Interest in business/process related **indicators** (ex.: When do users/costumers do what to which extend?)
    - Planning IT-infrastructure resources (operations management)
    - Optimized customer processes (resource allocation)
    - Recognize possible bottlenecks
  - **IT-Security** interests

What is Auditing?

- **Auditing generally deals with the question (two different definitions):**

1. **Who has done What, When, Why and Where?**
2. **Who did What type of action on What?**

   When did he do it and Where, From Where and Where To?

Types of Auditing

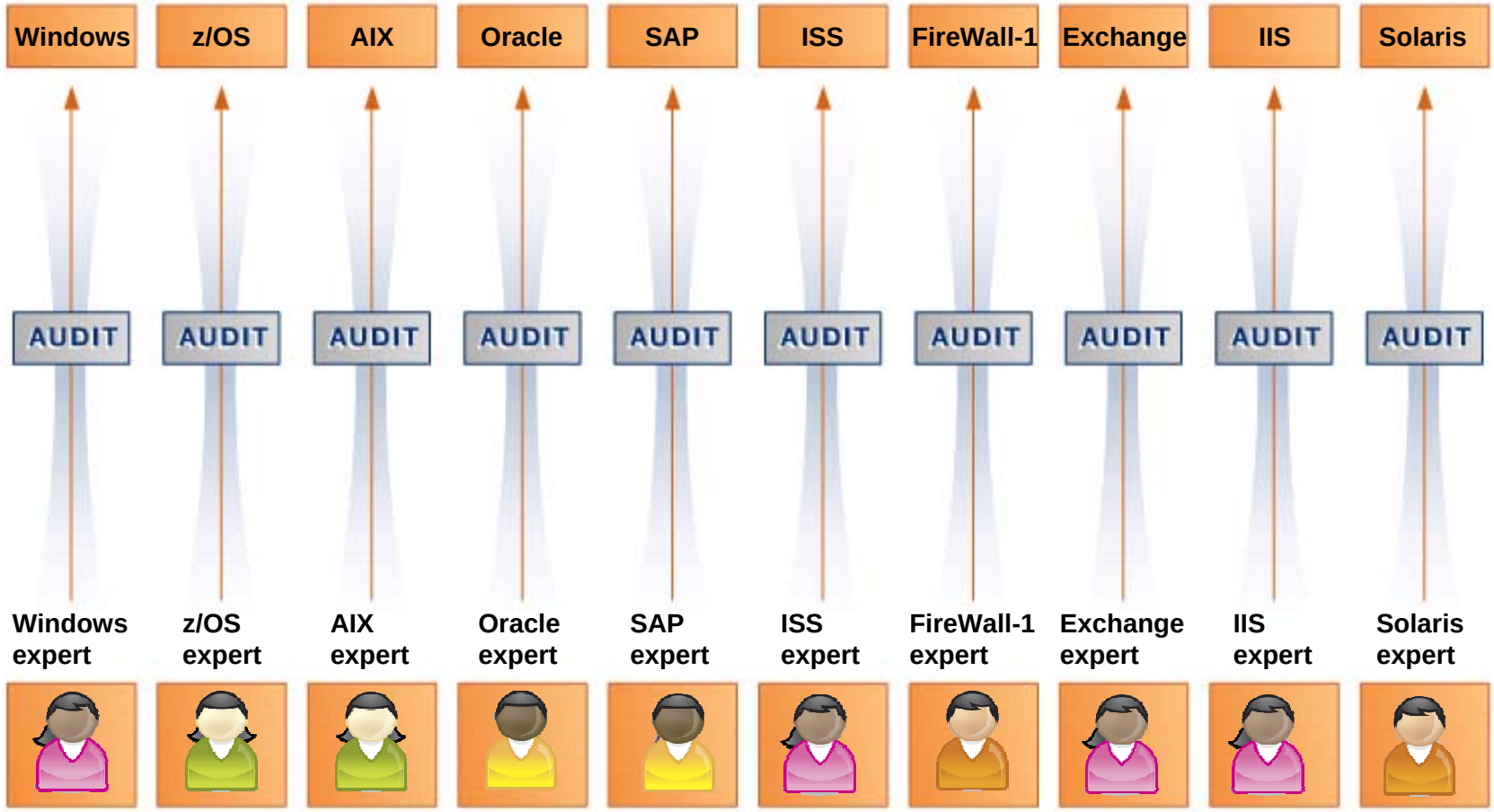
- **Incidident/Case (driven) audits**
- **Regularly performed audits**
  - Asynchronous: post mortem to an event
  - Synchronous: (near) real-time after an audit-entry is generated
    - Trigger based audits

Audit Sources

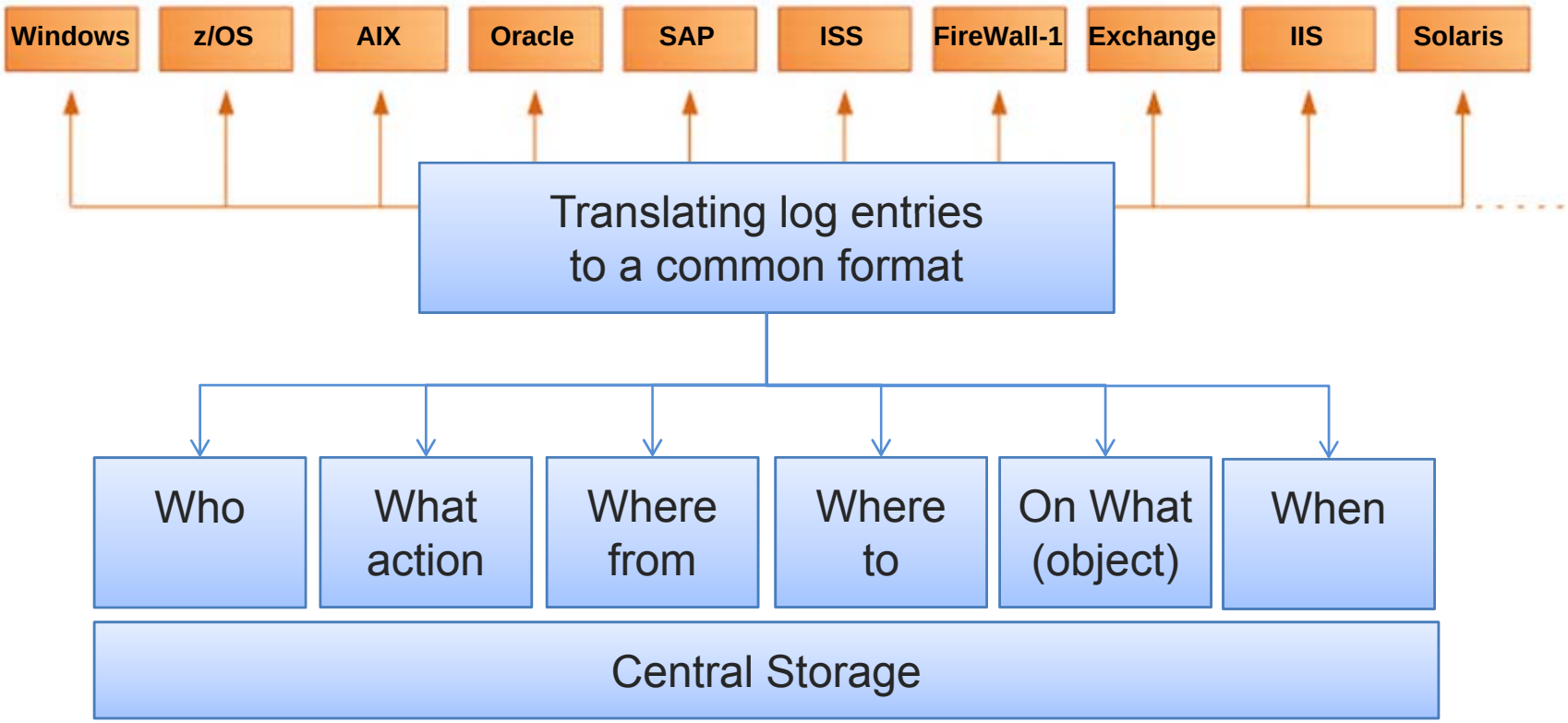
- **Logs created by various systems and applications**
  - Door entry systems (ID-cards, biometric authentication)
  - Operating System Logging (syslog, system event logging, RACF, etc.)
  - Database logging (transaction log, access log, …)
  - Network devices (Router logs)
  - Security devices (Firewalls, IDS)
  - Application specific logging (access-logs, VMWare Virtual Center Activity Log, custom logs defined by a administrator)

- **Cameras**
  - On-the-fly object categorization of a video-stream (surveillance camera) → Meta-Data for any object that moves within the camera scope

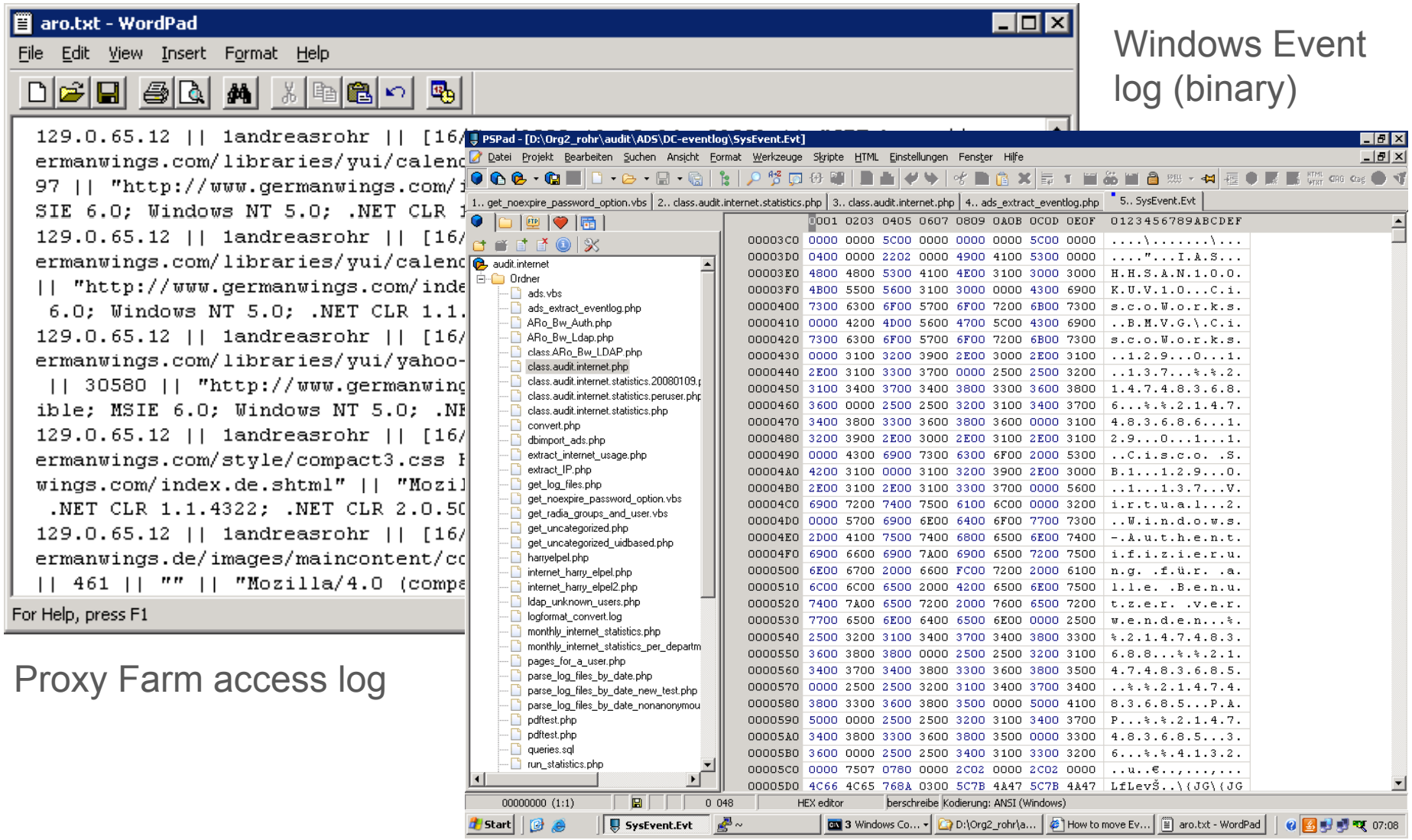
Audit Sources



Audit Sources – Translation



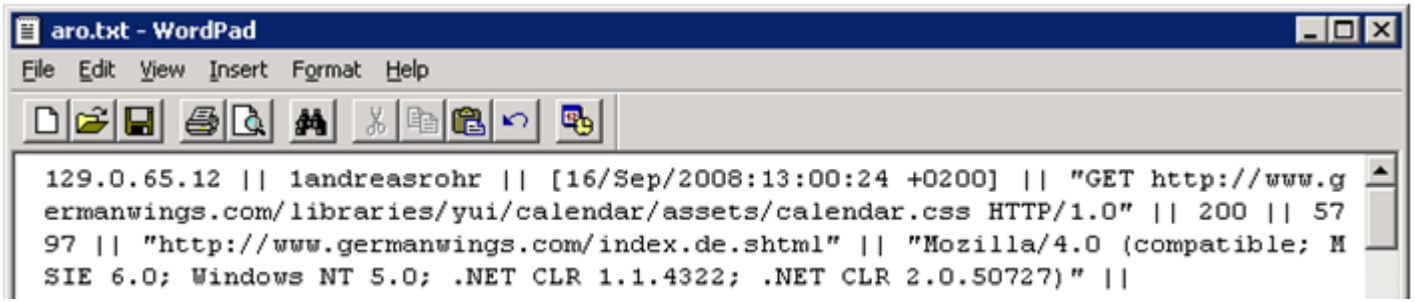
Audit Sources – Translation

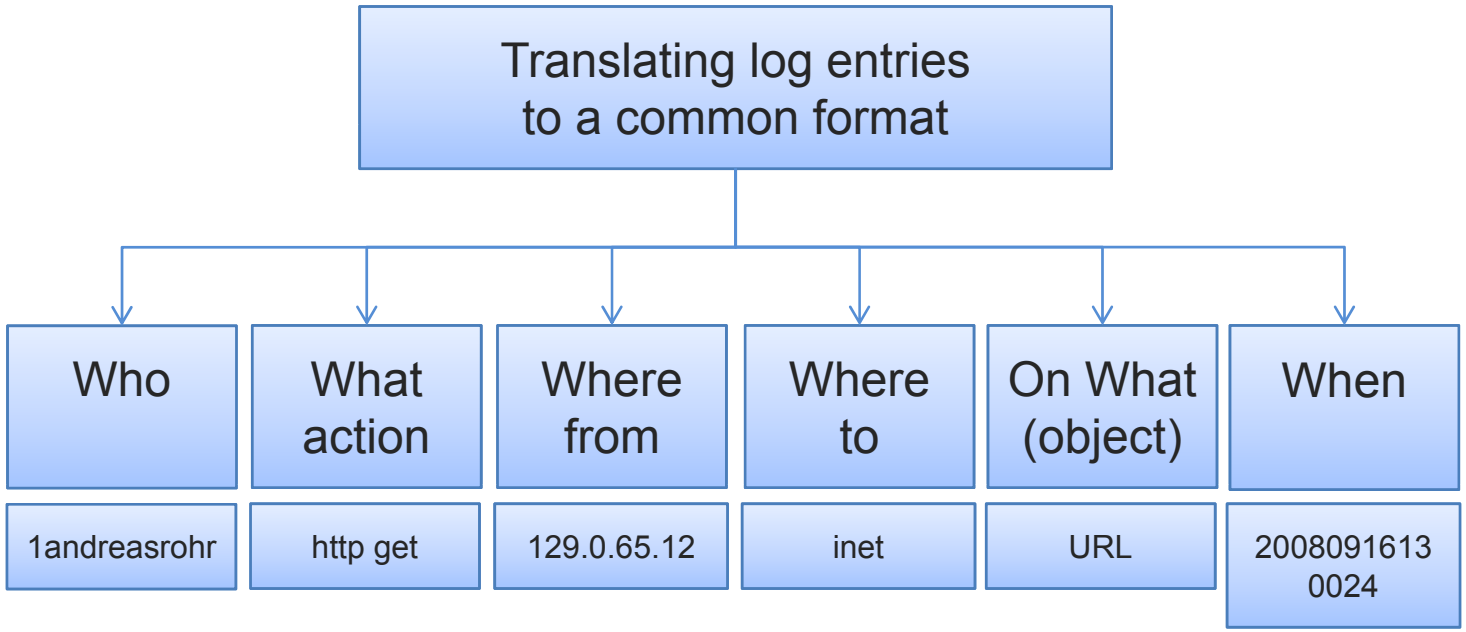


Windows Event log (binary)

Proxy Farm access log

Audit Sources – Translation (ex.)





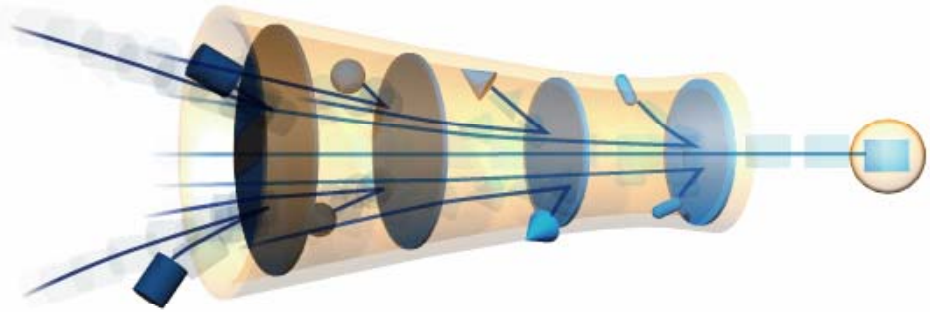
Aggregation / Correlation

- **Hundreds of thousands of audit/log entries**
- **Different audit sources**
- **User behavior (detection of normal vs. abnormal behavior measured to a certain baseline)**

⇨ **Aggregation / Correlation is needed!**

Statistical Analysis
- Detecting non typical behavior

Rules-based Correlation
- Detecting non typical behavior
- Detecting misuse
- Enforcing security policies



Accountability of Logs/Audit Sources

- **Spoofing of basically any attribute that is used by an auditor:**
  - ensuring that a log entry is not tricked/spoofed by the trigger (triggering user)
    - ID spoofing (who)
    - IP-Address spoofing (from where)
    - etc…
  - Timestamp service for logs / log entries (PGP, company PKI)
- **Cross-correlation of different sources**
  - Door entry systems / time registration systems
  - DNS log
  - OS login (LDAP log)

Auditing Acceptability

Required Skills for an Auditor

- Coheasive view on what and how systems / applications are tracking/logging
- At least one programming (scripting) language for automation (ksh, perl, php, vbscript, etc.)
- Forensic knowledge (capabilities) to cross-check results (if they are plausible to the rest of the data)

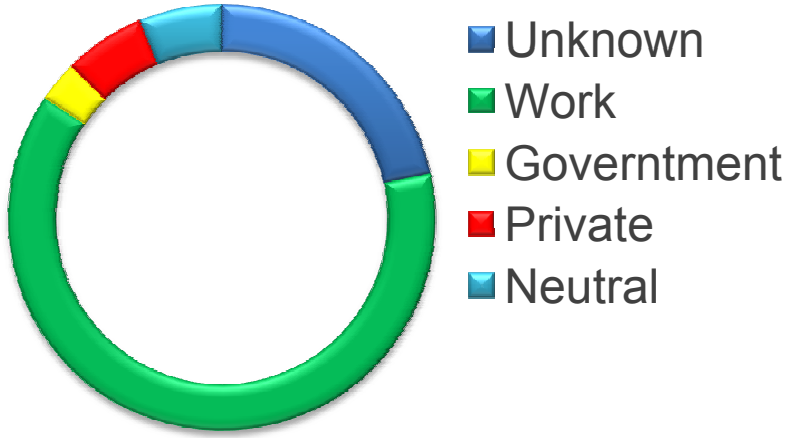
Internet-Audit

- **Case study of the FMoD Germany**

- **Assumptions:**
  - Internet usage is not permitted for private purposes
  - Authentication for internet access (proxy farm authenticates with LDAP directory)
  - MOU between MoD secretary of state and the employee committee
    - Audit on misuse (e.g. private usage) on a per user basis is permitted once a month
    - Anonymised /cummulated audit (statistics) on the department level and above; results in a executive summary report at the end of each month

Internet-Audit: Private Usage Def.

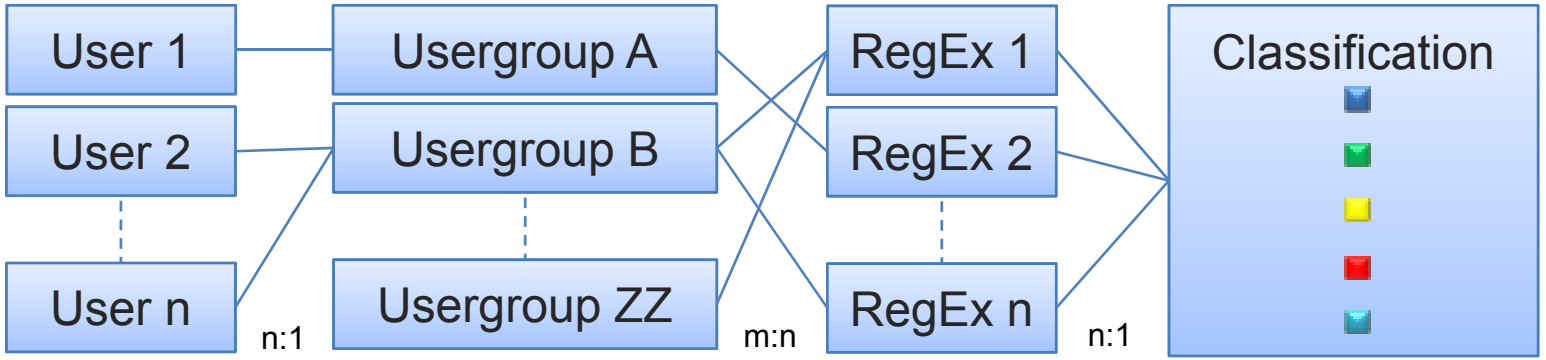
- **Definition of private usage within the FMoD:**
  - Classification based on manually rating of sites:
    - Unknown (unclassified yet)
    - Work relevant
    - Special subnet (governmental intranet)
    - Private
    - Not classifiable (neutral: not private, but also not work relevant)
  - Using categories of webwasher URL-filter database
  - Threshold for mitigation between man power to classify web sites and precision: 20% Unknown



Fictitious distribution!

Internet-Audit: Private Usage Def.

- **Definition of private usage within the FMoD (2):**
  - Sites are classified per second, third, … level domain as RegEx:
    - **.*.?google\.[de|com|fr] : Work**
    - **chat\.google\.[de|com|fr] : Private**
    - **.*google-analytics\.com : Neutral**
    - **.*ivbb\.bund\.de : Government**
  - All RegEx rules are related to user groups
  - Default: all RegEx applies to each user group
  - All RegEx rules are ordered; first match decides → like FW rules



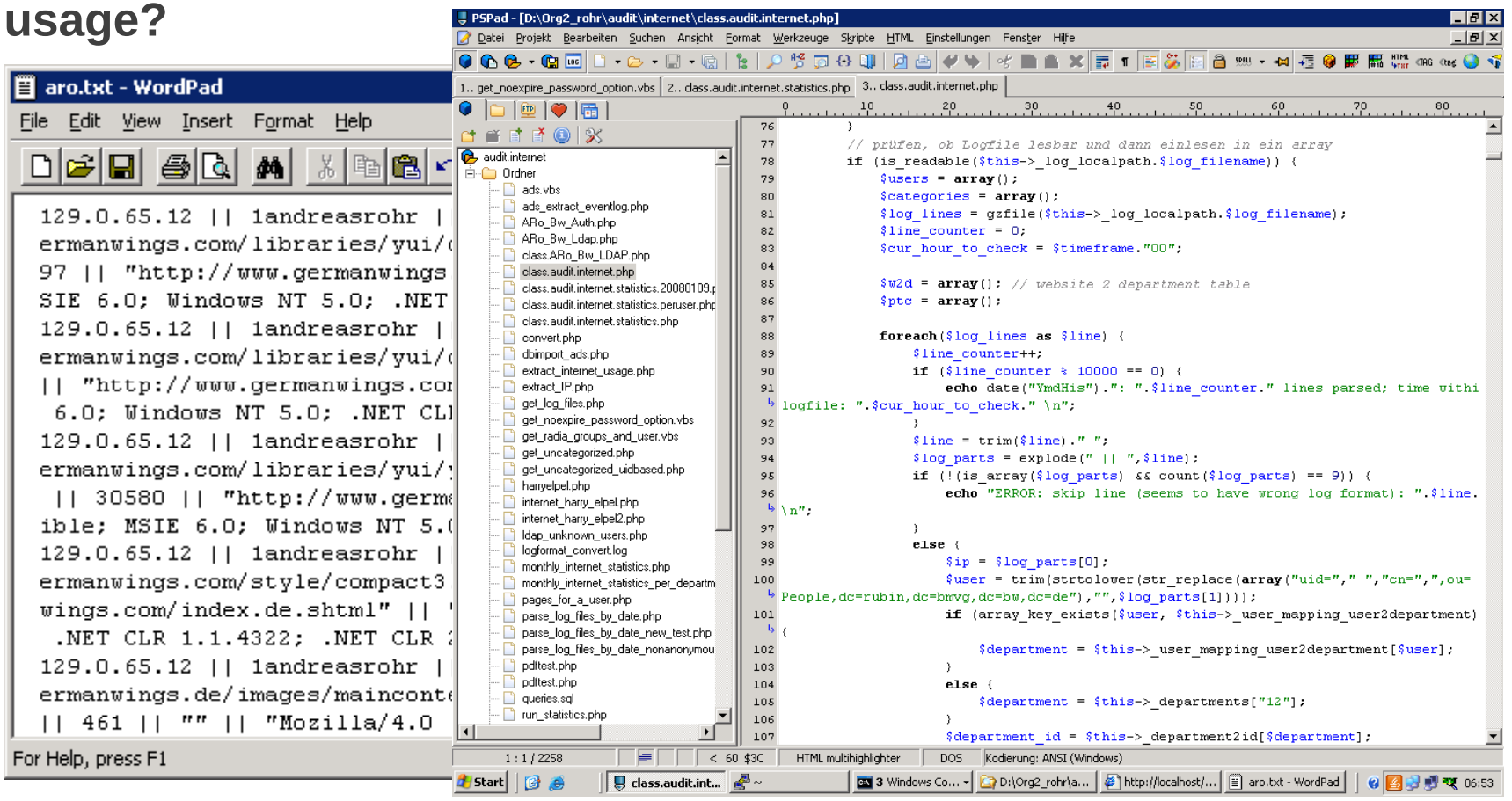
Internet-Audit: Measuring time usage

- **Discussion with our legal advisors for court proof (plausible) results**
  - For how long has a defendant used the internet for private purposes?

- **Problem:**
  - http is a stateless protocol
  - Within the access log there one line per http request
  - How long does a user sits actively in front of a loaded webpage?
  - Are there any disruptions (ex. phone calls, visitors)?

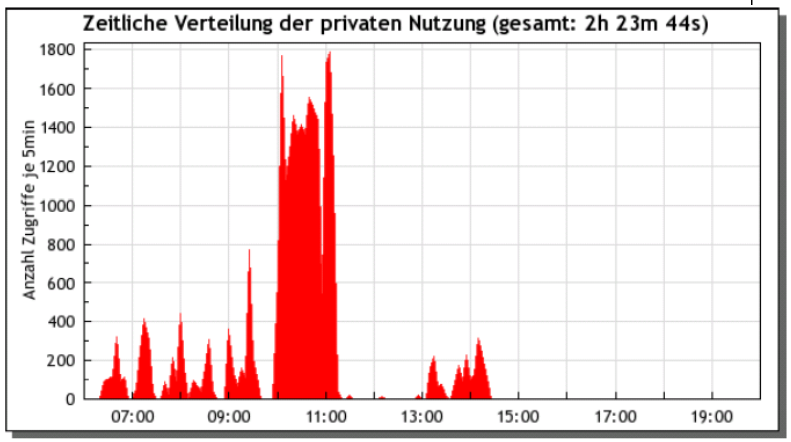
Internet-Audit: stateless requests

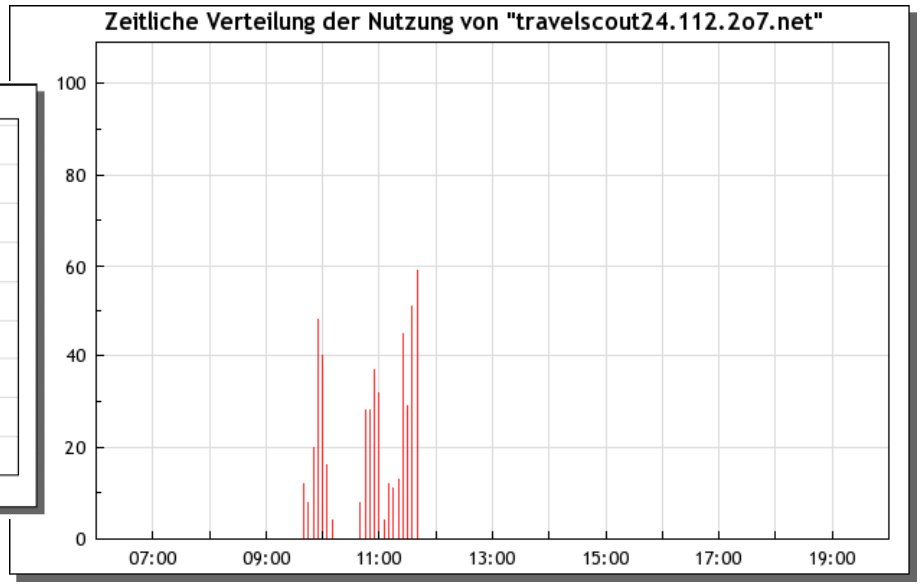
- **How to translate stateless requests (as stated within the proxy farm‘s access log) to a stateful timeframe in sense of real usage?**

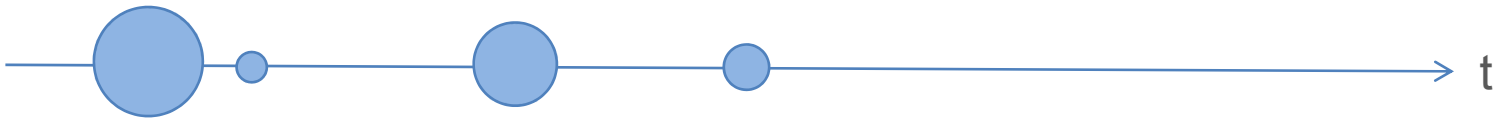


Internet-Audit: stateless requests

- **How to translate stateless requests (as stated within the proxy farm‘s access log) to a stateful timeframe in sense of real usage?**
  - now we have a set of translated access log entries that belong to a certain user



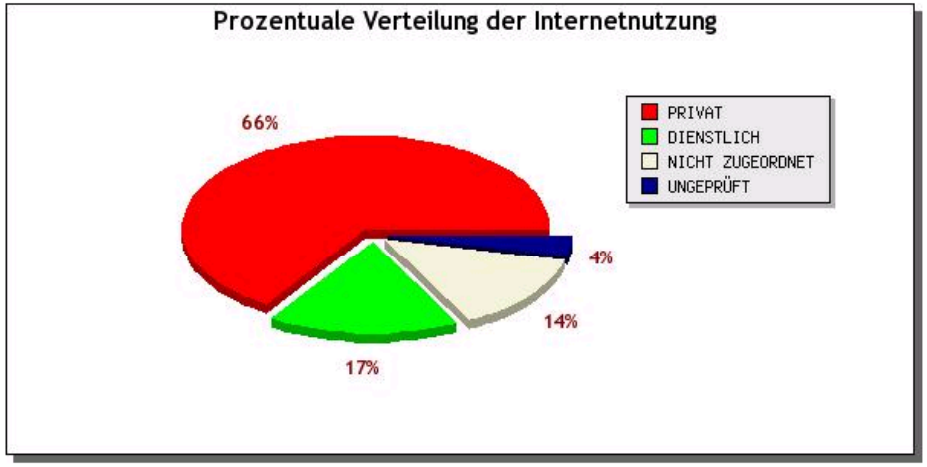


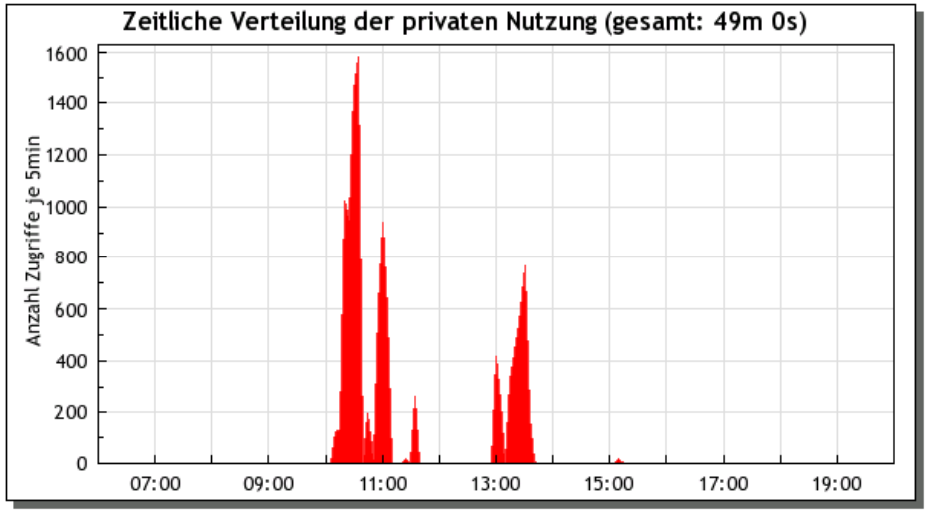


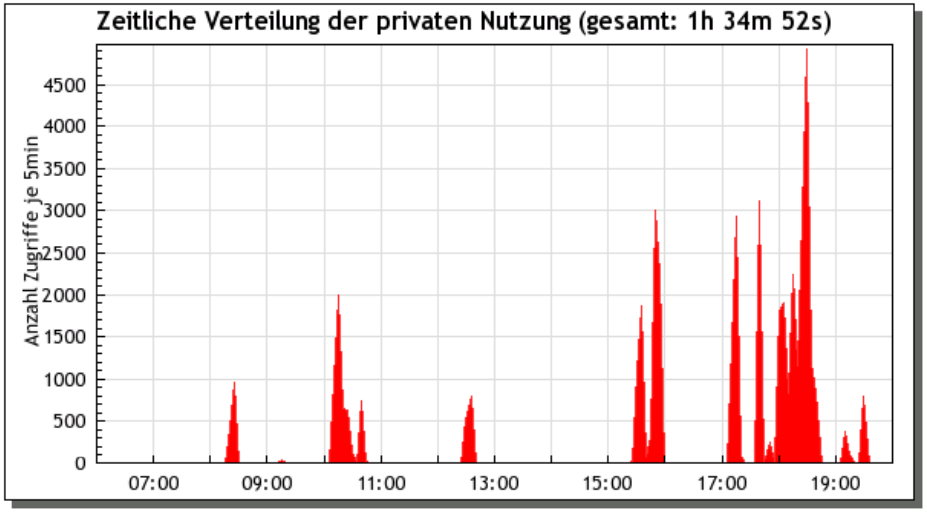
Internet-Audit: Determine time usage

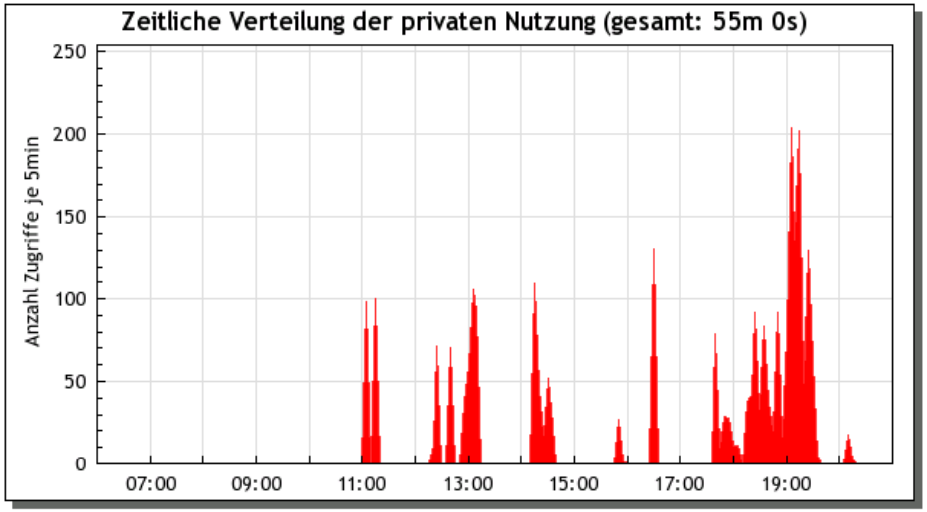
- **For calculating the time usage we developed a heuristic:**
  - Time usage consists of 1 or more intervals
  - Each interval starts with the first request within this interval
  - An interval spans over requests, where consecutive requests must be within a defined timeframe **TF** (of 1 minute)
  - If this timeframe **TF** is exceeded by two consecutive request the interval is closed and a new interval starts with the least of the two requests
  - To avoid marginal usage to be counted we only “charge” intervals with a length greater than 1 Minute
  - The calculation heuristic ignores the usage part after the last request of an interval

Internet-Audit: Determine time usage



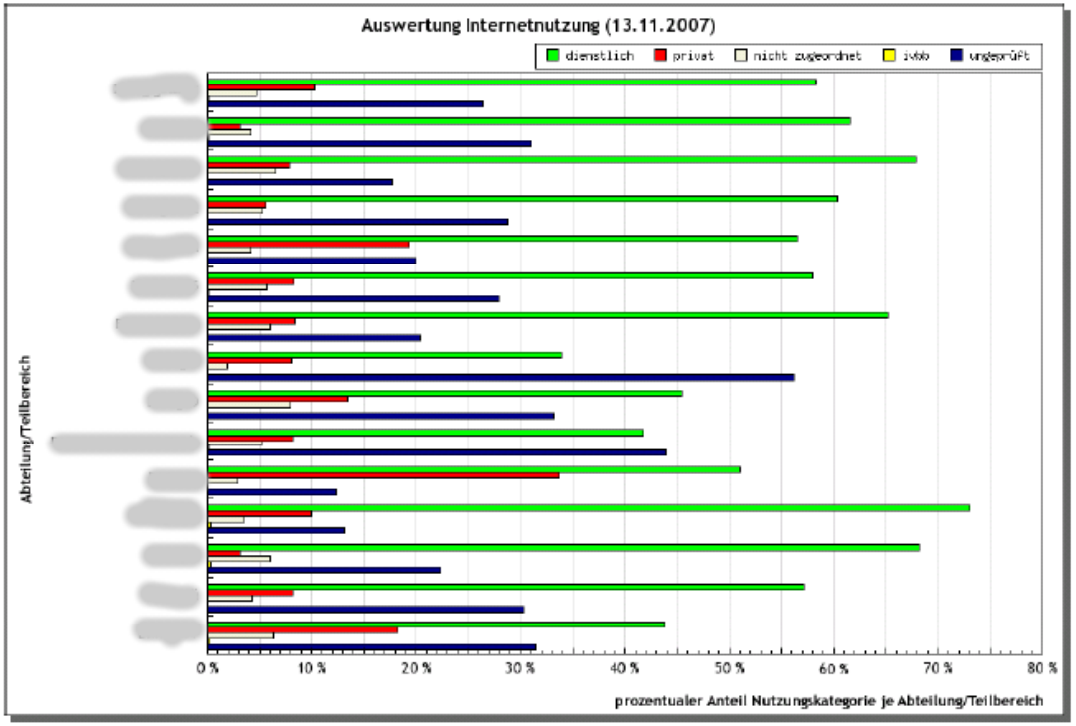






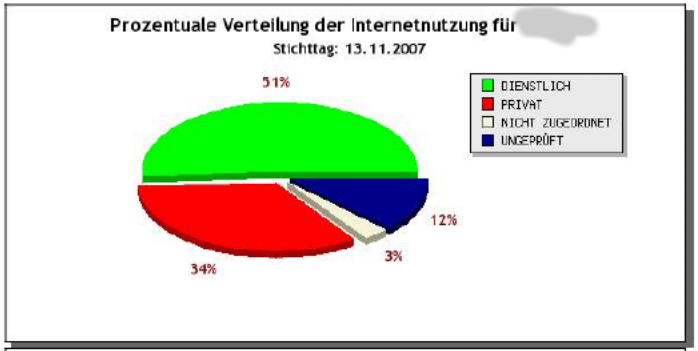
Internet-Audit: Anonymous Audit

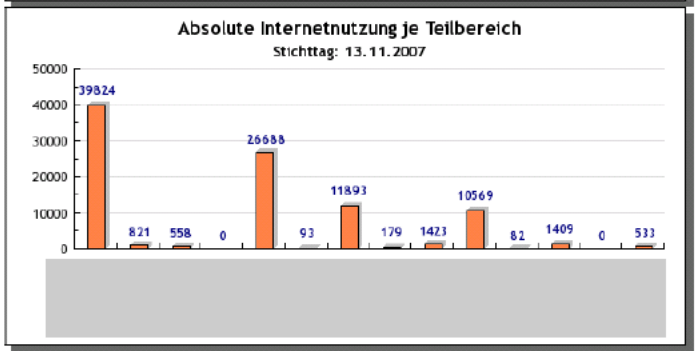
- **We audit the internet usage once a day**
  - We cumulate the distribution of classes for each department
  - According to the organisational structure we build the branch and division as well as the overall statistic

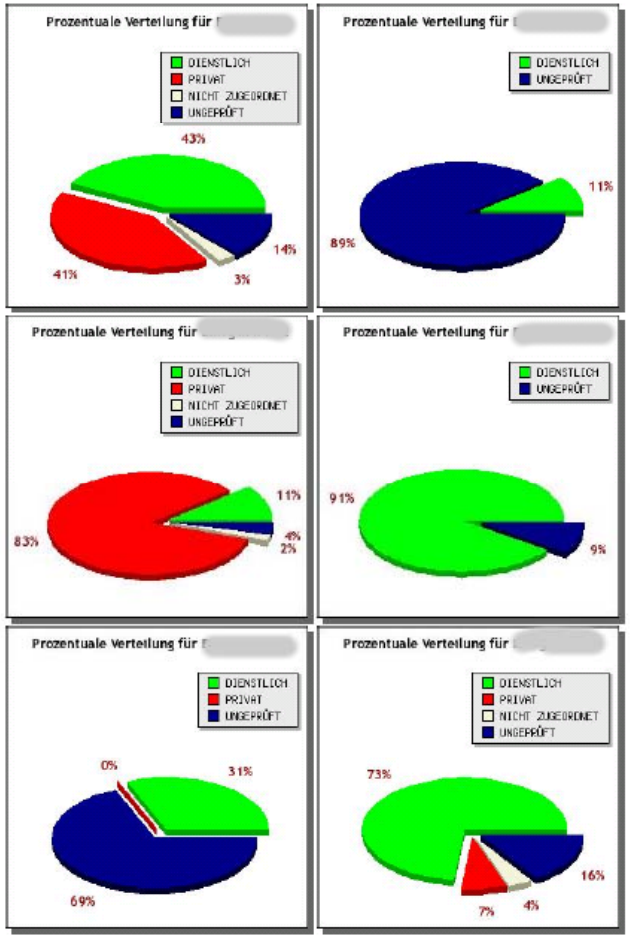


Internet-Audit: Anonymous Audit

- **Division drill down to the departments**

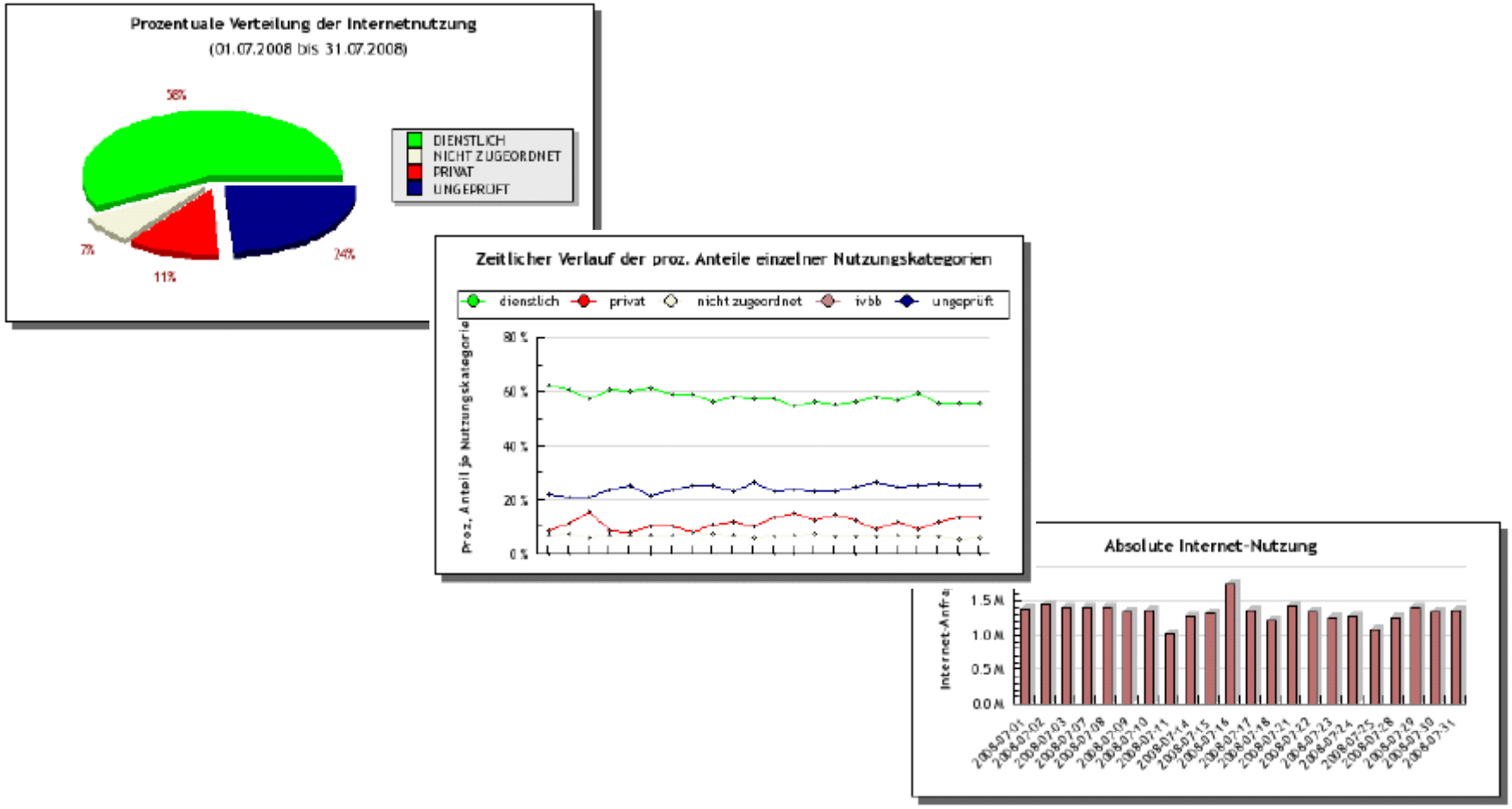






Internet-Audit: Anonymous Audit

- **Executive summery (once a month)**



Thanks for attention/participation

Questions ???

Contact:

Andreas Rohr
Federal Ministry of Defense
Fontainengraben 150
53125 Bonn

+49 228 12 9277
andreasrohr (at) bmvg.bund.de
rohr (at) hsu-hh.de