Data Leakage Prevention durch Einsatz geeigneter Tools

GI-SECMGT-Workshop Februar 2011

Dipl.-Inform.Med. Christian Uwe Götz

Definition „Tool“

- “Tool”
  - a piece of equipment which you use with your hands to make or repair something
  - something that helps you to do a particular activity

Quelle: Cambridge Dictionaries Online, http://dictionary.cambridge.org/

- Übersetzung in diesem Kontext: „Hilfsmittel“
  - Ziel: durch Einsatz eines „Hilfsmittels“ präventiv und/oder reaktiv etwas Positives erreichen

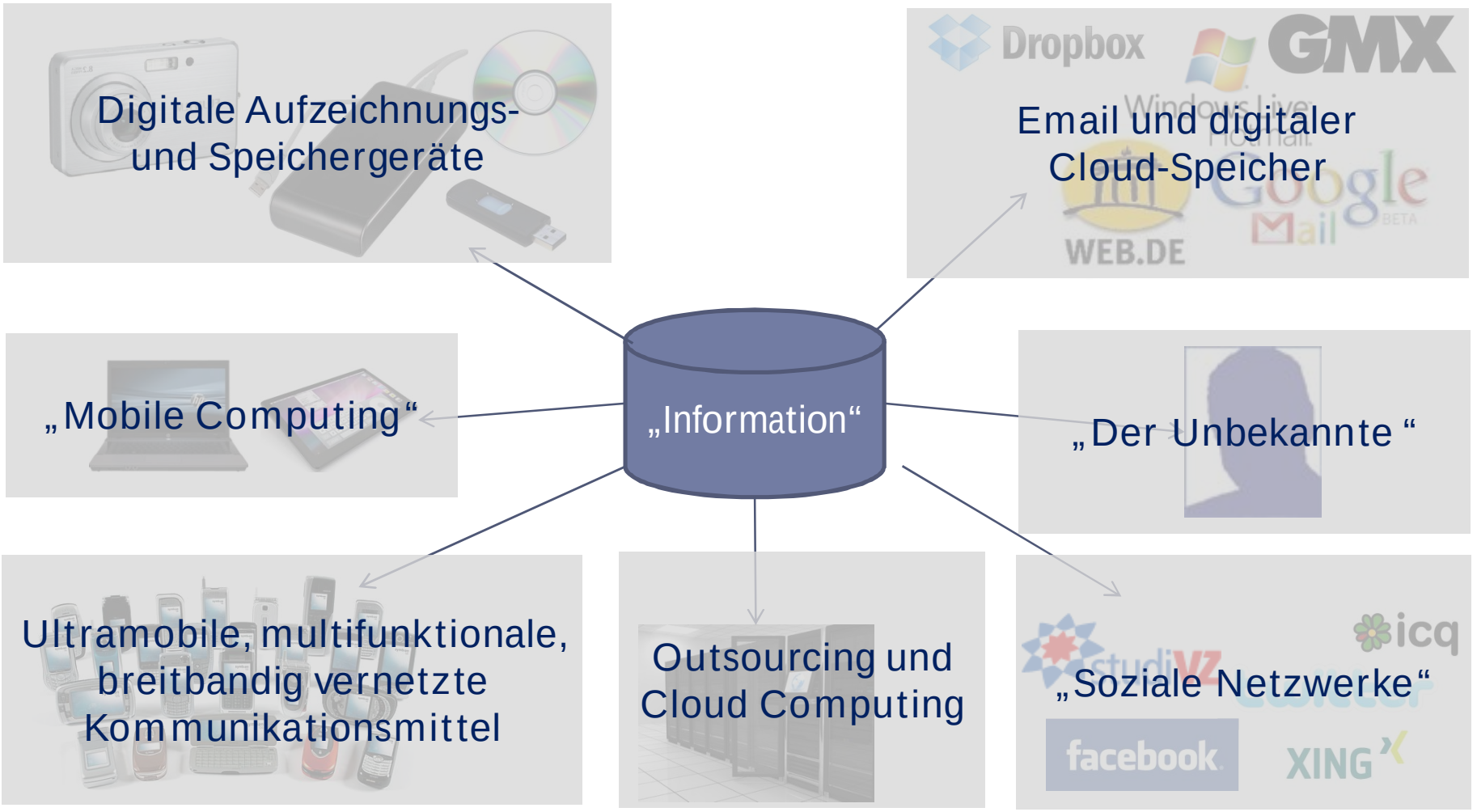
Begriffsdefinitionen für das Akronym „DLP“

- “Data Loss Prevention”
  - “Loss” – englisch für “Verlust”
  - Eher zufälliger, ungewollter Verlust von Daten und Informationen
  - Beispiele: Ein in der U-Bahn vergessener Laptop oder die CD-ROM mit dem Jahresabschluss auf dem Firmenparkplatz
- „Data Leakage Prevention“ (auch „Data Leak Prevention“)
  - „Leak“ – englisch für „Leck“
  - Bewußter, vorsätzlicher, gezielter Abfluss von Daten und Informationen
  - Beispiele: „Tweets“ aus der Vorstandssitzung heraus oder die Email mit den Ein- und Verkaufspreisen für die Produktlinie 2012-2013 an das private Postfach bei „Freemailer xyz“
- Dies sind freie Definitionen des Autors im Rahmen dieser Präsentation, die zunächst keine Allgemeingültigkeit haben!

Wir brauchen „Tools“!

Aber welches? Und wofür?

Moderner Informations- und Datenfluss ...

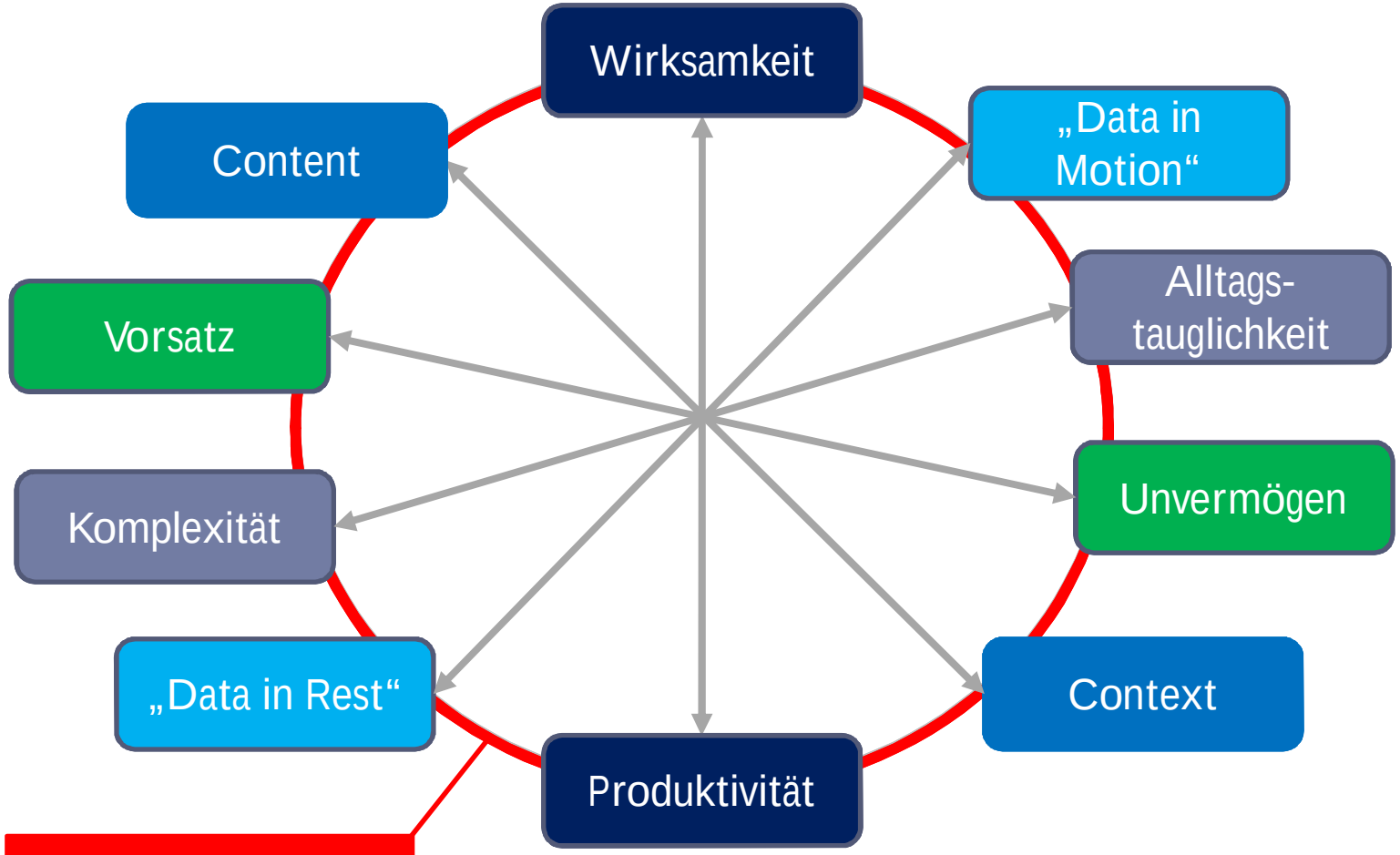


Hinweis: Bei den hier dargestellten Objekten und Lösungen handelt es sich um Beispiele!

Herausforderungen

- Alle zuvor genannten „Tools“ sind „Gewohnheitsrechte“ und integraler Bestandteil unsere Lebens geworden
  - „Statussymbole“, Ausdruck von Freiheit und persönlichem Stil
- Immense Rechenleistung und Speicherkapazitäten die zur Verfügung stehen
- Grenzen und Funktionen der einzelnen „digitalen Tools“ verschwimmen zunehmend
  - z.B. „Smartphone“ – Telefon, Kleincomputer, Massenspeicher, …?
- Viele dieser „Tools“ sind gratis oder sehr kostengünstig zu haben
- Sie sind nahezu für jedermann verfügbar und auch mit wenig IT-Knowhow zu nutzen und zu bedienen

Einige Antagonisten im DLP-Umfeld ...

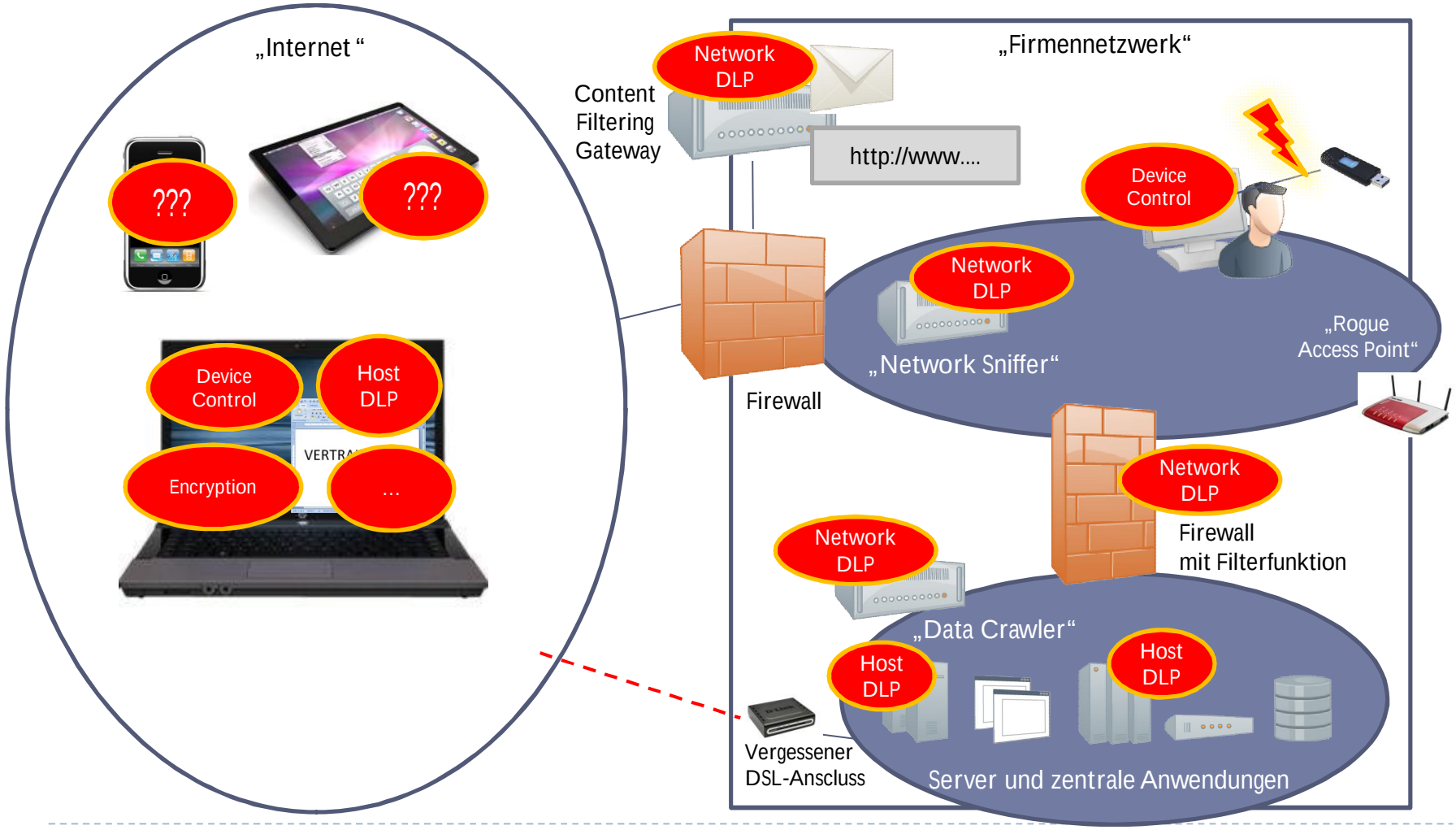


Typen von DLP-Tools

- Netzwerk-basierte DLP Lösungen
  - Fokussierung auf in der Übertragung befindliche Datenströme z.B. auf Gateways und Proxies
- Host-basierte DLP Lösungen
  - Einsatz von lokal auf Zielsystemen installierten „Agenten“
  - Kontrolle von Schnittstellen und Prozessen, etc. auf dem Endgerät
- „Information Identification“-basierte DLP Lösungen
  - Einsatz von Hilfsmitteln zur Identifikation und Klassifikation schützenswerter Daten
  - „Discovery-based“,„Data Crawling“, …

Einsatzbereiche von DLP-Tools

Beispielhafter Lösungsansatz (mit Potential zur Verbesserung)



Die Frage nach den richtigen „Tools“ …

- Warum glaube ich ein DLP Tool zu brauchen?
- Wonach suche ich? Gibt es eine „Datenklassifikation“?
  - Versuche auf letztere Frage eine Antwort zu finden ist oft der Anfang vom Ende eines DLP-Vorhabens …
- Was möchte ich finden bzw. verhindern?
- Für wen bzw. gegen wen sollen die durch Tools umgesetzten „Maßnahmen“ greifen?
- Welche Kommunikationskanäle muss ich betrachten?
- Wie wirksam sollen bzw. können die durch Tools umgesetzten „Maßnahmen“ sein?
- Gibt es denn wirklich eine „All-in-One“-Lösung dafür?

Beispiele für DLP Tools

Ein Ausflug

Hinweis

- Dies soll keine Marktanalyse oder vollständige Übersicht über heute am Markt verfügbare DLP-Lösungen sein!
- Ziel ist es vorhandene technische Lösungsansätze gegenüberzustellen und einige, wenige Lösungen exemplarisch vorzustellen!

DLP-Tools für „Data in Rest“

- „Data in Rest“ – Daten die sich auf einem Speichermedium, z.B. einer Festplatte oder eine Storage Area Network (SAN) befinden und nicht in der Übertragung sind
- Anforderungen an die Funktionsweise eines solchen Tools:
  - „Der sichere Hafen für meine Daten! “
  - Geschützte Ablage von Daten (Verschlüsselung)
  - Implementierung ausgefeilter Rechte- und Rollenkonzepte (Access Control)
  - Umsetzung des „*least* privilege“ und „need-to-*know*“-Prinzips („Ich sehe nur, was sehen soll“)
  - Auditierung und Nachvollziehbarkeit
  - …

Beispiel: Cyber-Ark Digital Vault®

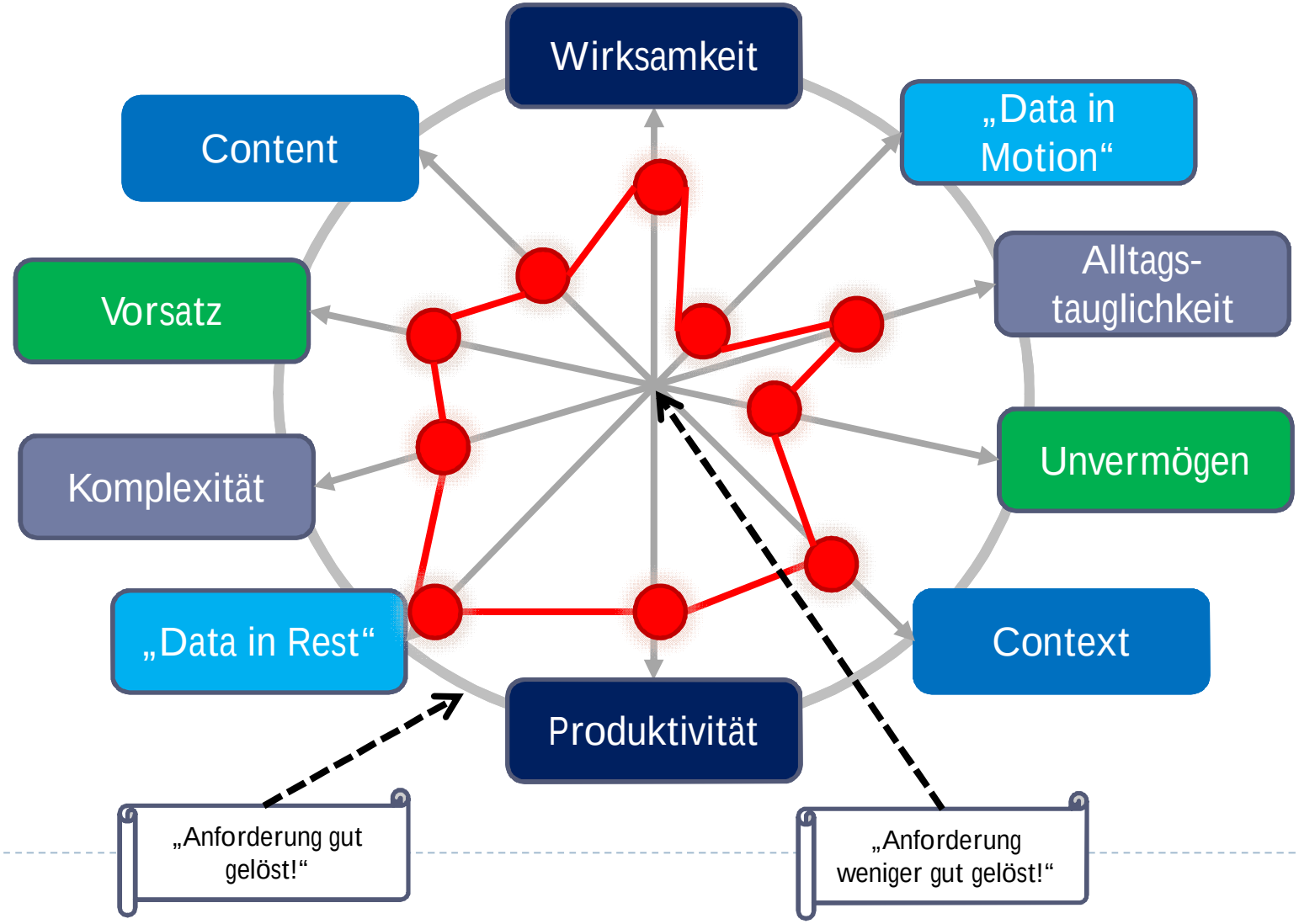
- Digitaler Datentresor
  - Verschlüsselung aller Daten ohne aufwändiges Key Management
  - Datenzugriff nur über Applikationsebene, nicht über OS
  - Klare Trennung „Administrator“-“Datenbesitzer“
  - Volle Auditierung aller Zugriffe
  - …

  - Aber: Schützt nur Daten die auch im “Vault“ liegen

Weitere Beispiele für „Data in Rest“-DLP …

- Sauberes Rollen- und Berechtigungskonzept
  - Files & Folders, auf Applikationsebene, …
  - Rechte werden nicht nur vergeben sondern auch wieder entzogen!
- „File & Folder“-Encryption Lösungen
  - Flächendeckender Einsatz von „Verschlüsselungslösungen“ für alle Typen von Datenspeichern
- „Device Control“
  - Welche „Devices“ sollen legitim sein, d.h. müssen von der DLP-Lösung unterstützt werden?
- …

Einstufung „Data in Rest“-DLP-Tools



„Content Aware“-DLP-Tools

- „Content Aware“ – Suche nach klassifizierten „Inhalten“ durch Methode wie z.B. Pattern Matching, Fingerprinting, statistische Methoden, Reguläre Ausdrücke, …
- Werden oft auch als „Data in Motion“-DLP-Tools bezeichnet
  - „in vitro“-Anwendung der Methoden auf in Übertragung befindliche Daten
- Einsatz auch in Kombination mit „Information Identification“-basierten DLP-Lösungen
  - Durchsuchen von „Data in Rest“ als Basis zur Definition von „Suchmustern“, die dann auf „Data in Motion“ angewendet werden

„Content Aware“-DLP-Tools

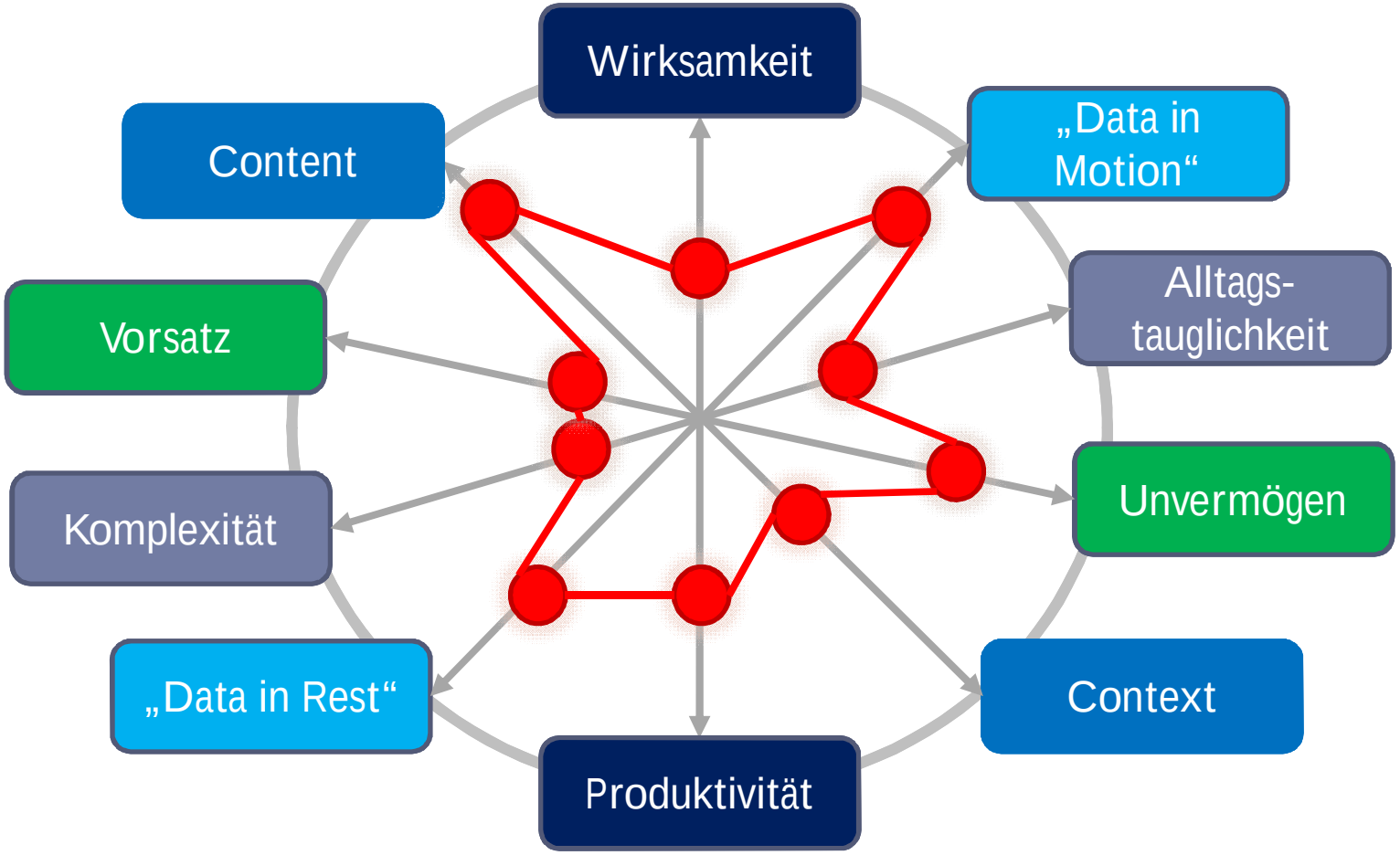
- Anforderungen an die Funktionsweise eines solchen Tools:
  - Unterstützung möglichst vieler Datenkanäle und Gerätetypen
    - „Alle Operating Systems“, „Email“, „Surfen“, „Handy“, …
  - Hohe Erkennungsrate
    - Unterstützung unterschiedlichster Datenformate und –typen
  - Robustheit der Algorithmen
    - „Deep Content Inspection“
    - „Akquisition Firma X“ vs. „Akqu1s1t1on F1rm8 X“
  - Optimale Positionierung im Datenstrom
  - Übergabe von validem Input, was als schützenswert anzusehen ist
    - „Daten“ vs. „Information“
    - „Wer sagt wem, wonach er suchen soll?“

Beispiel: Websense Triton Solutions

- Gateway-basierter Ansatz (u.a.)
- Fokus auf Email + Web und darüber übertragene Inhalte
- Möglichkeit zur granularen Definition von Suchmustern

Abbildungen von http://www.websense.com

Einstufung „Content Aware“-DLP-Tools



„Context-Based“-DLP-Tools

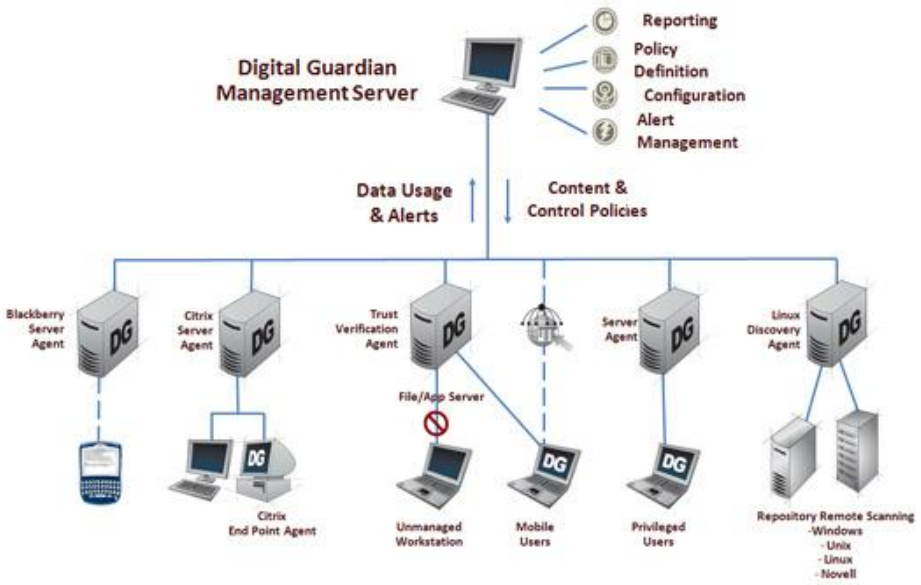
- „Context-based“ – Es werden primär Datenquelle und –ziele analysiert, sowie alle Prozesse auf einem Endgerät, die Daten aus einer schützenswerten Quelle verarbeiten.
  - Beispiel: „Office Document aus Share „Entwicklung“ wurde geöffnet; Anwender will Passage mit Copy & Paste in Email einfügen und an einen unbekannten Empfänger schicken“
- Einsatz meist auf dem Endgerät, da nur dort die auf die zu schützenden Daten angewandten Operationen kontrollierbar sind.

„Context-Based“-DLP-Tools

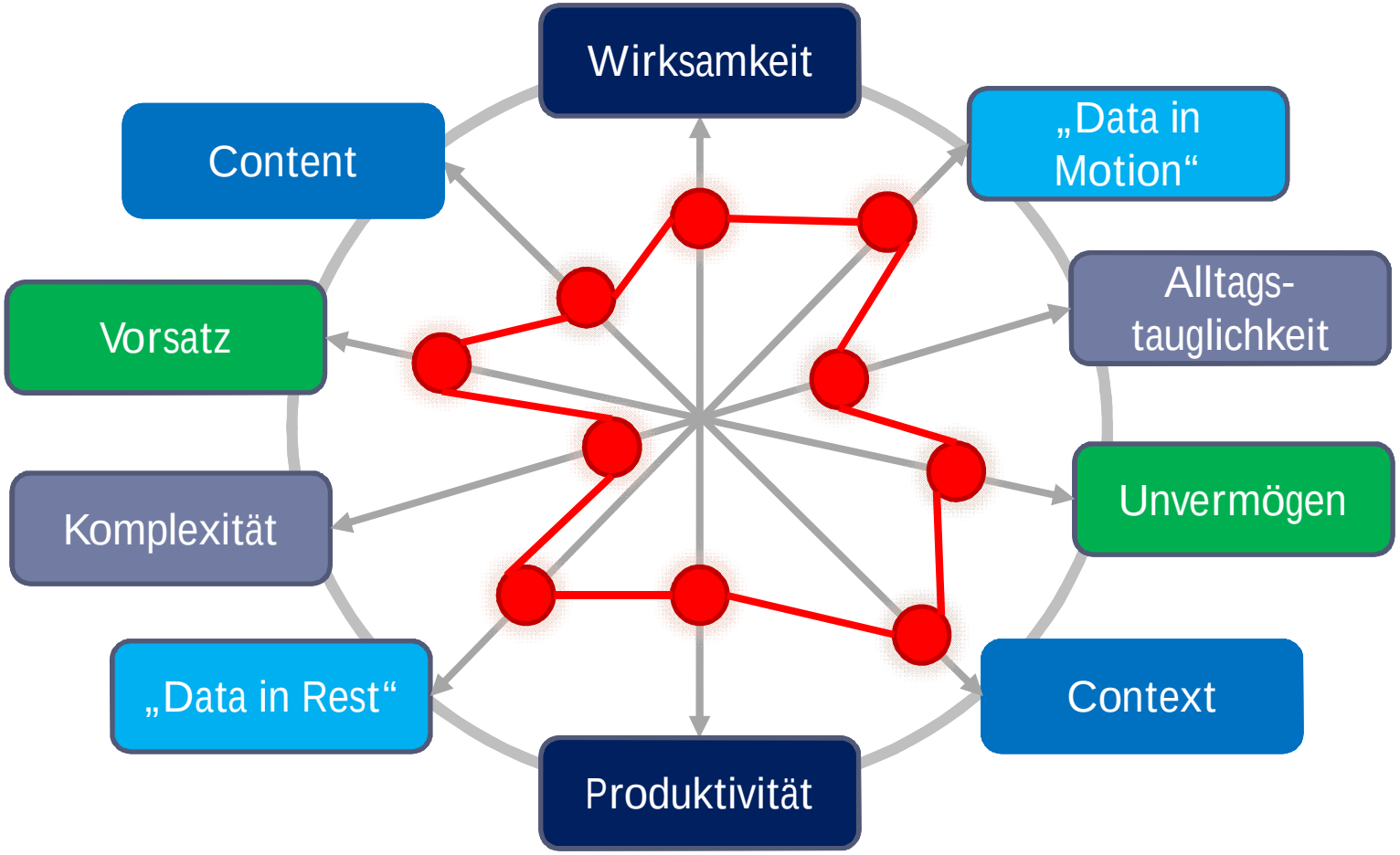
- Anforderungen an die Funktionsweise eines solchen Tools:
  - Unterstützung möglichst vieler Datenkanäle und Gerätetypen
    - „Alle Operating Systems“, „Email“, „Surfen“, „Handy“, …
  - Hohe Erkennungsrate
    - Unterstützung unterschiedlichster Datenformate und Prozesstypen
  - Robustheit der Anwendung
    - Anwender darf DLP-Tool nicht einfach deaktivieren können
  - Agenten-basiert
    - „In the middle of the action“
  - Übergabe von validem Input, was zu verhindern ist
    - „Wo liegen sensitive Daten (Quellen)?“
    - „Was darf damit (nicht) gemacht werden?“
    - …

Beispiel: Verdays Digital Guardian

- Hostbasierter Lösungsansatz
  - Überwachung von Schnittstellen und Prozessen
- Policy gibt vor was erlaubt ist
  - Eigene Skriptsprache
- Agentenfunktion im und außerhalb des Firmennetzes gegeben
- Optional Verschlüsselung von Daten
- Lösungssuite wurde um „Content Aware“-Funktionen erweitert
- Aber: Plattformabhängigkeiten durch Agentenbasierten Ansatz



Einstufung „Context-Based“-DLP-Tools



Fazit

- Die perfekte Lösung scheint es (noch) nicht zu geben ☹
  - Aber müssen es immer 100% Zielerfüllung sein?
  - Wie wäre es für (den Anfang) mit 80-20?
- „DLP“ ist technisch wie auch organisatorisch ein sehr vielschichtiges Thema
  - Oft fehlt es an der konsequenten Umsetzung
  - … und man hat (noch) nicht alle technischen Anforderungen im Griff
- Es gibt nicht die eine, allumfassende technische Lösung
  - Ohne Klassifikation, Policies und Richtlinien, gesunden Menschenverstand, … und eine Kombination mehrerer technischer Lösungsansätze wird man nicht oder nur bedingt erfolgreich sein
- Für DLP wird es in einer freien Gesellschaft nie eine 111%-ige Lösung geben!
  - Das allumfängliche Einschränken und Verhindern wird kaum möglich sein, ohne die Freiheit des Einzelnen im Umgang mit Daten und Informationen massiv zu beeinträchtigen.