



teaching, assessment, learning

# Welche Software rettet die Welt?

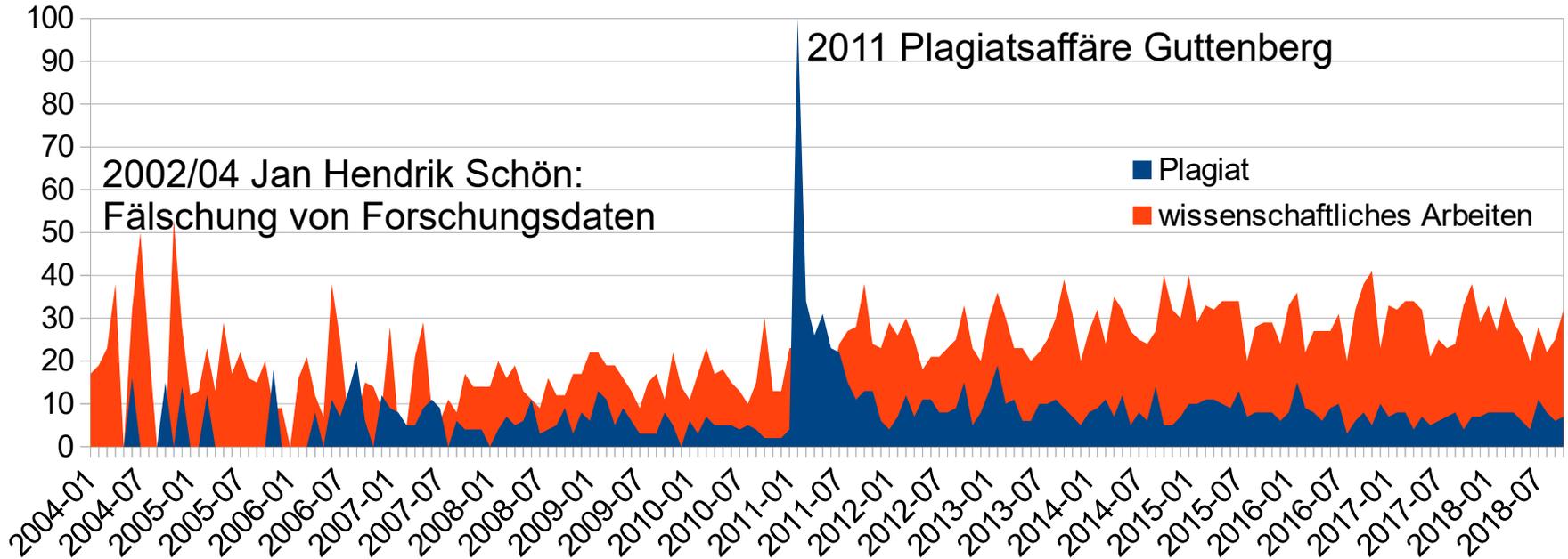
## Ein Werkstattbericht

Armin Glatzmeier



## Präludium: Gutten-Peak

*Häufigkeit der Suchanfragen deutscher IPs bei Google in Relation zum Anfragevolumen.*



## Auftrag und Ziel des Projekts

- (1) Test und ggf. Empfehlung von ‚Anti-Plagiatssoftware‘ (APS) für den zentralen, universitätsweiten Einsatz.
- (2) Evaluation strategischer, organisatorischer, technischer und rechtlicher Grundlagen und Rahmenbedingungen eines APS-Einsatzes.
- (3) Identifikation, Planung und Konzeption zentraler und/oder dezentraler Unterstützungsstrukturen zur Plagiatsprävention bzw. Sicherung guter wissenschaftlicher Praxis.

# Softwaretest: Vorbemerkung

- (1) Abgleich schriftlicher Prüfungsleistungen mit veröffentlichten Texten
- (2) Abgleich schriftlicher Prüfungsleistungen mit Arbeiten anderer Studierender (,Querverprobung‘)
- (3) Textgattungen
  - (a) Wissenschaftliche Abhandlungen
  - (b) Programmcode

# Softwaretest: Erkenntnisinteresse

- (1) Primär
  - (a) Wie viele Plagiate der Testdokumente werden erkannt?
  - (b) Welche Art von Plagiaten wird erkannt?
- (2) Sekundär
  - (a) Dateiformate?
  - (b) Zugänglichkeit der Quellen?
  - (c) Qualität der referenzierten Quellen?

# Softwaretest: Testportfolio

## (1) Vorgehen

(a) Marktevaluation

(b) Einsatz an vergleichbaren Hochschulen

(c) Andere Softwaretests

## (2) Getestete APS

CopyLeaks, Docoloc, iThenticate, PlagAware, PlagiarismChecker X, Plagius, PlagScan, SafeAssign, Turnitin, Unicheck und Urkund

## Softwaretest: Testsetting

- (1) Zwei Testphasen
- (2) Zwei vollständig dokumentierte Testdokumente
- (3) Unterschiedliche Plagiatsformen (C&P, Paraphrase, Übersetzung, Struktur)
- (4) 23 Fächer (Schwerpunkt: Geistes- und Sozialwissenschaften)
- (5) Fünf Sprachen (Deutsch, Englisch, Französisch, Italienisch und Spanisch)
- (6) Vorab definiertes Kategorienschema für Bewertung möglicher Fundstellen

# Softwaretest: Bewertung der Referenzstellen

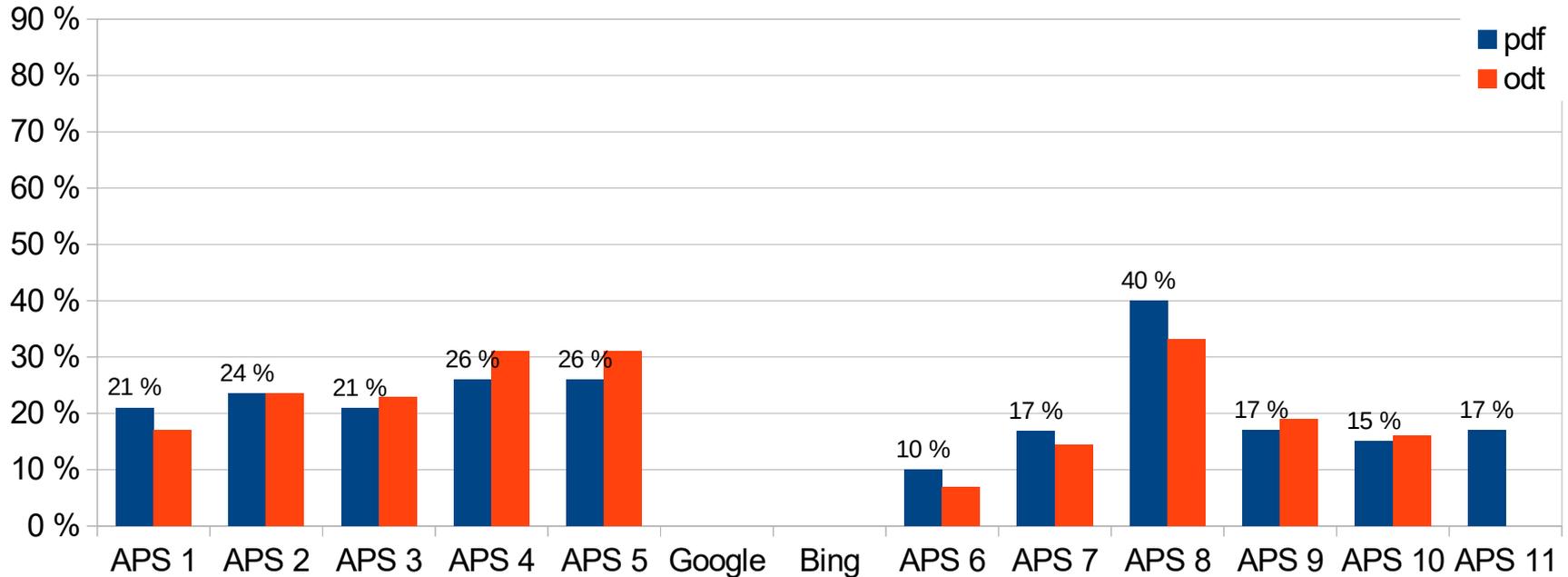
- (1) Direkte Quellen
- (2) Indirekte Quellen (Typ1)
- (3) Indirekte Quellen (Typ 2)
- (4) Irreguläre Quellen

# Softwaretest: Auswertung (Ausschnitt)

	Funde laut Prüfbericht		Erkannte Plagiatsinstanzen (abs.)		Erkannte Plagiatsinstanzen (%)		Effektiv erkannter Zeilenumfang (%)	
	pdf	odt	pdf	odt	pdf	odt	pdf	odt
APS 1	21 %	17 %	5 (+1)	4 (+1)	19 % (22 %)	15 % (19 %)	19 % (22 %)	15 % (17 %)
APS 3	21 %	23 %	8 (+1)	7 (+2)	30 % (33 %)	26 % (33 %)	31 % (35 %)	28 % (36 %)
APS 4	26 %	31 %	7	8	26 %	30 %	30 %	33 %
Google	—		21 von 27		78 %		81 %	
APS 8	40 %	33 %	10 (+2)	9 (+2)	37 % (44 %)	33 % (41 %)	42 % (49 %)	37 % (44 %)

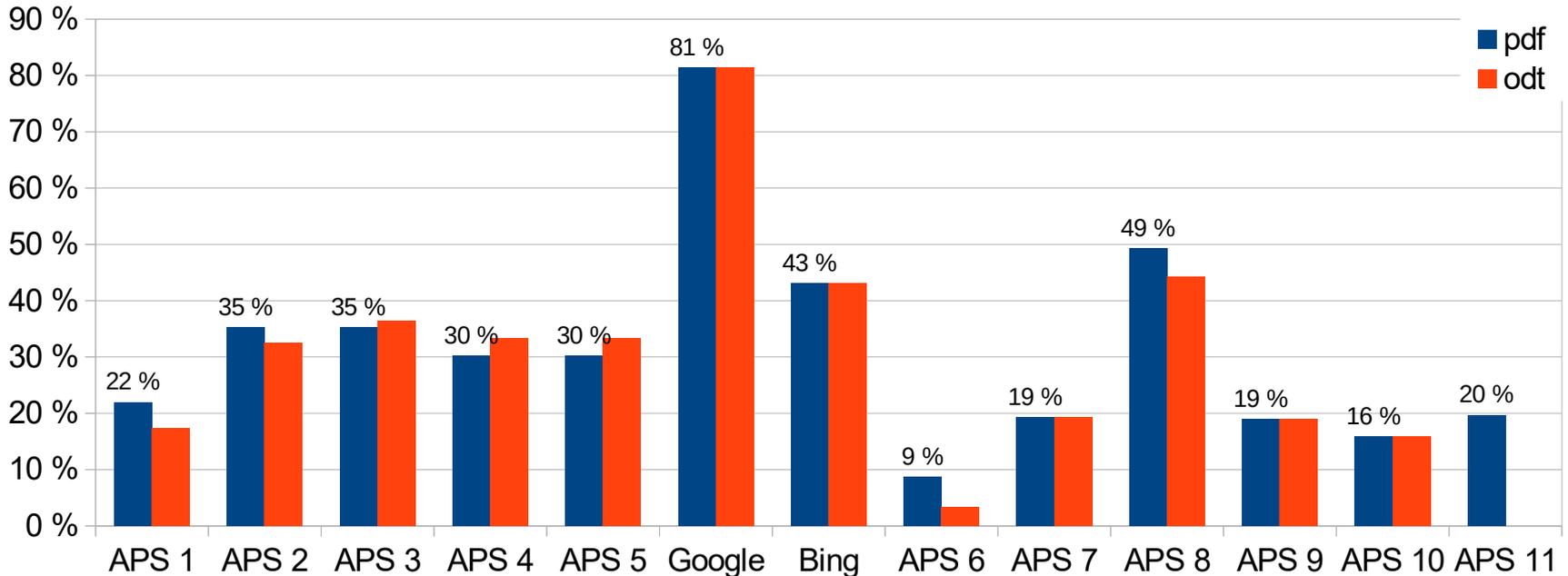
# Testdokument 1: Befunde laut Prüfberichten

*Funde in Testdokument 1 laut Prüfbericht bei 100% Plagiatsanteil. Entfällt bei Google und Bing.*



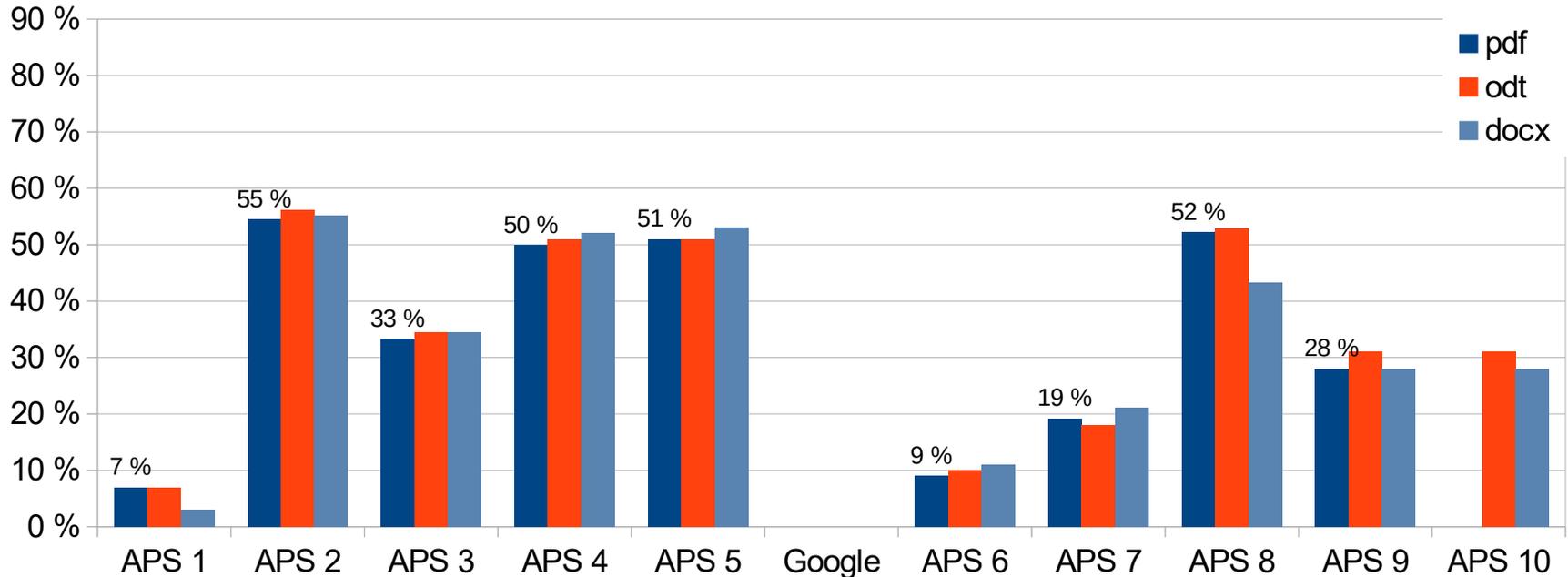
# Testdokument 1: Effektiv erkennbare Plagiate

*Effektiv erkennbarer Zeilenumfang (von 100%). Einschließlich Funde aus irregulären Quellen.*



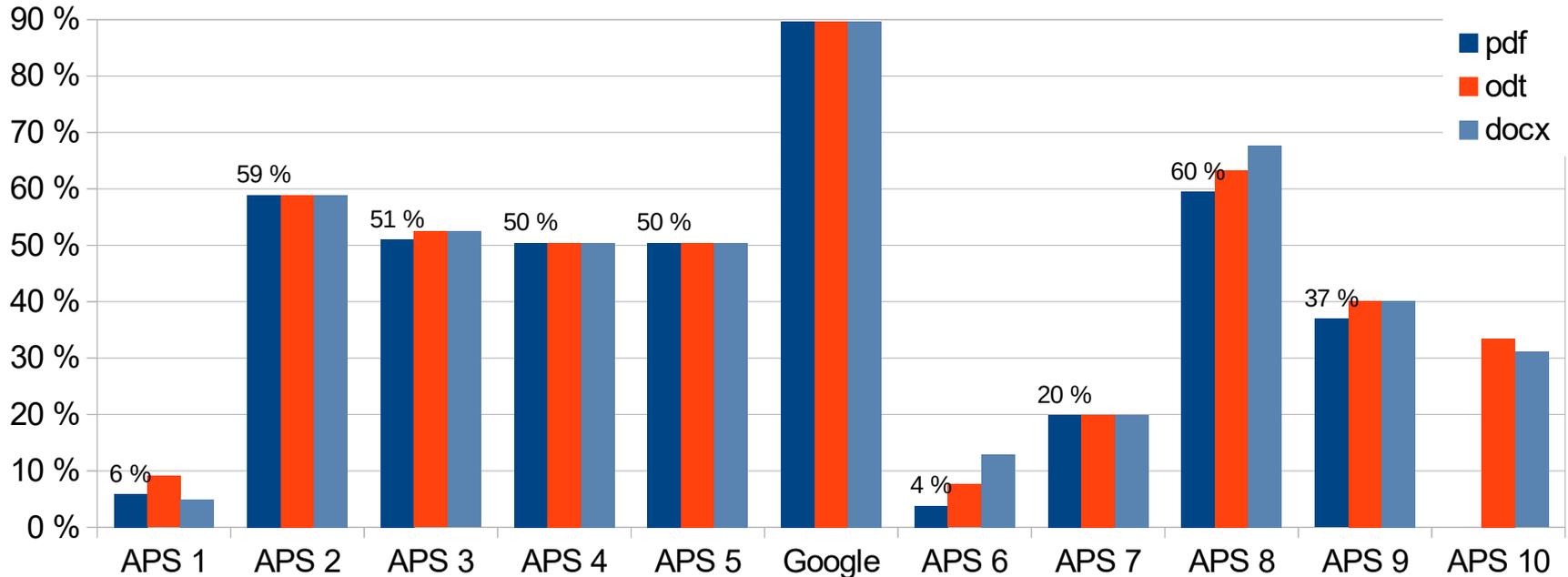
# Testdokument 2: Befunde laut Prüfberichten

*Funde in Testdokument 2 laut Prüfbericht bei 100% Plagiatsanteil. Entfällt bei Google und Bing.*



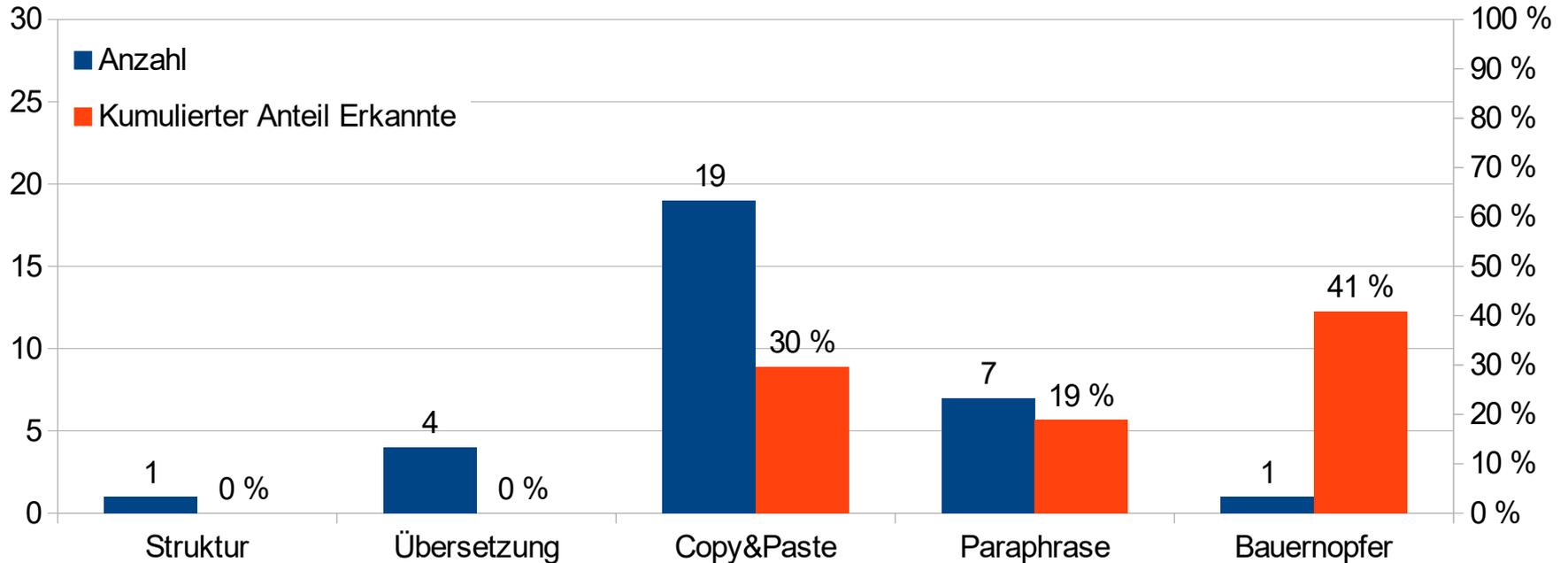
# Testdokument 2: Effektiv erkennbare Plagiate

*Effektiv erkennbarer Zeilenumfang (von 100%). Einschließlich Funde aus irregulären Quellen.*



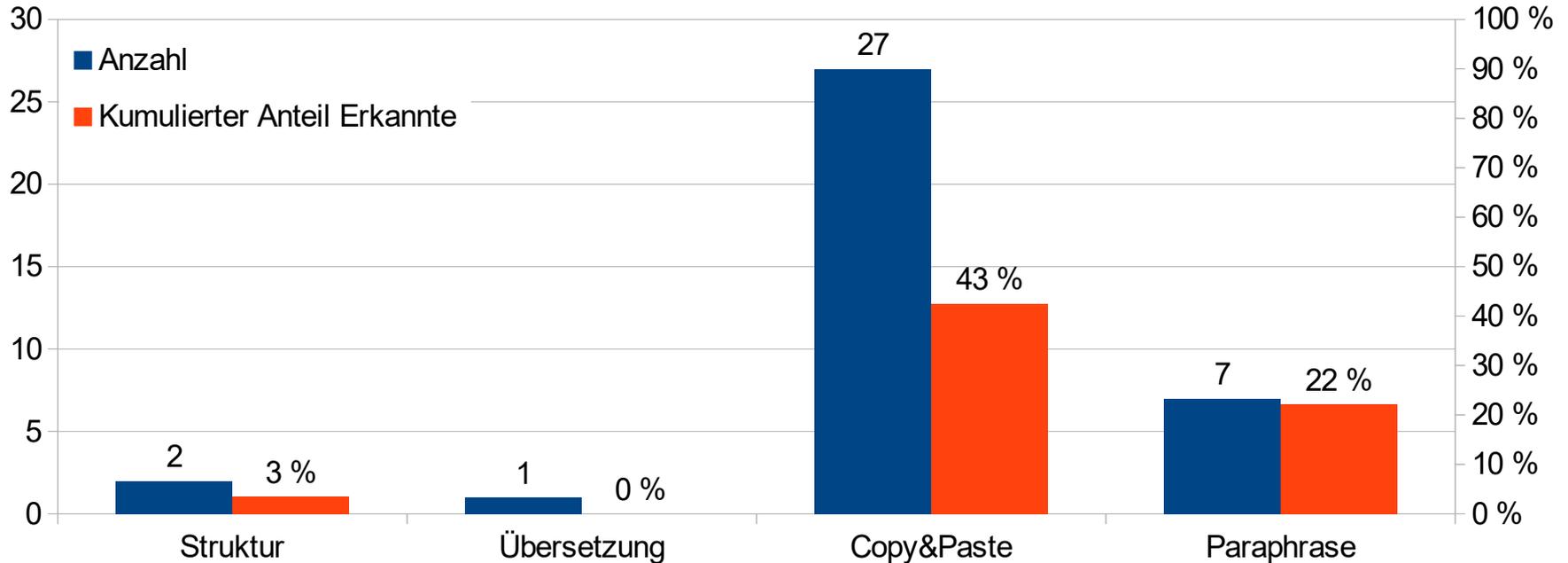
# Testdokument 1: Plagiatsvarianten

*Verwendete Plagiatsvarianten und Erkennungsrate. Kumulierte Funde aller APS im Test.*



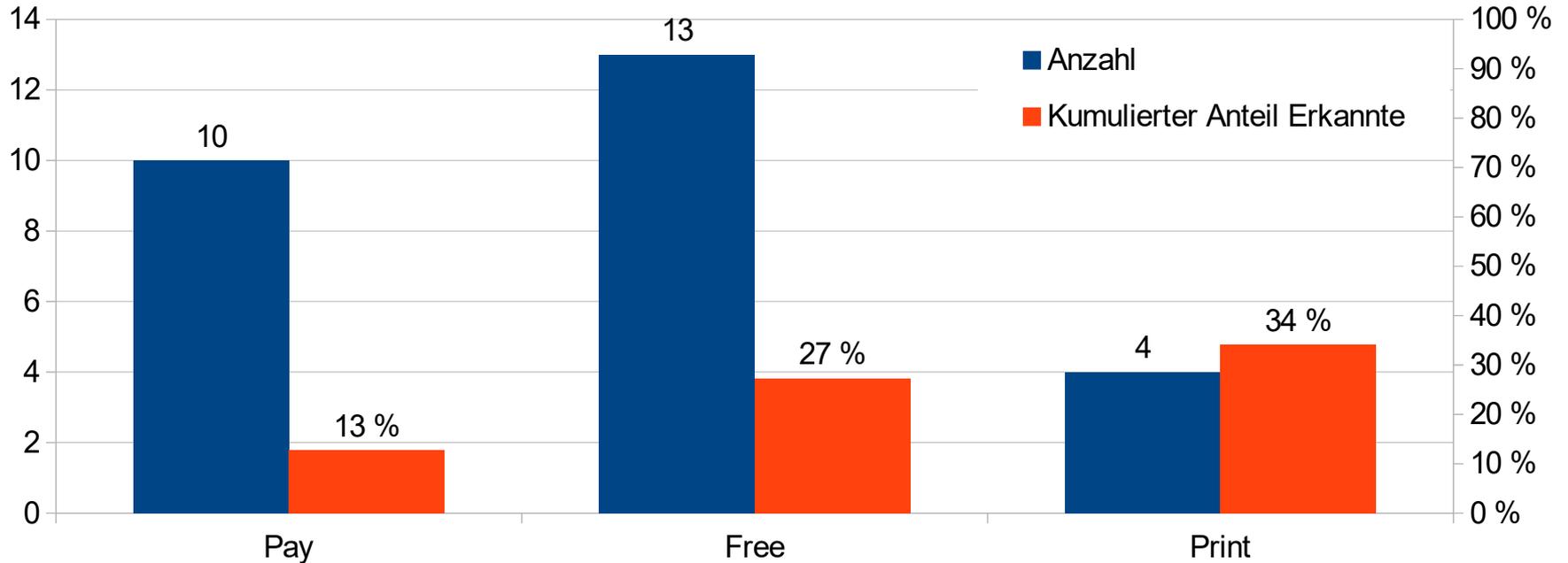
## Testdokument 2: Plagiatsvarianten

*Verwendete Plagiatsvarianten und Erkennungsrate. Kumulierte Funde aller APS im Test.*



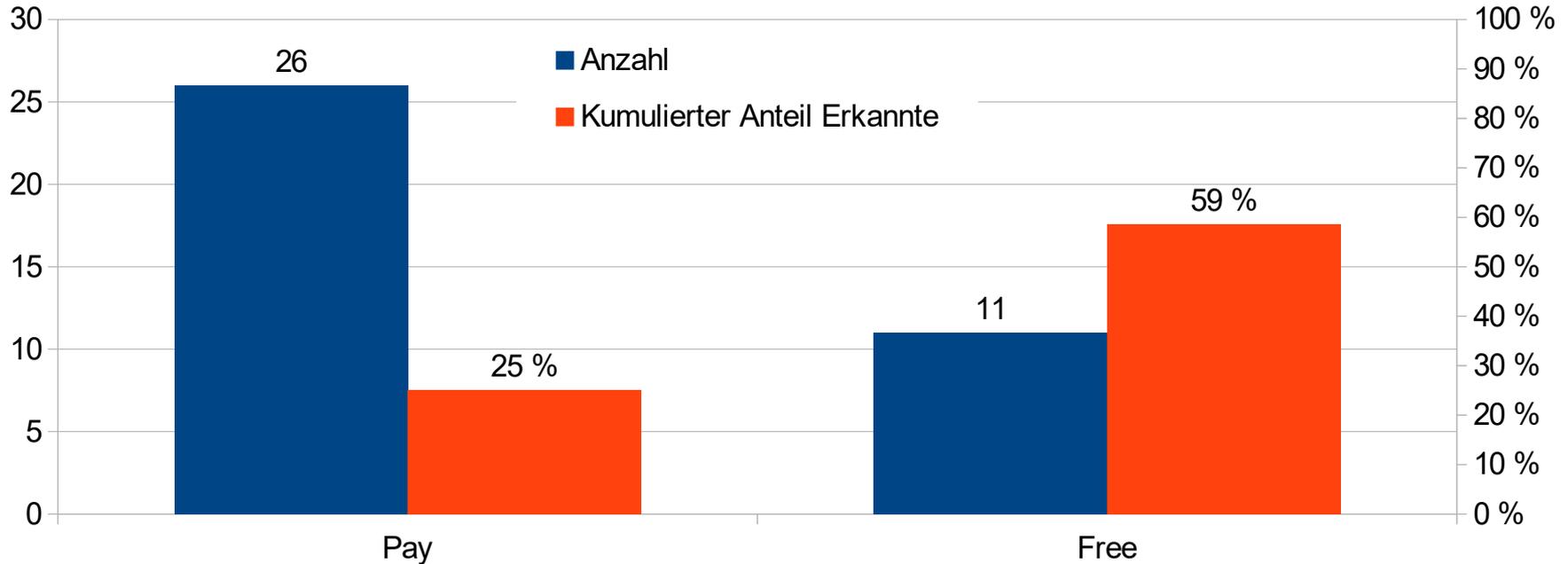
# Testdokument 1: Art der Quelle

*Verwendete Quellen: Art des Zugangs und Erkennungsrate. Kumulierte Funde aller APS.*



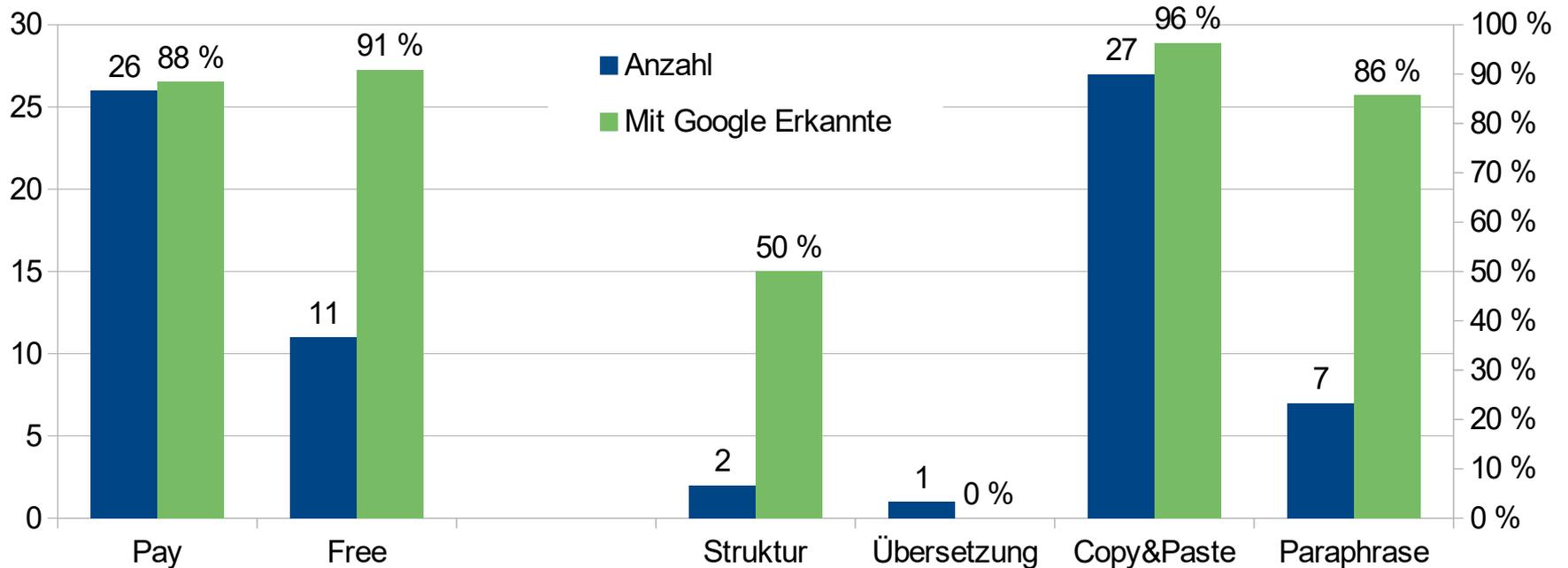
## Testdokument 2: Art der Quelle

*Verwendete Quellen: Art des Zugangs und Erkennungsrate. Kumulierte Funde aller APS.*



## Testdokument 2: Art der Quelle / des Plagiats

*Ergebnisse der Rekonstruktion mit Google.*



## Zentrale Ergebnisse

- (1) Eine vollständige Erkennung kann nicht geleistet werden
- (2) Frei zugängliche Netzinhalte werden besser erkannt
- (3) C&P-Plagiate aus freien Quellen werden nicht immer erkannt
- (4) Paraphrasen sind für alle APS im Test problematisch
- (5) Übersetzungsplagiate werden nicht erkannt
- (6) Sinnentstellungen werden nicht erkannt
- (7) Strukturplagiate werden nicht erkannt



teaching, assessment, learning

**Vielen Dank für Ihre Aufmerksamkeit!**

Kontakt: [academic-integrity@fu-berlin.de](mailto:academic-integrity@fu-berlin.de)

