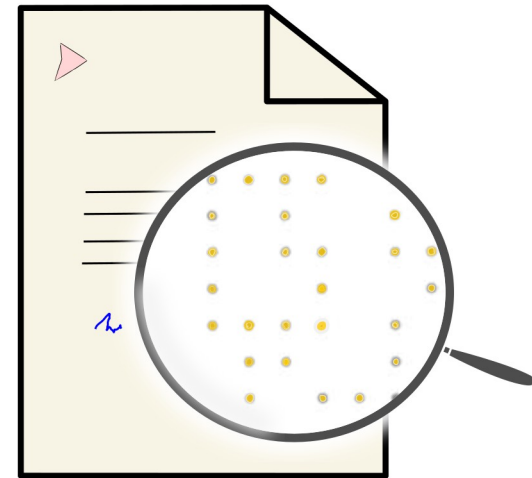


PRINTER FORENSICS



Stephan Escher
09.05.2019

Metadata

- Stephan Escher (stephan.escher@tu-dresden.de)
- Lehrstuhl Datenschutz und Datensicherheit – TU Dresden
- Projekt: Duplikatsprüfung und Forensik an gedruckten Dokumenten
- <https://dfd.inf.tu-dresden.de>

- Kooperationspartner: Dence GmbH (dence.de)
- Förderer: BmWi

Supported by:



on the basis of a decision
by the German Bundestag

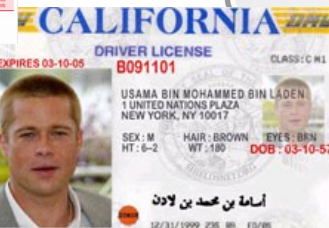
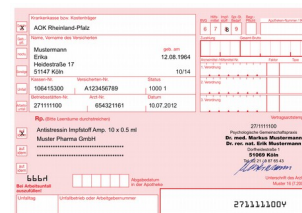
Motivation

- Last 30 years: development of high quality and low-cost printers, scanners, image manipulation tools

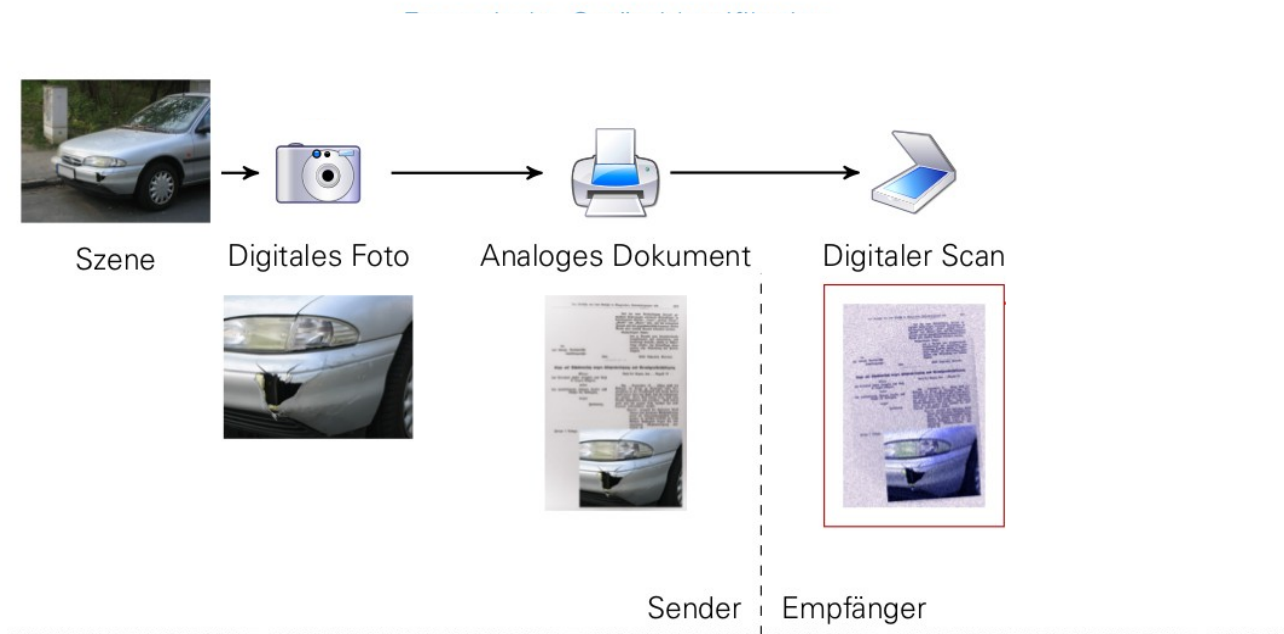


Motivation

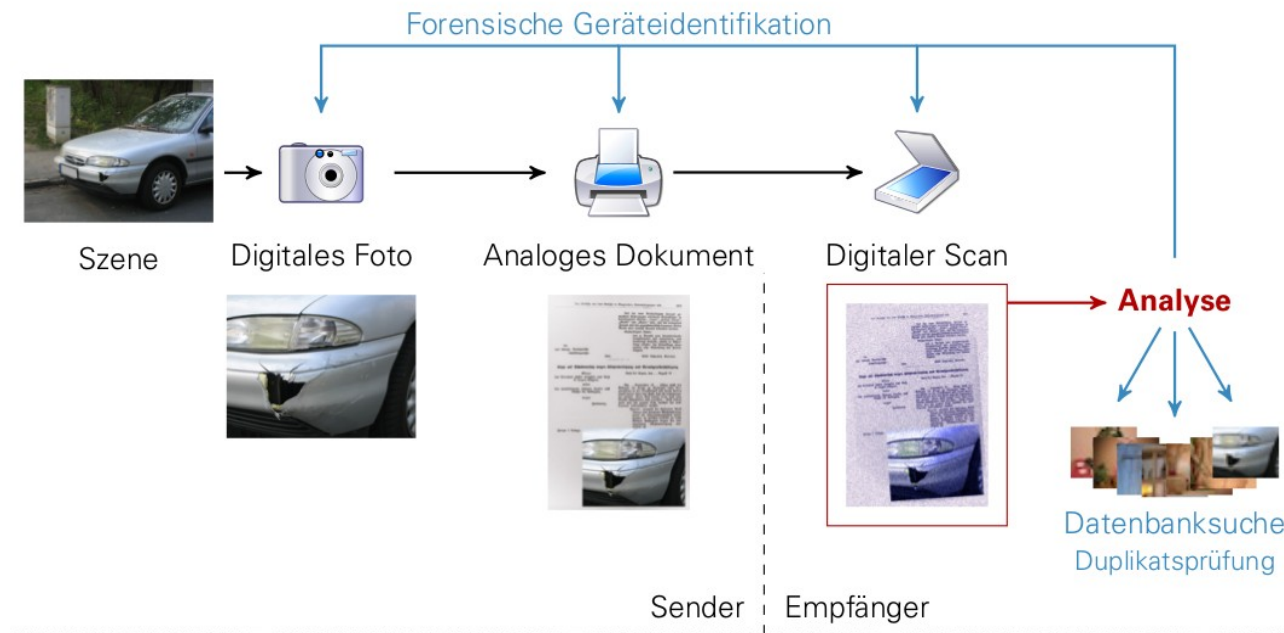
- Last 30 years: development of high quality and low-cost printers, scanners, image manipulation tools
- Used everywhere: PIDs, credentials, money, certificates, contracts, ...
- Anybody can create, manipulate and duplicate documents and images



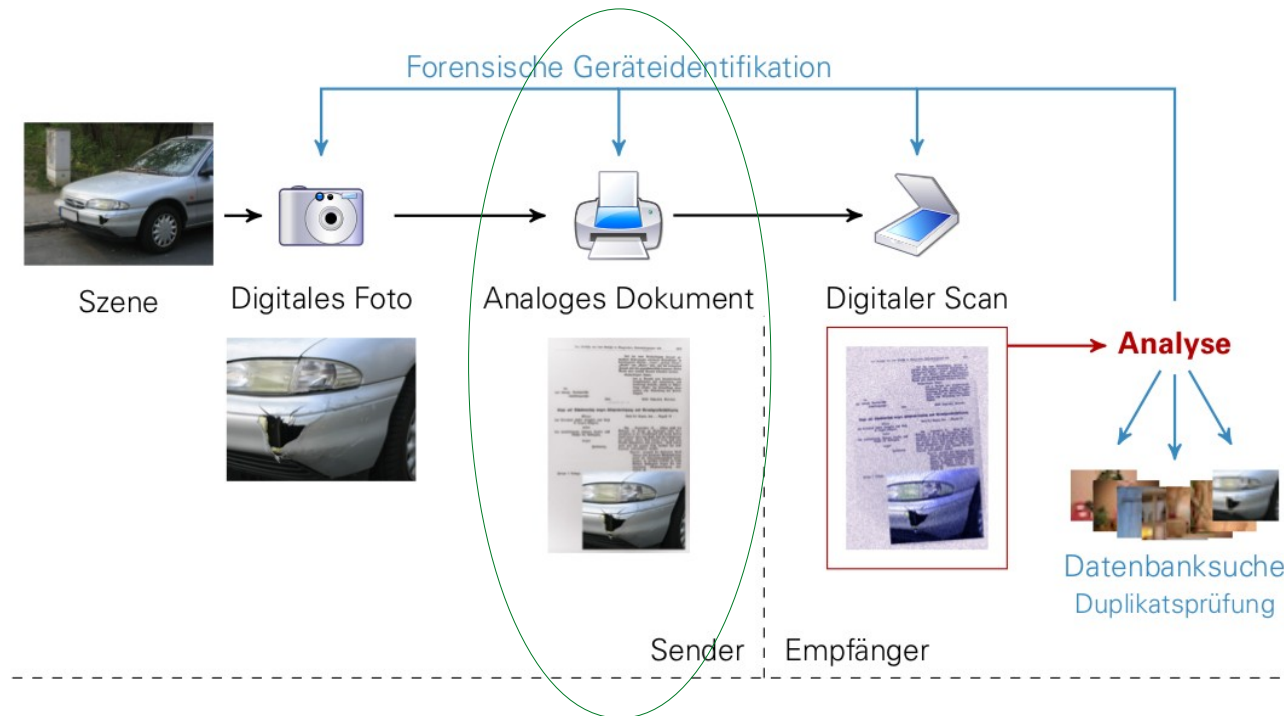
Project Use Case - Insurance



Project Use Case - Insurance

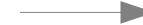


Printer Forensics



Printer Forensic Questions

- **Printer technology**
- Printer device
- Comparison of multiple documents
- Forgery detection
- Age of a document



Laserprinter
Inkjet
Photocopier
Offset
...



Printer Forensic Questions

- Printer technology
- **Printer device**
- Comparison of multiple documents
- Forgery detection
- Age of a document

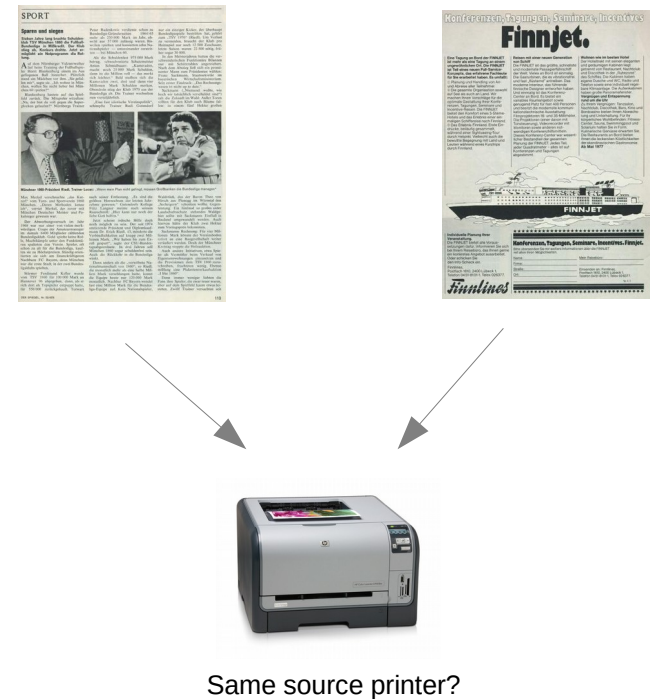


Brand HP
Modell M553
Device CNCXF526



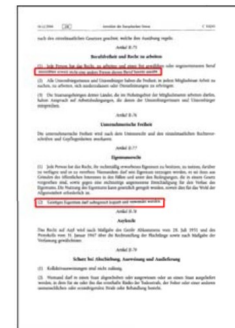
Printer Forensic Questions

- Printer technology
- Printer device
- **Comparison of multiple documents**
- Forgery detection
- Age of a document



Printer Forensic Questions

- Printer technology
- Printer identification
- Comparison of multiple documents
- **Forgery detection**
- Age of a document



Was the document
forged, ...?

Possible Solutions

- **Active Techniques**

embed proactively information (extrinsic signatures) in documents before or while printing

- **Passive Techniques**

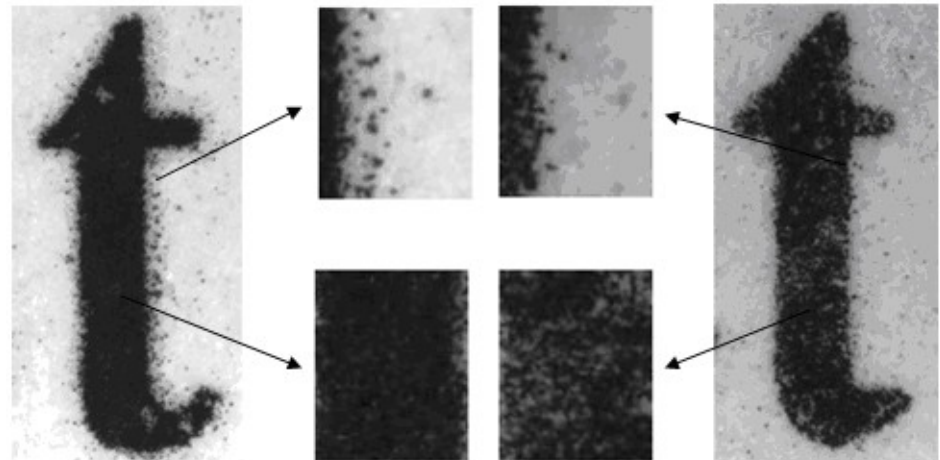
use print artifacts (intrinsic signatures) caused by the printing mechanism

Passive Techniques

- Use of intrinsic signatures
 - printing artifacts which are technology / brand / model / device dependent
 - electromechanical / mechanical imperfections, differences between constructions of printer models
 - should be stable over several printouts

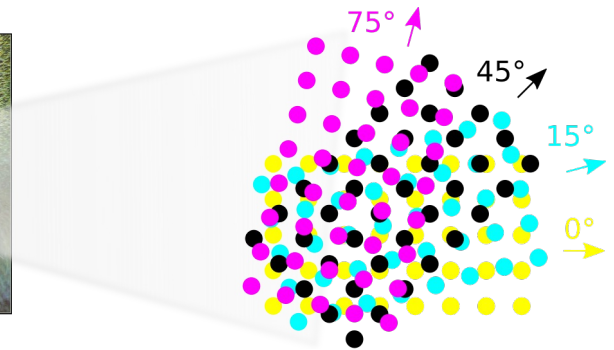
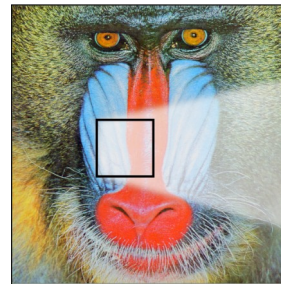
Intrinsic Signatures - Text

- **Micro textures**
- **Edge structur**
 - Roughness
 - Gradient
- Overspray
- Geometric distortion



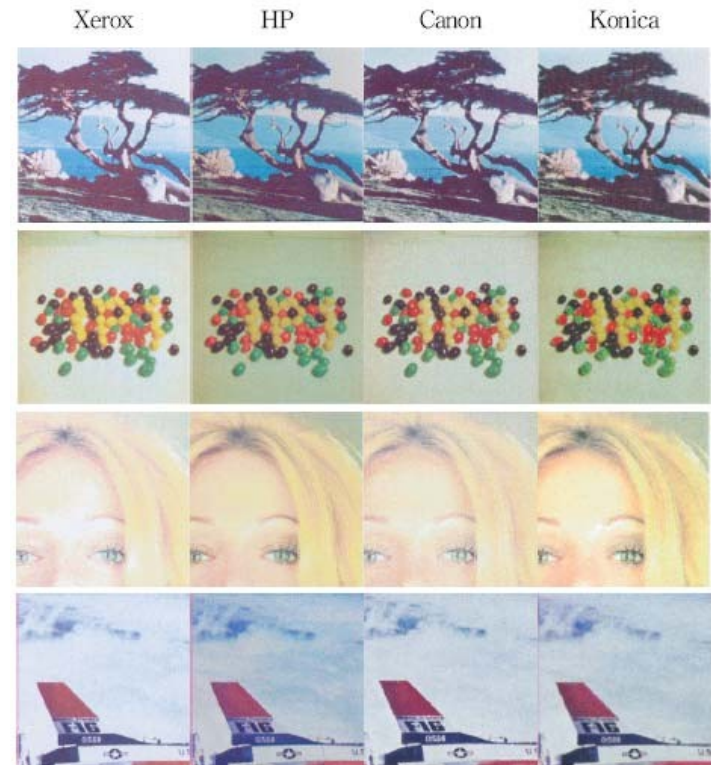
Intrinsic Signatures - Images

- **Halftoning**
 - Arrangement (AM)
 - Dot shape
- Color noise
- Geometric distortion (Banding, ...)



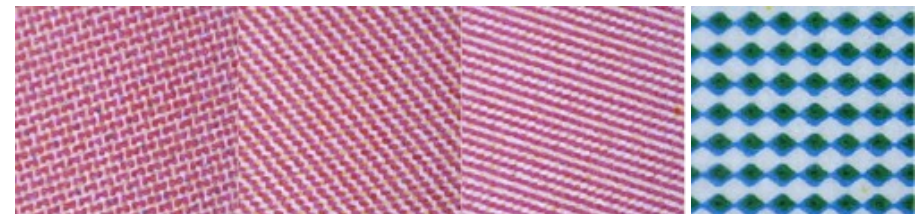
Intrinsic Signatures - Images

- Halftoning
 - Arrangement (AM)
 - Dot shape
- **Color noise**
- Geometric distortion (Banding, ...)



Findings

- Many potentially influencing parameters which could change the signature itself
 - Driver settings (e.g. toner save modi, resolution), age of the toner, used paper (plain vs. recycled), different font types, ...
 - Forgery of signature sometimes possible (e.g. halftoning)
- Possible overlaps for large datasets
- Max. identification rate: printer model
 - active methods are more accurate



Active Techniques

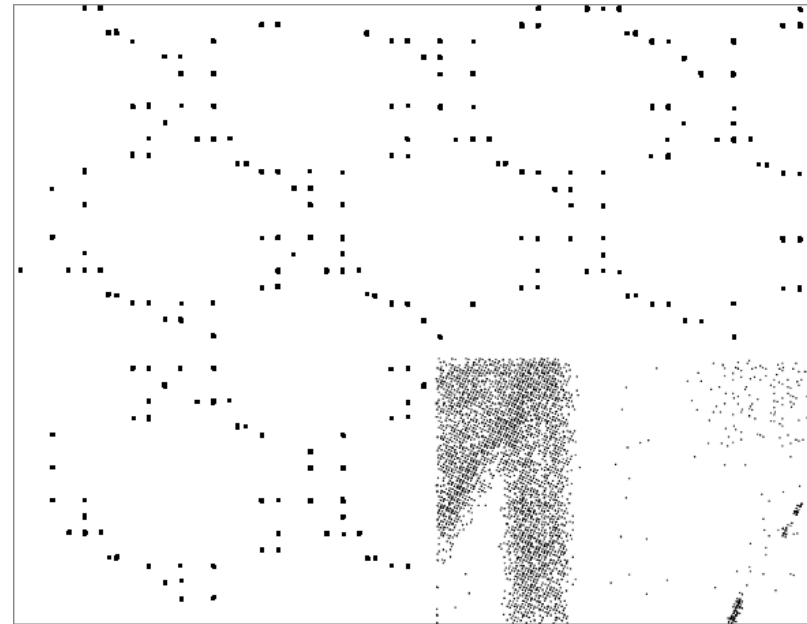
- Embed information in documents before or while printing (extrinsic signatures)
- Concrete information within the document
- Needs access to the document or printer device
- Unusable for project use case – except **Tracking Dots**

Tracking Dots

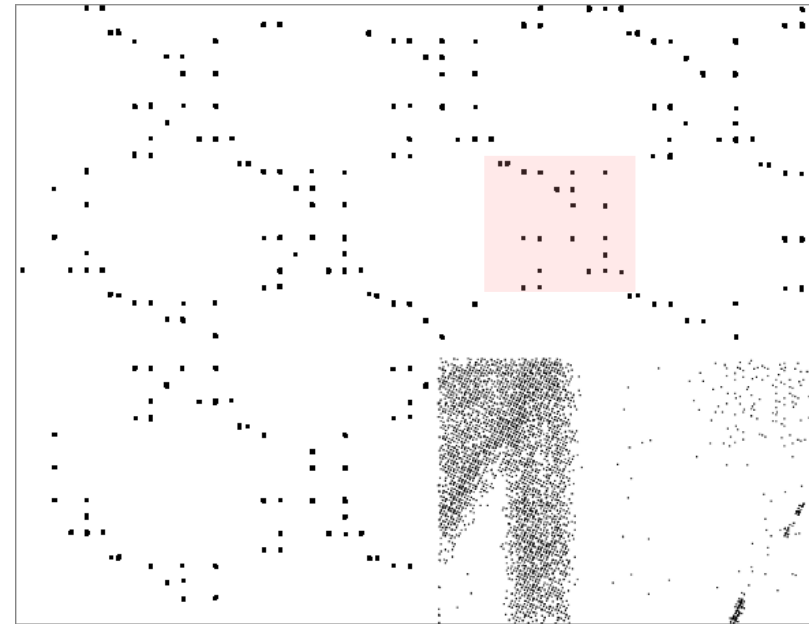
- Tiny Yellow Dots (~0.003 mm - not visible to the naked eye) ordered in matrices
- Repeated over entire document
- Implemented in colour laser printers itself
→ embedded while printing
- Found in 2005 by EFF and DFKI
 - Decoded 1 Pattern
 - Tracking dot pattern contains a serial number, date and time

→ *Reuse Tracking Dots for Project Use Case?*

Tracking Dots



Tracking Dots



Tracking Dots

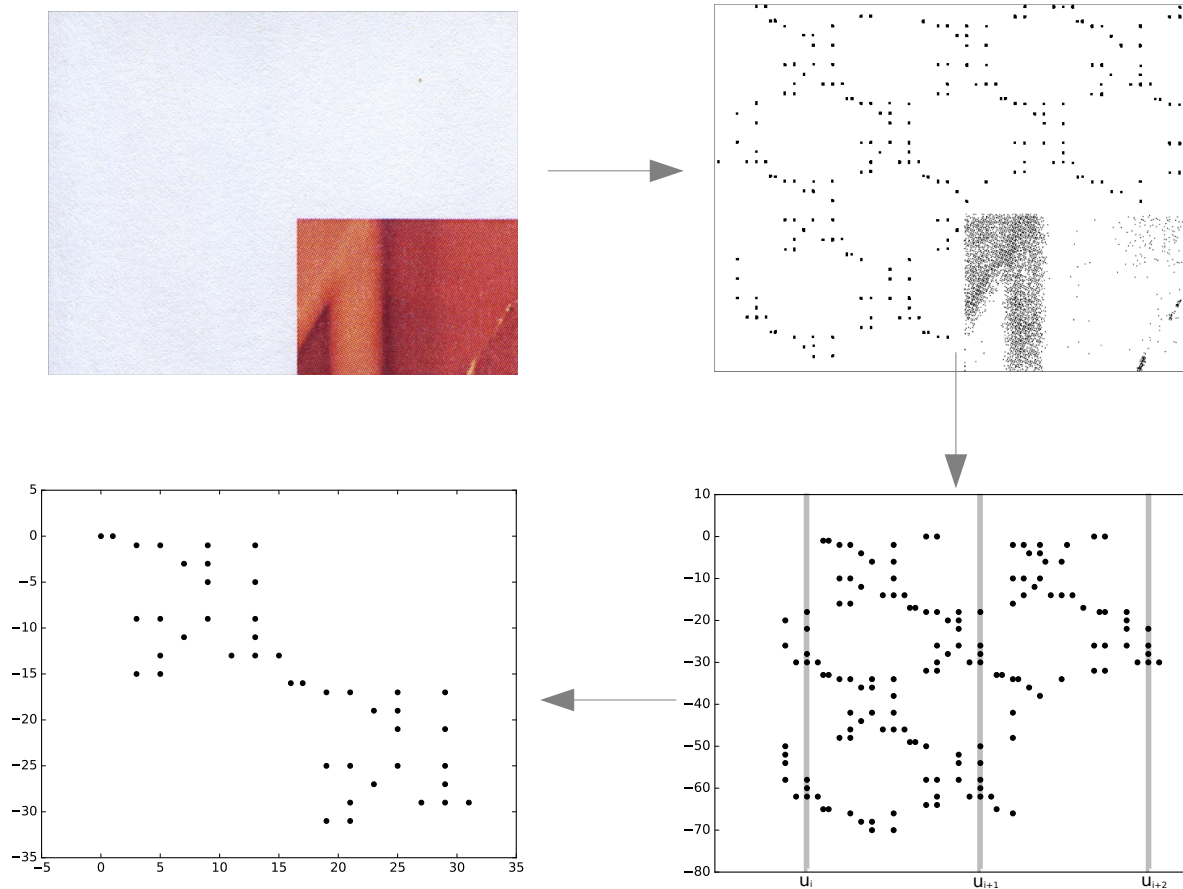
- Official reason, embedded information and structure unknown
- Several manufacturers contacted
- Printer manufacturer (document from 2010): Please contact the following institutions:
 - Central Bank Counterfeit Deterrence Group (CBCDG)
 - German Federal Bank
- CBCDG: „Not a CBCDG product“



Tracking Dots - Extraction

- Developed own extraction algorithm
- Scanned printout → digital tracking dot matrix

Tracking Dots - Extraction



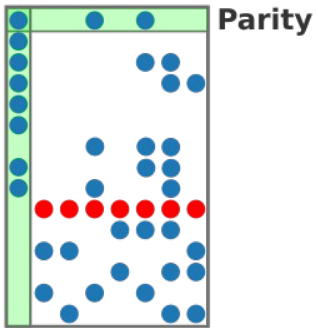
Tracking Dots - Dataset

- 1286 prints with images and text from
 - 141 colour laser printers a 106 models by 18 manufacturers
 - Own dataset and from DFKI
- Extracted all tracking dot pattern

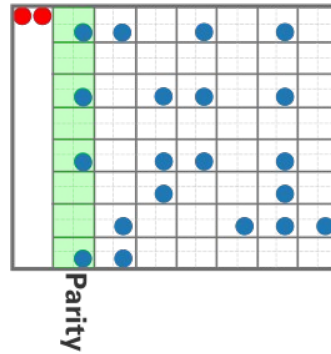
Tracking Dots – Patterns

- 5 different patterns found in dataset
- 4 Pattern structure decoded (marker, information bits, error detection bits, ...)
- 2 Pattern fully decoded (information)
- Nearly all colour laser printers affected

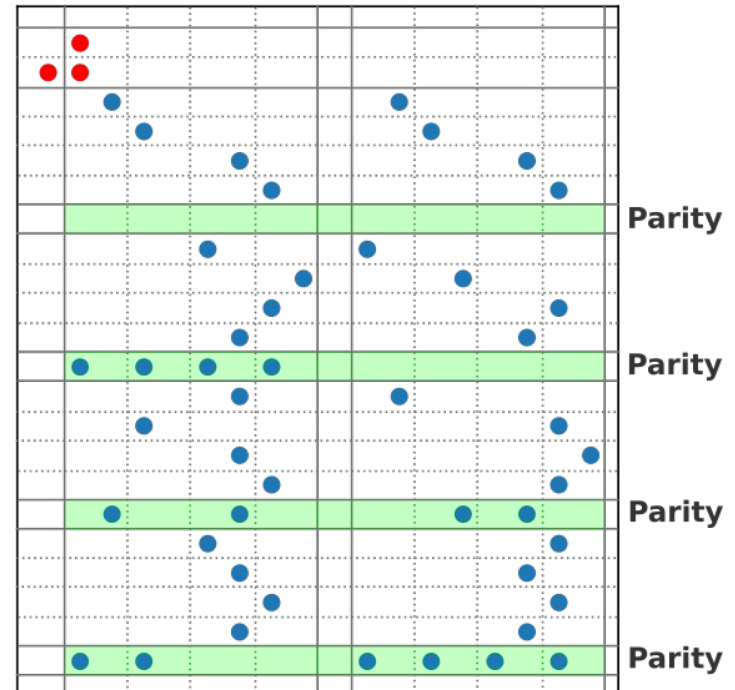
Pattern 4



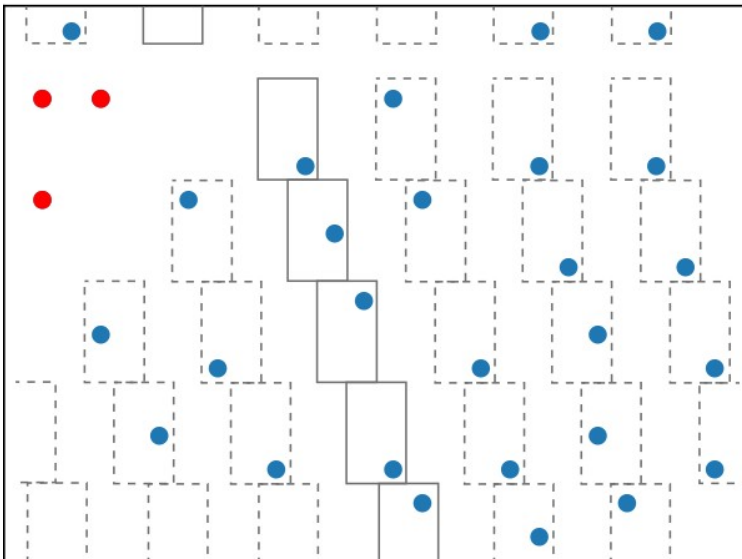
Pattern 1



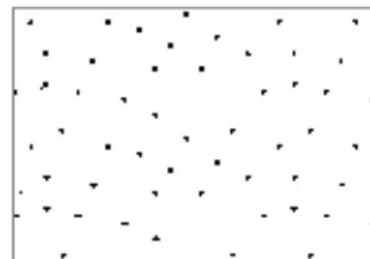
Pattern 2



Pattern 3



Pattern 5



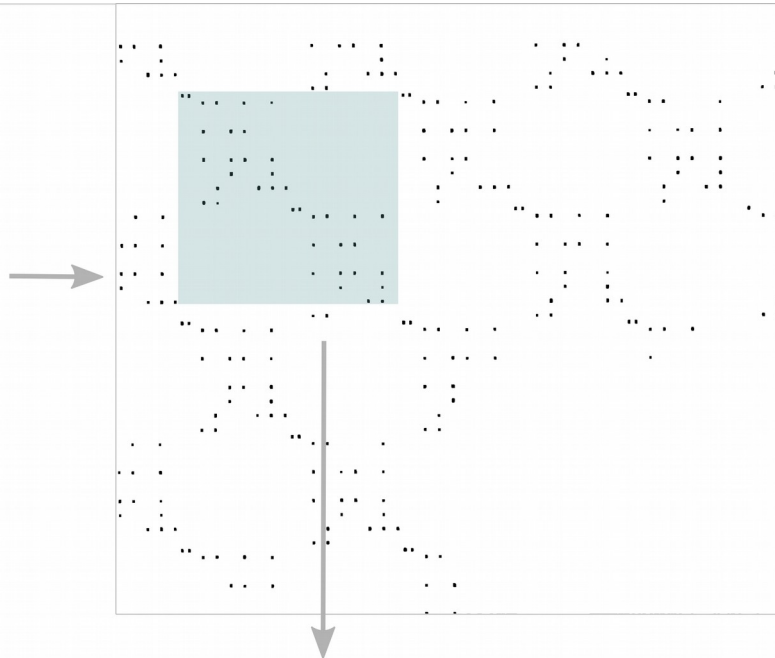
Tracking Dots – Patterns

Pattern	Manufacturer
1	Lanier, NRG, Ricoh, Savin
2	HP, Kyocera, Lexmark, Okidata, Ricoh
3	Epson, Konica Minolta
4	Dell, Epson, Xerox
5	Canon

Samsung, Tektronix and Brother not affected (only small quantity in dataset)

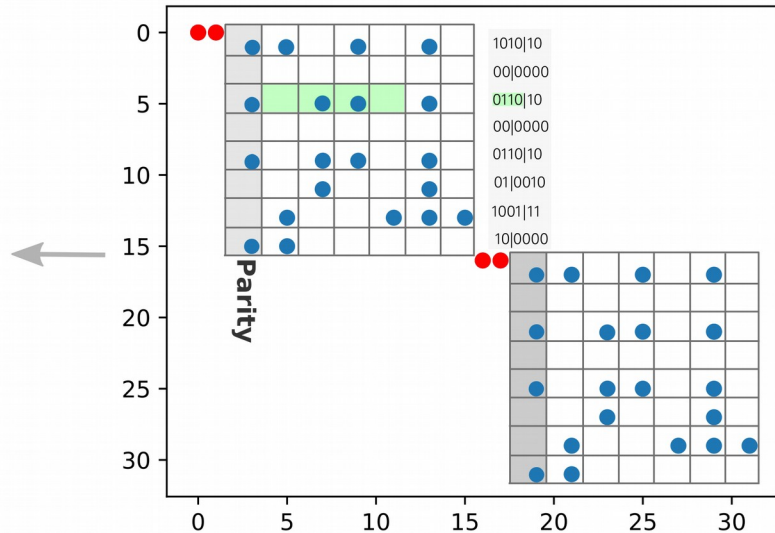


Pattern 1

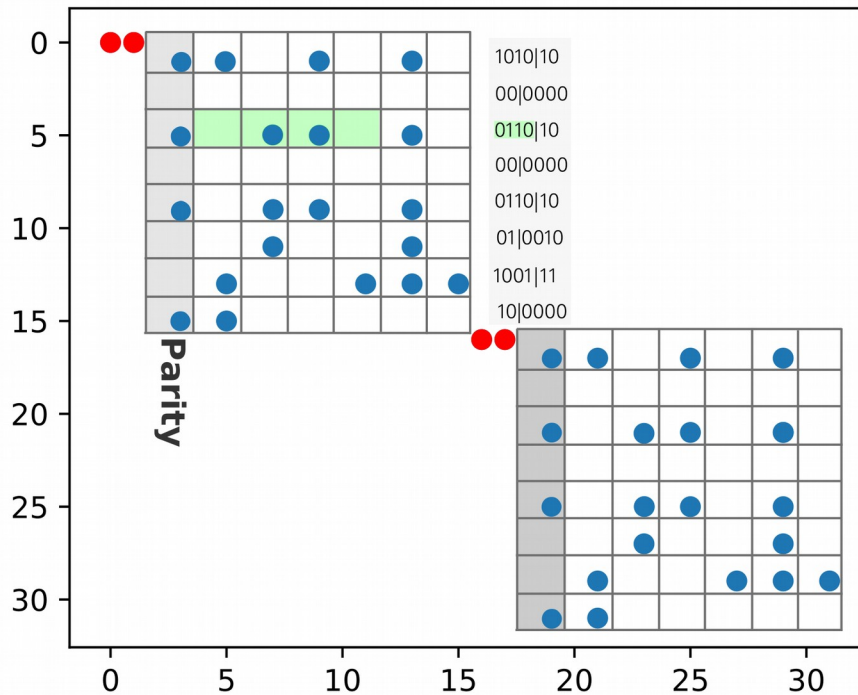


0000	→	0
0111	→	7
1001	→	9
0100	→	4
1001	→	9
0110	→	6
0000	→	0
0001	→	1
0110	→	6
0000	→	0
0001	→	1
0101	→	5

Seriennummer des Druckers:
W794P601601

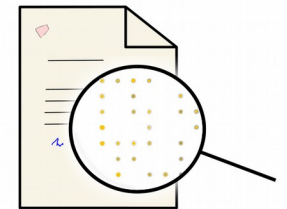


Pattern 1



0000	→	0
0111	→	7
1001	→	9
0100	→	4
1001	→	9
0110	→	6
0000	→	0
0001	→	1
0110	→	6
0000	→	0
0001	→	1
0101	→	5

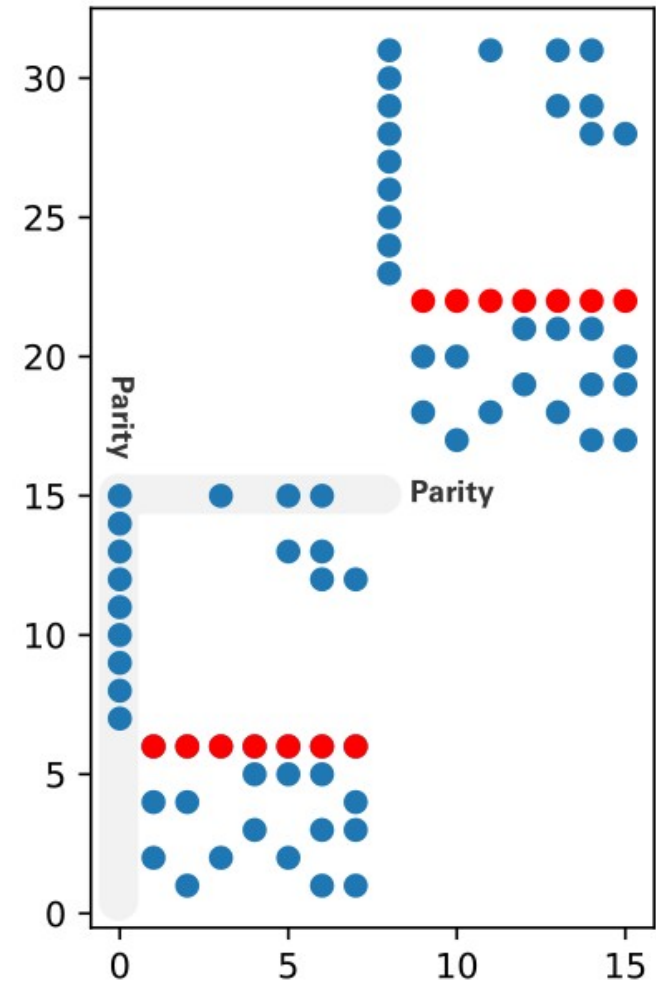
Seriennummer des
Druckers:
W794P601601



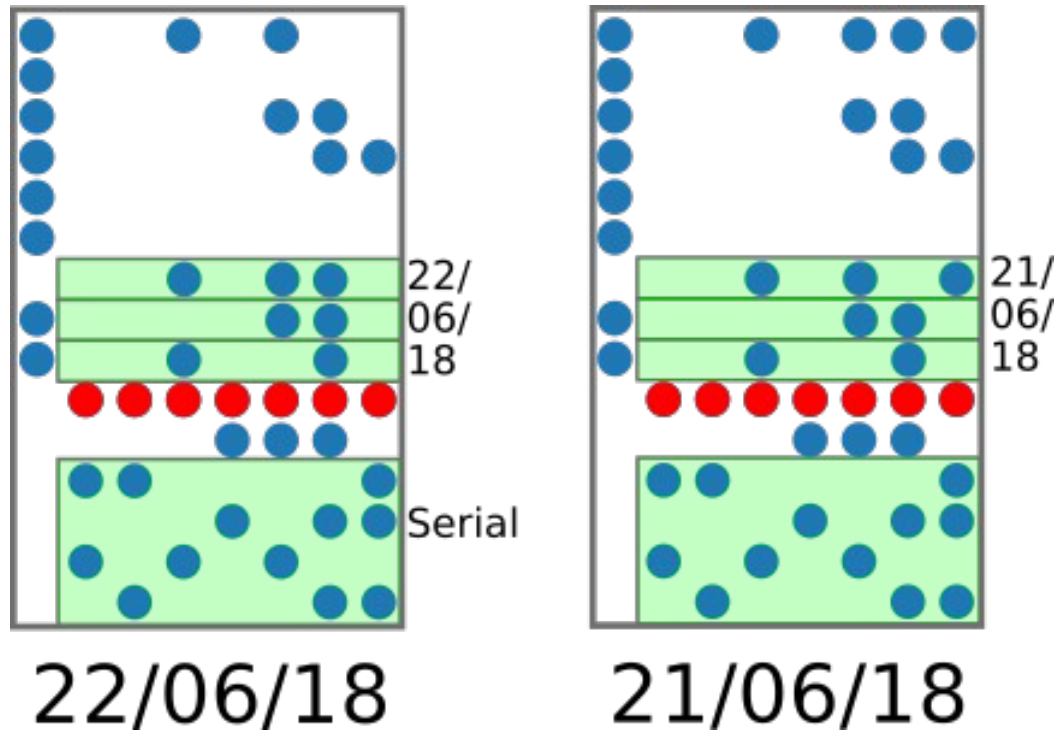
- (7,6,2) even parity code
- Red: marking dots
- *Serial number* as 4 binary bit blocks

Tracking Dots - Pattern 4

- 6 digits of *serial number*,
date and time
- (8,7,2) odd parity code - (15,14,2)
odd parity code
- Repeated in offset



Tracking Dots - Pattern 4



Tracking Dots - Privacy

- No access control: Tracking data can be read by anyone
- Privacy and Security Chair!
 - Prevent arbitrary tracking
 - Developed also anonymization methods



Tracking Dots - Privacy

- Question by Satu Hassi (Verts/ALE) 1 : „Does the Commission believe that the current practices of manufacturers [...] are consistent with relevant Community law on data protection and consumer protection?“
- EU Parliament:
The dots „might violate the right to protection of personal data“
- Time Stamp: 2008



The screenshot shows a page from the European Parliament website. At the top, there is the European Parliament logo and the text "European Parliament". Below this is a navigation bar with language options: BG, ES, CS, DA, DE, ET, EL, EN, FR, GR, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV. The main content area is titled "Parliamentary questions" and shows a question from 23 January 2008 (E-5724/2007) answered by Mr Frattini on behalf of the Commission. The answer text is as follows:

The Commission is not aware of any specific laws either at national or at Community level governing tracking mechanisms in colour laser printers and photocopiers.

In the cases outlined in the Honourable Member's question, the information based on tracking printed or copied material does not necessarily include data relating to identified or identifiable individual, i.e. personal data.

To the extent that individuals may be identified through material printed or copied using certain equipment, such processing may give rise to the violation of fundamental human rights, namely the right to privacy and private life. It also might violate the right to protection of personal data.

The protection of privacy is ensured by Article 8 of the Convention of Human Rights and Fundamental Freedoms. The Charter of Fundamental Rights of the European Union, in Article 7, provides for the protection of private and family life, home and communication, and in Article 8, for the protection of personal data.

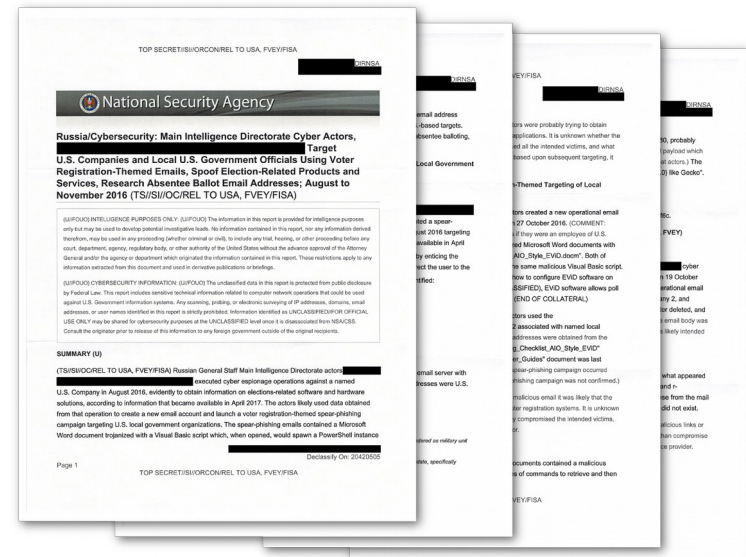
Directive 95/46/EC of Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data ('Data Protection Directive')⁽¹⁾ ensures the protection of personal data and applies regardless of the technology used for the processing of personal data.

⁽¹⁾ OJ L 281, 23.11.1995.

At the bottom right of the page, it says "OJ C 191, 29/07/2008". At the very bottom, there is a footer that says "Last updated: 6 March 2008" and "Legal notice".

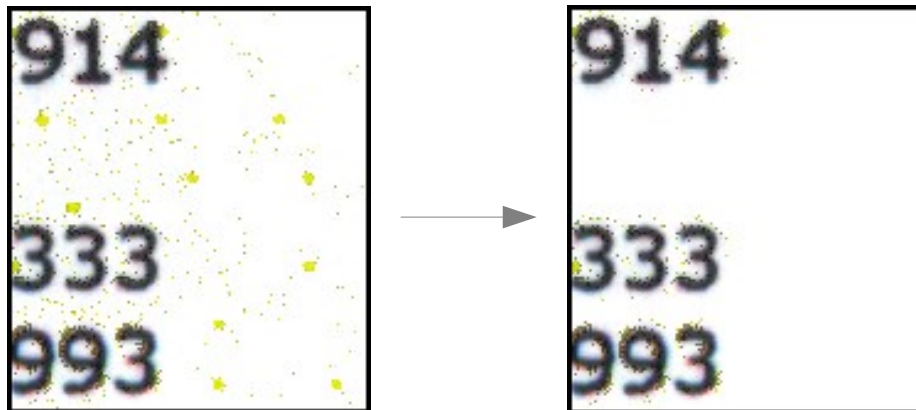
Tracking Dots - Privacy

- Serial Number = unique identification number
- Possible linkability to e.g. credit card number, IP address, ...
- Dots possibly used in court by NSA in 2017 because of leaking secret documents



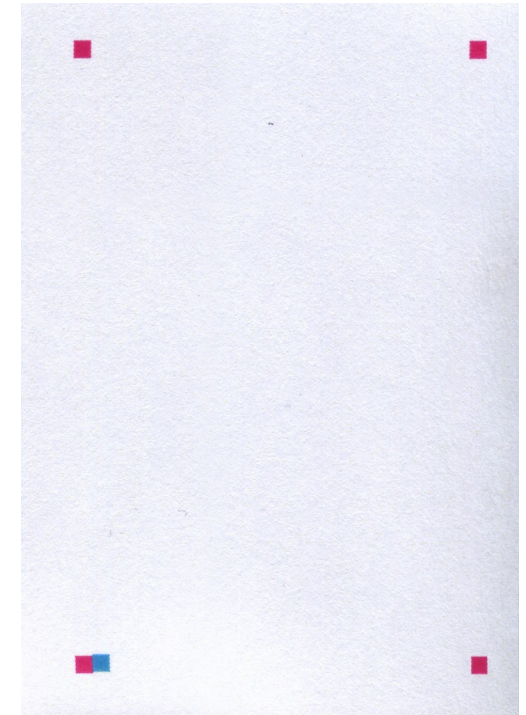
Remove Tracking Dots on Scans

- Mask printed area of the document
- Invert Mask
- Fill white



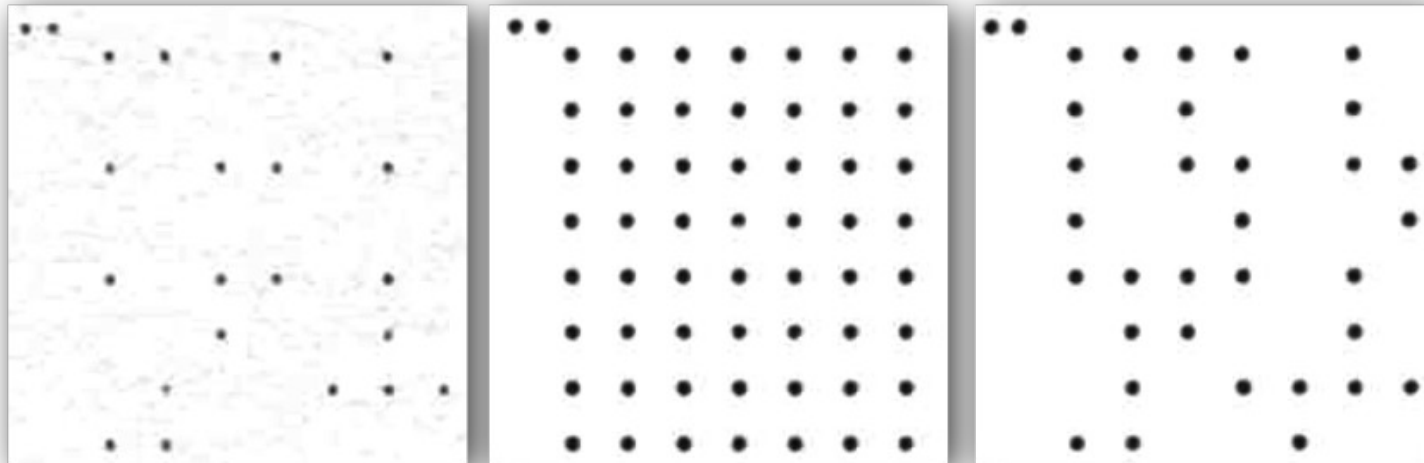
Mask Tracking Dots on Print outs

- Overlaying the tracking dots
- Position of dots must be known
 - Print calibration page with position markers
- Scan it
- Tracking dot extraction
- Measure distance between tracking dots and markers



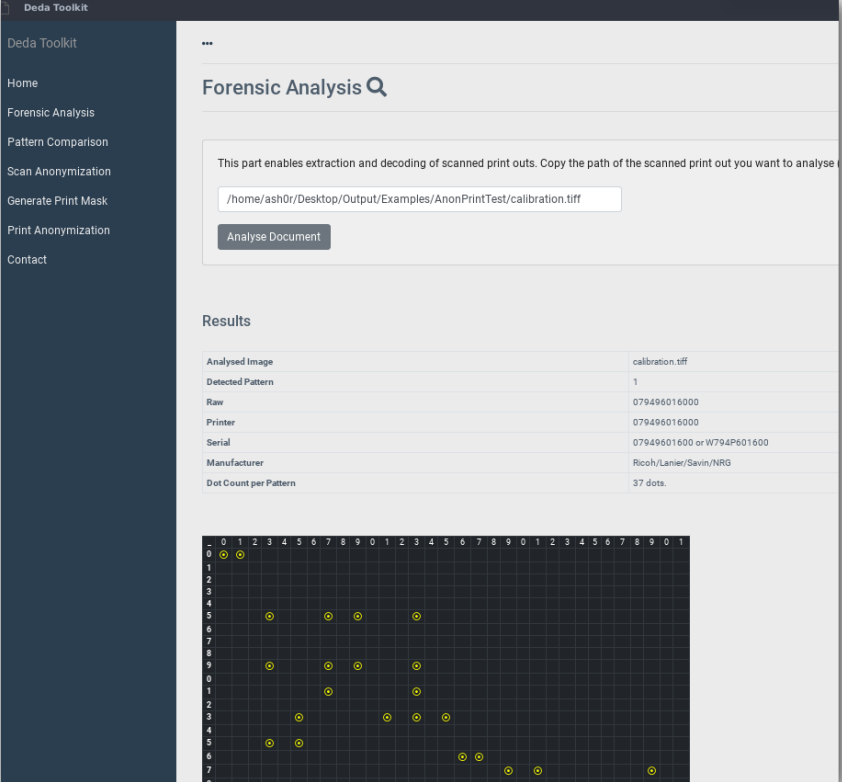
Mask Tracking Dots on Print outs

- Fill extracted tracking dot matrixe with additional dots to destroy encoded information
- Embed tracking dot mask in document with correct distances and overprint existing tracking dots



DEDA

- Toolkit for whole workflow of extracting, decoding and anonymization of tracking dots
- Install Python 3
- `$ pip3 install deda`
- `$ deda_gui`
- `dfd.inf.tu-dresden.de`



The screenshot displays the Deda Toolkit web interface. On the left is a dark sidebar with navigation links: Home, Forensic Analysis, Pattern Comparison, Scan Anonymization, Generate Print Mask, Print Anonymization, and Contact. The main content area is titled 'Forensic Analysis' and contains a text input field with the path `/home/ash0r/Desktop/Output/Examples/AnonPrintTest/calibration.tiff` and an 'Analyse Document' button. Below this is a 'Results' section with a table of extracted data:

Analysed Image	calibration.tiff
Detected Pattern	1
Raw	079496016000
Printer	079496016000
Serial	07949601600 or W794P001600
Manufacturer	Ricoh/Lanier/Savin/IRG
Dot Count per Pattern	37 dots.

At the bottom of the results section is a grid visualization of tracking dots. The grid has 10 columns and 10 rows, with columns and rows indexed from 0 to 9. Yellow circles represent detected dots at various positions on the grid.

Summary

- Tracking dots reusable for project use case
 - If tracking dots detected and decodable – use these
 - Else use intrinsic signatures
- Content still unknown / hidden by manufacturers
- We have
 - Identified codes
 - Boosted data privacy
 - Designed anonymisation method
 - Created and evaluated own toolkit
- In work: Pattern 5, Decoding of Pattern 2 and 3