DATANOMIQ
Applied Data Science

DATA SECURITY CHECKLIST
for Data Engineers / Data Scientists / Data Analysts

»»» PASSWORD MANAGEMENT
Passwords should stay unique and secret, never share or reuse them!

- » Choose strong passwords with at least 14 signs (containing letters, numbers & special characters).
- » Do not re-use passwords across several systems.
- » For web/cloud systems, change your password at least once all 12 months.
- » Consider using Multi-Factor Authentication!
- » Attention to valid SSL Encryption when you use your password to access external systems!
- » Do not store your password in online or offline notes. Consider using a password manager tool.
- » If you need to share passwords, use split passwords (e. g. TAN-lists or password tables) or use password manager for teams (e.g. 1Password).

»»» BE AWARE OF HACKING BY SOCIAL ENGINEERING
Always stay skeptical if someone asks you for information or access!

- » Do not trust any strange person who claims to be part of any team. Ask identifying questions before giving trust.
- » Never share passwords to people without cleared responsibility and purpose of usage.
- » Be aware of phishing e.g. via e-mail! Have a closer look on the sender of an e-mail.
- » Do not trust open WLAN and be aware of Wi-Fi phishing (evil twins).

»»» ENCRYPTION OF PERSISTENTLY STORED DATA
Personal and enterprise data should be encrypted per default!

- » Have at least one partition of your hard disk encrypted with Bitlocker (Windows), FileVault (macOS) or VeraCrypt.
- » Store critical or private data on encrypted partitions and/or in encrypted Linux systems.
- » Use Virtual Mashines (VMs) as extra Encryption Level.
- » For very critical data create own encrypted partitions and use steganography (e.g. hide partition or use OpenStego)!
- » Use hardware encryption for external devices optionally additionally, but not as replacement of software encryption.

»»» ENCRYPTION OF E-MAILS & FILE TRANSFERS
Be aware of how your E-mails are encrypted!

- » Use asymmetric encryption for e-mails to customers.
- » Client-based encryption preferred, avoid gateway-based systems.
- » SSH or SFTP instead of FTP!
- » Use symmetric encryption for file-transfers.
- » Attention to valid SSL-encryption for transfer of data!

»»» BACKUP STRATEGY
Backup your work but not data!

- » Store backups on hardware at different locations.
- » As Data Analyst / Engineer / Scientist you are not the administrator, backup your work, not the source data!
- » Do not backup confidential data! If data is lost, request a new instance from the data source managers / admins.

»»» SERVER & CLOUD SYSTEMS
Do not trust external systems without a deeper look!

- » Do not store personal or customer data in the cloud or on public server without allowance by the owner.
- » For servers, disallow root user login.
- » Reduce the allowed amount of login attempts.
- » Change standard ports (e.g., ssh port 22 to 31232).
- » Do not use public file converters (e.g., PDF to Word file transformer on the internet).

»»» USER ACCESS MANAGEMENT
Access is key for security!

- » Be aware to store passwords hashed in SHA-1/-2/-3.
- » Attention to clear user role definitions, e.g., super-admin vs admin vs owner vs user.
- » User roles and users should be documented by the systems itself, otherwise via separated documents.
- » Monitor user access and data consumption by users.
- » Delete not needed users and user roles, especially root-users or super-users.

»»» DATA PRIVACY
Respect personal data!

- » Read, understand and consider the GDPR and your regional laws!
- » Do not store personal data without allowance from each person.
- » Do not analyze personal data without allowance from each person.
- » Use anonymization / pseudonymization of data with personal context.
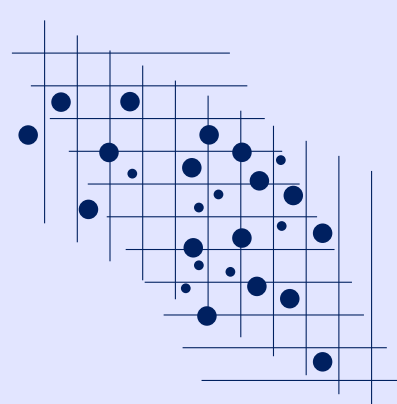
»»» PROTECTIVE SOFTWARE
Use professional tools for protection!

- » Activate the Anti-Virus-Protection of your operating system.
- » Consider using Firewall tools for critical servers / devices.
- » Use Proxy Server / SSH tunnels / Virtual Private Networks in between data providers (e. g. database servers) and clients.
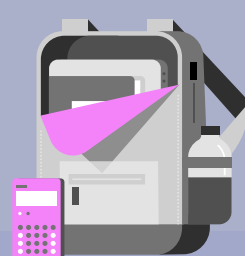
»»» NON-DISCLOSURE AGREEMENT & COPYRIGHT
Respect data and their owners!

- » Define, understand and sign a strict Non-Disclosure Agreement (NDA).
- » Do not start accessing data before the NDA is signed by both parties.
- » Independent of the NDA, always handle data from clients with high priority regarding data security.
- » Respect the ownership of data. Do not use data for purposes other than negotiated.
- » Respect the copyright of data stored in external data sources as well. Do not use open data if the allowance for this is unclear.

DATANOMIQ
Applied Data Science