

# Lernen zu Löschen

Vergessen digitaler Objekte als Gemeinschaftsaufgabe  
von Mensch und KI



Friedrich-Alexander Universität Erlangen-  
Nürnberg

Lehrstuhl Psychologie im Arbeitsleben

Cornelia Niessen

Kyra Göbel

Otto-Friedrich-Universität Bamberg

Professur Angewandte Informatik insb.  
Kognitive Systeme

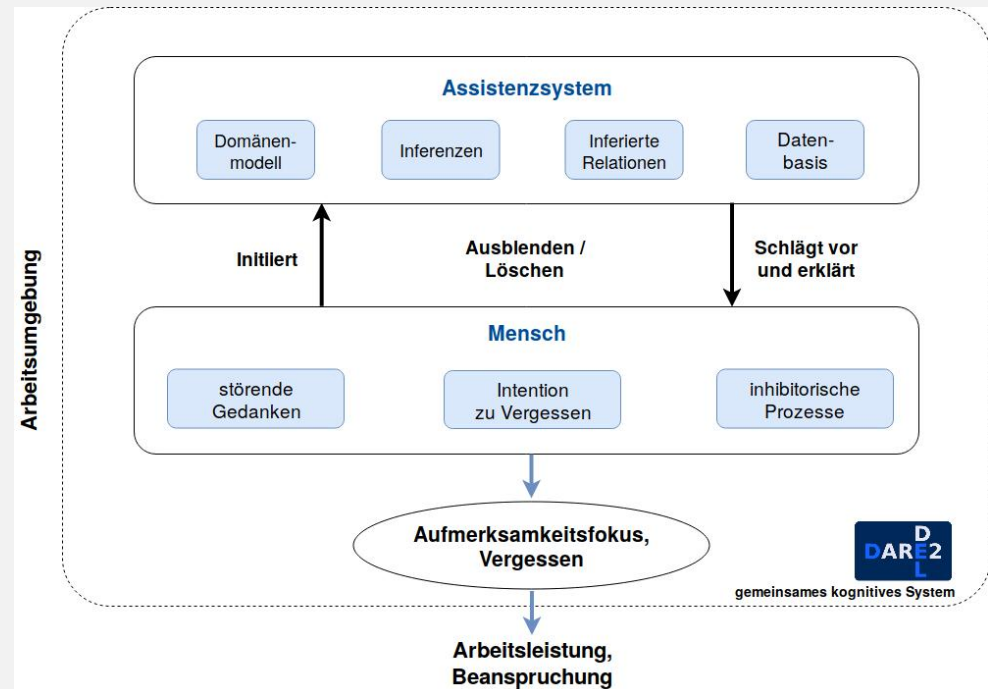
Ute Schmid

N.N

# Ausgangspunkt

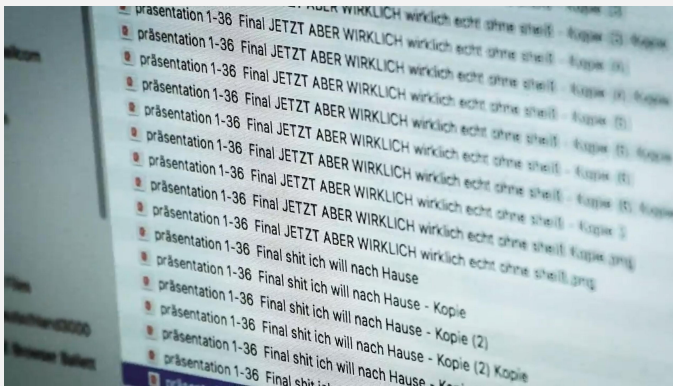
Speicherung von Wissen im Gedächtnis und in externen Speichern

- Interaktives, intelligente Assistenzsystem zur Regulation digitalen Wissens:  
Dare2Del -> Ausblenden, **Löschen**



# Problem

- Unsortiertes Speichern von zu vielen digitalen Objekten
- *Digital hoarding*
- Vermeidung vom Löschen, eher reaktiv, wenn Arbeitsabläufe behindert werden
- Dare2Del: Transparenz und Nachvollziehbarkeit von Löschempfehlungen
- Auseinandersetzung mit Entscheidungsvorschlägen -> Arbeitsleistung, Beanspruchung



## ZEIT ONLINE

[<https://www.zeit.de/serie/digitale-aengste>]  
AUS DER SERIE

### Digitale Ängste

#### Digital Hoarding

## Ist das krank?

10.000 ungelesene Mails sind völlig normal. Nur leider stresst Datenmüll. Und manche horten Fotos, Videos und digitale Dinge sogar krankhaft. Zeit, endlich auszumisten.

Von **Leonie Sontheimer**

3. September 2019, 15:10 Uhr / 134 Kommentare

# Ziele des Projektes Dare2Del

- Gestalten des Löschens irrelevanter Dateien in einem vergessenden Informationssystem
  - Mithilfe von Erklärungen für die Löschvorschläge
- Untersuchung personenbezogener (z.B. Fähigkeit zur Inhibition) und organisationaler Bedingungen (z.B. soziale Normen) der Nutzung und Wirkung des *cognitive companions*

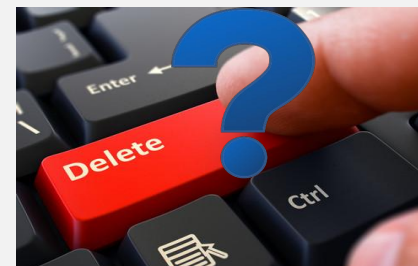
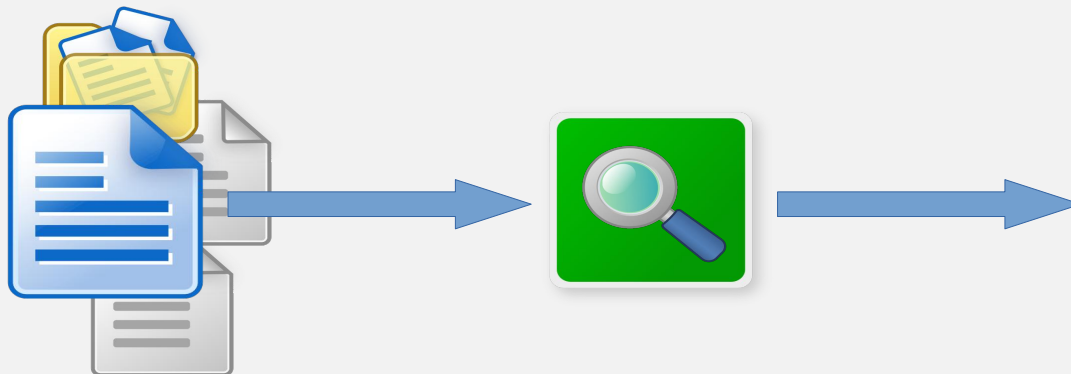


Foto: Colourbox.de / Tashatuvango

# Forschungsbeiträge

## **Informatik:**

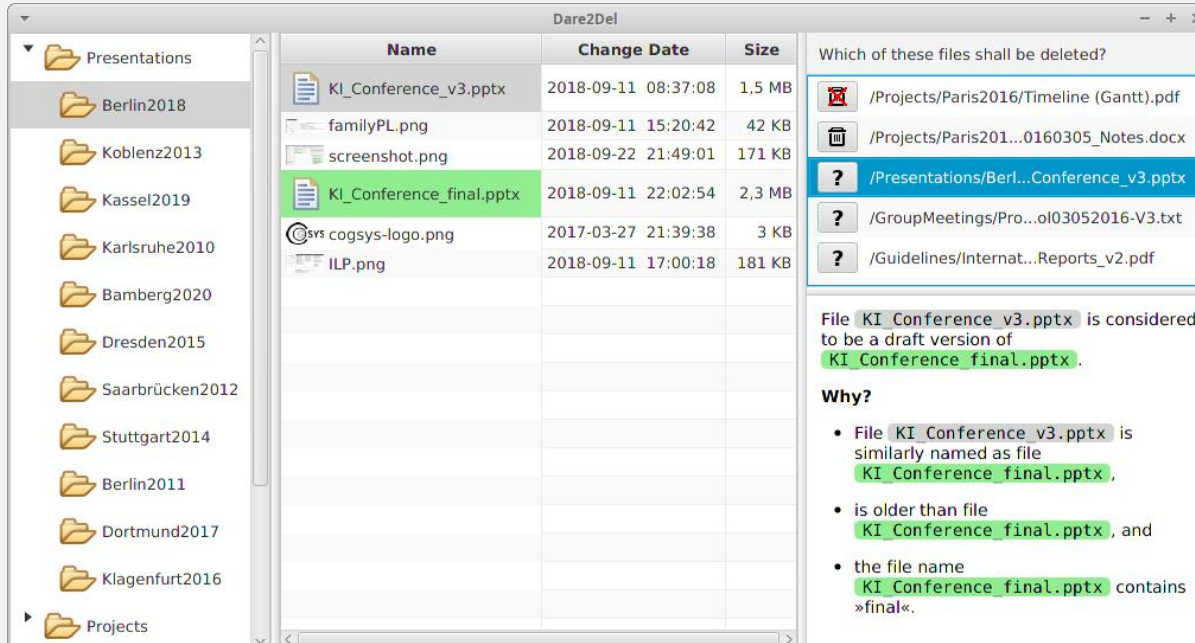
- Erweiterung des auf Induktiver Logischer Programmierung (ILP) basierenden Ansatzes zur Klassifikation digitaler Objekte als potentiell irrelevant
- Erzeugung von Erklärungen
- Berücksichtigung von Konfidenzen
- Korrigierbarkeit durch Inkrementelles/kooperatives Lernen
- Gezielte Auswahl von Objekten die dem Nutzer präsentiert werden

## **Psychologie:**

- Untersuchung des Effekts der aktiven Handlung des Löschens externer Objekte auf das intentionale Vergessen im Gedächtnis (Rolle von Erklärungen)
- Konsequenzen der Löschassistenz für die Arbeitsleistung und Beanspruchung

# Erklärungen

- grundlegend für Nachvollziehbarkeit von KI-Systemen (XAI)
- Kein „one size fits all“: verbale Erklärungen für aktuelle Instanz sowie near-miss examples (vgl. counter factual reasoning)
- Verschieden detailliert (need for cognition, Konsequenz der Entscheidung)



Name	Change Date	Size
KI_Conference_v3.pptx	2018-09-11 08:37:08	1,5 MB
familyPL.png	2018-09-11 15:20:42	42 KB
screenshot.png	2018-09-22 21:49:01	171 KB
KI_Conference_final.pptx	2018-09-11 22:02:54	2,3 MB
cogsys-cogsys-logo.png	2017-03-27 21:39:38	3 KB
ILP.png	2018-09-11 17:00:18	181 KB

Which of these files shall be deleted?

- /Projects/Paris2016/Timeline (Gantt).pdf
- /Projects/Paris201...0160305\_Notes.docx
- /Presentations/Berl...Conference\_v3.pptx
- /GroupMeetings/Pro...o103052016-V3.txt
- /Guidelines/Internat...Reports\_v2.pdf

File **KI\_Conference\_v3.pptx** is considered to be a draft version of **KI\_Conference\_final.pptx**.

**Why?**

- File **KI\_Conference\_v3.pptx** is similarly named as file **KI\_Conference\_final.pptx**,
- is older than file **KI\_Conference\_final.pptx**, and
- the file name **KI\_Conference\_final.pptx** contains >final<.

# Erklärungen

- grundlegend für Nachvollziehbarkeit von KI-Systemen (XAI)
- Kein „one size fits all“: verbale Erklärungen für aktuelle Instanz sowie **near-miss examples** (vgl. counterfactual reasoning)
- Verschieden detailliert (need for cognition, Konsequenz der Entscheidung)

```
irrelevant(itemA) :- similar_name(itemA,itemB), older(itemA,itemB),  
name_contains(itemA,final)
```

*„Datei A kann gelöscht werden, weil  
Datei A einen ähnlichen Namen wie Datei B,  
Datei A ist älter als Datei B,  
Datei A hat 'final' im Dateinamen.“*

*„Datei A könnte gelöscht werden, wenn sie älter wäre als Datei B.“*

# Psychologische Auswirkungen von Erklärungen

## Nachvollziehbarkeit von Systementscheidungen

- Vertrauen (Mercado et al., 2016; Miller, 2018; Pu & Chen, 2007; Ribeiro et al., 2016; Wang & Benbasat, 2007)
- Reduzierung von informationaler Unsicherheit (van den Bos, 2009)  
(„warum bekomme ich gerade diese Datei vorgeschlagen“)
  - Unsicherheit wird häufig negativ erlebt (z.B. Wilson, Centerbar, Kermer, & Gilbert, 2005)
  - Grübeln (Berenbaum et al., 2008; Kofta & Sedek, 1999)
  - Entscheidung ist kognitiv nicht vollständig abgeschlossen (Leroy, 2009)
- Negativer Affekt
- Reduzierte Aufmerksamkeit
- Besseres Vergessen der gelöschten Informationen



# Auswirkung der Löschassistenz

- Löschen als Handlung zur Strukturierung der eigenen Aufgabenumgebung
- Experience Sampling Untersuchung: (1) Von der Unterstützung des Löschens könnten besonders diejenigen Personen profitieren, die Schwierigkeiten haben Irrelevantes auszublenden. (2) Ein *cognitive companion* sollte besonders dann hilfreich sein, wenn die unmittelbare Arbeitssituation zum Löschen auffordert (arbeits- und organisatorische Probleme)
  - Unterstützung der Informationssuche durch Löschen -> langfristige Reduktion von Regulationshindernissen bei Arbeitstätigkeiten
  - Bessere Aufgabenleistung, weniger Beanspruchung
- Organisationale Rahmenbedingungen (soziale Normen und organisationale Regularien) sollen das Löschen unterstützen

# Kooperatives Lernen mit wechselseitigen Erklärungen

Dare2Del
— + ×

Presentations

- Berlin2018
- Koblenz2013
- Kassel2019
- Karlsruhe2010
- Bamberg2020
- Dresden2015
- Saarbrücken2012
- Stuttgart2014
- Berlin2011
- Dortmund2017
- Klagenfurt2016
- Projects

Name	Change Date	Size
KI_Conference_v3.pptx	2018-09-11 08:37:08	1,5 MB
familyPL.png	2018-09-11 15:20:42	42 KB
screenshot.png	2018-09-22 21:49:01	171 KB
KI_Conference_final.pptx	2018-09-11 22:02:54	2,3 MB
cogsys-logo.png	2017-03-27 21:39:38	3 KB
ILP.png	2018-09-11 17:00:18	181 KB

Which of these files shall be deleted?

- /Projects/Paris2016/Timeline (Gantt).pdf
- /Projects/Paris201...0160305\_Notes.docx
- /Presentations/Berl...Conference\_v3.pptx
- /GroupMeetings/Pro...ol03052016-V3.txt
- /Guidelines/Internat...Reports\_v2.pdf

File **KI\_Conference\_v3.pptx** is considered to be a draft version of **KI\_Conference\_final.pptx**.

**Why?**

- File **KI\_Conference\_v3.pptx** is similarly named as file **KI\_Conference\_final.pptx**,
- is older than file **KI\_Conference\_final.pptx**, and
- the file name **KI\_Conference\_final.pptx** contains »final«.

# Korrigierbarkeit von Erklärungen

- Irrelevanz ist (überwiegend) subjektiv (keine ground truth)
  - Modifikation von Erklärungen erlaubt Wissen in den Lernprozess einzubringen
  - Übergenerelle Regel: Hinzufügen von Attributen, Relationen, Einschränken von Wertebereichen
    - Reduktion von false alarms
  - Überspezifische Regel: Entfernen von Attributen, Relationen, Erweitern von Wertebereichen
    - Reduktion von misses
- ➔ Constraints bei der Modelladaptation

# Studien zu den Effekten von Erklärungen

## Studie 1a

### Fragestellungen

- Führen Erklärungen eher zur Annahme des Vorschlages von Dare2Del?
- Wenn gelöscht wird, führen Erklärungen dazu, dass die Information (Ort und Inhalt) besser vergessen werden?

**Stichprobe:** 36 Personen

**Design:** Erklärung (mit/ohne)

### Aufgabe:

- Kategorisierung und Speichern von Abschlussarbeiten in der Bibliothek
- Aufforderung durch *cognitive companion* irrelevante, doppelte, veraltete Dateien zu löschen (selbst bearbeitete versus vorhandene Dateien)
- Löschvorschläge mit und ohne Erklärung
- Wiedergabetests: Inhalt und Ort der Dateien

**AV:** Annahme des Vorschlages, Unsicherheit, Affekt und Aufgabenwechsel, Erinnerungsleistung

# Studien zu den Effekten von Erklärungen

## Studie 1b (Replikation)

### Fragestellung:

- Fördern Erklärungen das Vergessen auch von gut gelernten Informationen (hohe Verarbeitungstiefe)?

### Stichprobe: 54 Studierende

- Design: 2 (keine Erklärung vs. Erklärung, within-subject Faktor) x 2 (geringe vs. hohe Verarbeitungstiefe) between-subject Faktor)
- Gleiche Aufgabe wie Studie 1a + Lernphase (Dateinamen durchlesen versus erinnern)

# Effekte des Löschens

## Studie 2a

### Fragestellung:

- Helfen Vorschläge zum Löschen Aufgabenleistung und kognitive Beanspruchung zu verbessern?
- Welche Rolle spielen dabei interindividuelle Unterschiede (Fähigkeit zur Inhibition)?

**Stichprobe:** 158 Probanden

**Design:** Faktor 1: Löschen mit Unterstützung vs. selbstinitiiertes Löschen vs. kein Löschen; Faktor 2: kognitive Fähigkeit, beide between-subject

# Studie 2: Effekte des Löschens

## Aufgabe:

- Gesundheitsbereich: Prüfung von Medikationen -> Suchen von Patientendatei und Regelwerk in einem Ordnerverzeichnis, email schicken

## UV:

### Bedingungen

- „mit Unterstützung“: Vorschlag des *cognitive companion* irrelevante Dateien zu löschen (irrelevant, störend versus nur irrelevant)
- „selbstinitiiertes Löschen“: Instruktion: Löschen ist hilfreich für die Aufgabenbearbeitung, daher sollten irrelevante Dateien gelöscht werden
- „kein Löschen“

Interindividuelle Differenzen: Fähigkeit zu vergessen, Arbeitsgedächtniskapazität

**AV:** Aufgabenleistung (Dauer der Bearbeitung, Fehler), die mentale Beanspruchung, Anzahl der gelöschten Dateien

# Effekte des Löschens

## Studie 2b

### Fragestellung:

- Moderieren soziale Normen in Organisationen den Effekt der Unterstützung durch Dare2Del?

**Stichprobe:** 113 Probanden

### Design:

- Löschen mit Unterstützung vs. selbstinitiiertes Löschen; soziale Normen vs. keine sozialen Normen, beide Faktoren between-subject)
- Soziale Normen: Löschvorschläge mit Erklärung, dass andere Studierende eine vergleichbare Datei in der überwiegenden Anzahl der Fälle auch gelöscht hätten
- In der Bedingung „keine Norm“ fehlt diese Zusatzerklärung

**Aufgabe:** wie 2a



# Effekte des Löschens

## Studie 2c

### Fragestellung:

- Moderieren organisationale Regularien den Effekt der Unterstützung durch Dare2Del?

**Stichprobe:** 113 Probanden

### Design:

- Löschen mit Unterstützung vs. selbstinitiiertes Löschen; soziale Normen vs. keine sozialen Normen, beide Faktoren between-subject)
- In der Bedingung „Speicherbeschränkung“ werden die Probanden instruiert, dass ihr Speicher fast voll ist (zu 90%), in der Bedingung „keine Beschränkung“ bekommen die Probanden keine spezifische Instruktion.

**Aufgabe:** wie 2a

