

# EXTRAHIERUNG VON ANFORDERUNGEN AUS NATÜRLICH-SPRACHLICHEN LASTENHEFTEN

Was erschwert eine KI-basierte Extrahierung?

Daniel Preuß, Dr.-Ing. Jens Pottebaum

Paderborn | 26.11.2021

# Extrahierung von Anforderungen aus natürlich-sprachlichen Lastenheften

## Motivation

- In der Entwicklung komplexer technischer Systeme werden ca. **80% der Anforderungen** in natürlich-sprachlichen Anforderungsdokumenten wie **Lastenheften** (oder auch Interview-Transkripte, narrative Szenarien, ...) dokumentiert [1, 2].
- Die manuelle Extrahierung von Anforderungen aus natürlich-sprachlichen Spezifikationsdokumenten ist zeitaufwändig, fehleranfällig, monoton und führt zu **hohen Kosten** für Unternehmen [3].
- Bei der Entwicklung komplexer technischer Systeme können Anforderungslisten über 700 Anforderungen umfassen, wodurch der Aufwand hoch ist [4]. Darum ist in diesem Anwendungskontext eine **automatisierte Extrahierung** notwendig.

[1] Mich, L.; Franch, M.; Novi Inverardi, P.L.: "Market research for requirements analysis using linguistic tools".

In: *Requirements Engineering*, Jg. 9, Nr. 1. Springer, 2004, S. 40-56.

[2] Neill, C.J.; Laplante, P.A.: "Requirements engineering: The state of the practice". In: *IEEE Software*, Jg. 20, Nr. 6, 2003, S. 40-45.

[3] Ambriola, V.; Gervasi, V.: "On the Systematic Analysis of Natural Language Requirements with C IRCE".

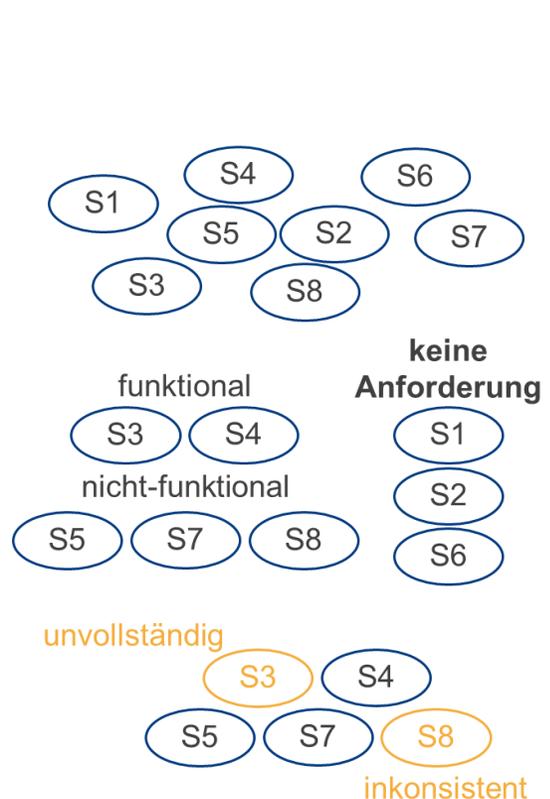
In: *Automated Software Engineering*, Jg. 13, Nr. 1. Kluwer Academic Publishers, 2006, S. 107-167.

[4] Fernandes, J.; Henriques, E.; Silva, A.; Moss, M.A.: "Requirements change in complex technical systems: an empirical study of root causes".

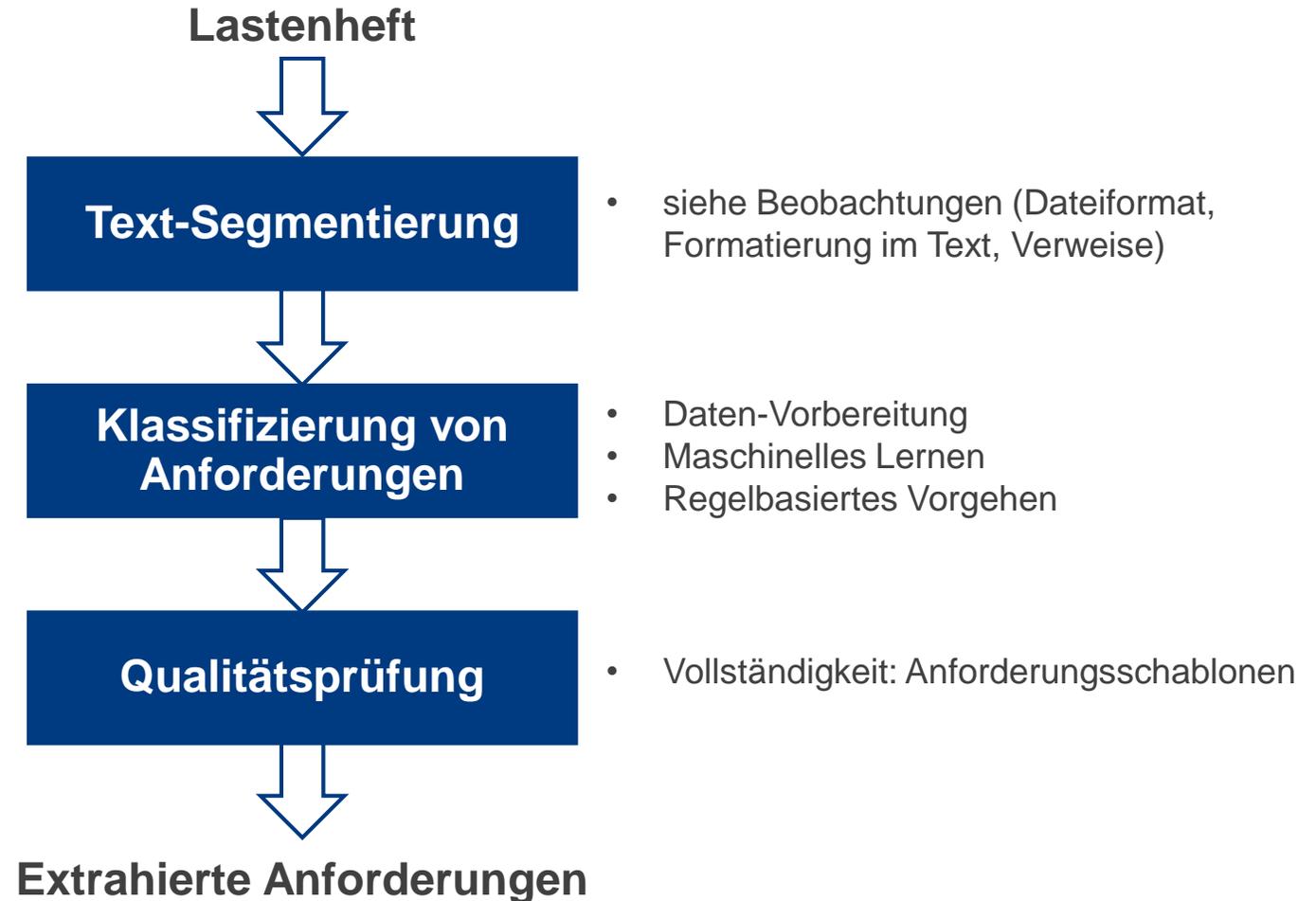
In: *Research in Engineering Design*, Jg. 26, Nr. 1. Springer, London, 2015, S. 37-55.

# Extrahierung von Anforderungen aus natürlich-sprachlichen Lastenheften

## Automatisierter KI-basierter Ansatz



SX: Segment X



# Extrahierung von Anforderungen aus natürlich-sprachlichen Lastenheften

## Beispiel

### Auszüge aus dem Spezifikationsdokument des DUAP-Systems

- A<sub>1</sub>:** „Das DUAP-System soll eine Java-Softwarebasis verwenden.“  
**nicht-funktionale Anforderung**
- A<sub>2</sub>:** „Die Eingabedienste des DUAP-Systems sollen Sondendaten von einzelnen Fahrzeugen empfangen.“  
**funktionale Anforderung**
- A<sub>3</sub>:** „Das DUAP-System muss in der Lage sein, Wetterdaten zu erfassen.“  
**funktionale Anforderung**
- A<sub>4</sub>:** „Die Dokumentenkontrollnummer für dieses Dokument ist in der Fußzeile des Dokuments enthalten.“  
**keine Anforderung**



- **Input:** Natürlich-sprachliche Text-Segmente aus Spezifikationsdokumenten
- **Pre-Processing:** Entfernen von Stopp-Wörtern; POS-Tagging
- **Processing:** Klassifikation der Text-Segmente
- **Output:** als „nicht-funktionale Anforderung“, „funktionale Anforderung“, „keine Anforderung“ ausgezeichnete Text-Segmente

### Beispiel-Input für A<sub>4</sub>:

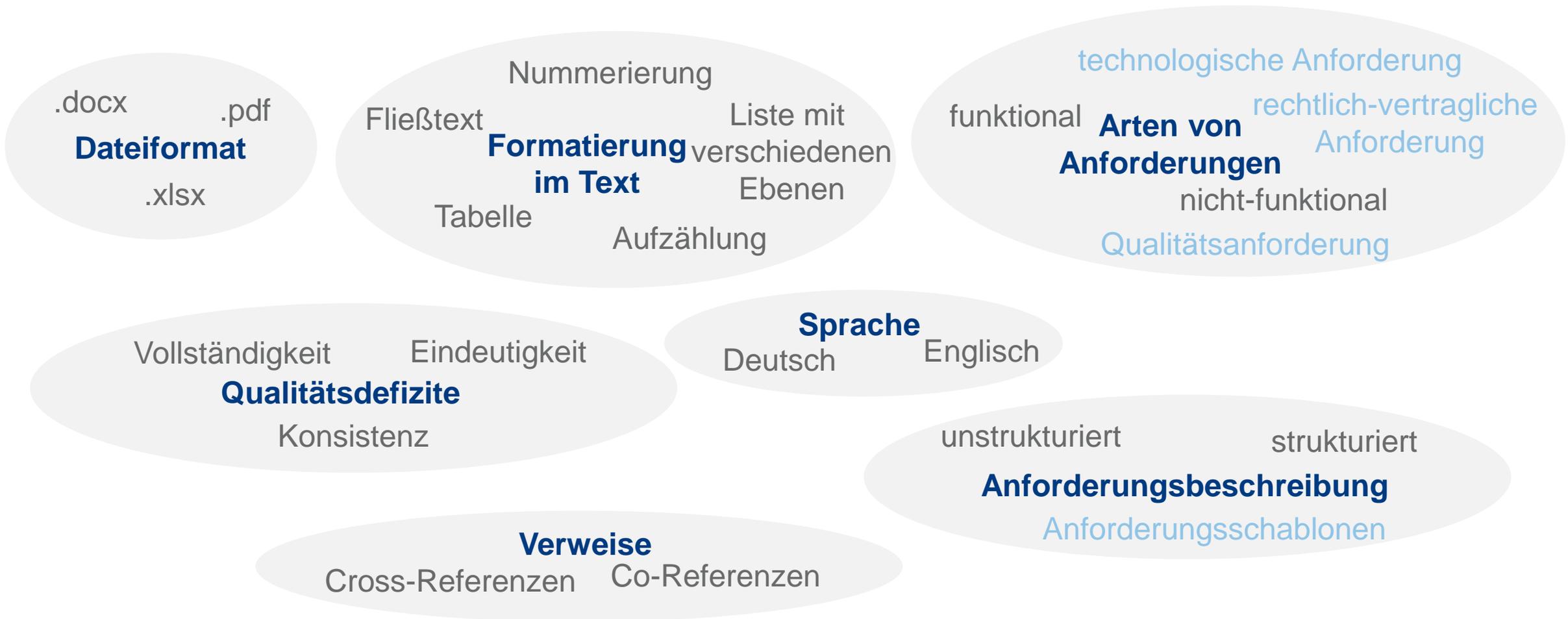
**Wörter:** {Dokumentenkontrollnummer, Dokument, Fußzeile, Dokuments, enthalten}

**POS-Tags:** {NN, NN, NN, NN, VB}      NN: Nomen, Singular; VB: Verb

*DUAP = Data Use Analysis and Processing*

# Extrahierung von Anforderungen aus natürlich-sprachlichen Lastenheften

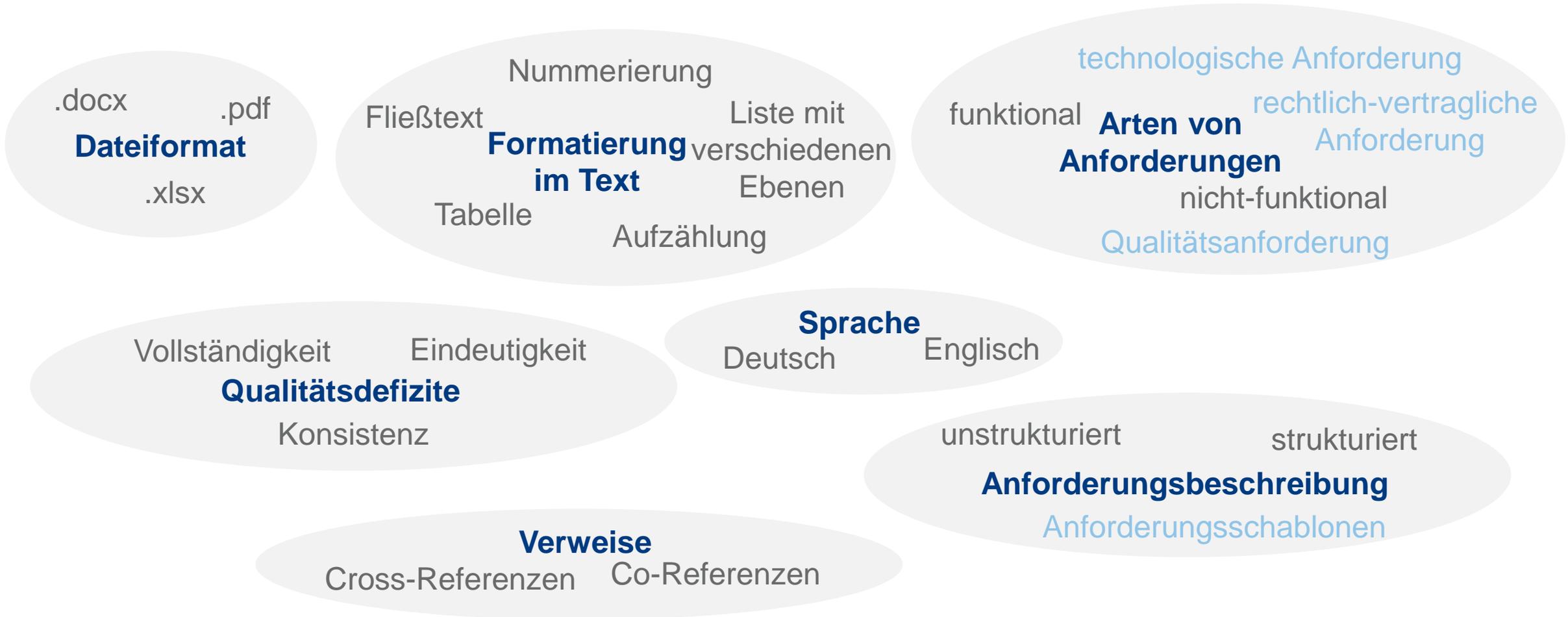
## Beobachtungen: Was erschwert eine KI-basierte Extrahierung?



# Extrahierung von Anforderungen aus natürlich-sprachlichen Lastenheften

## Beobachtungen: Was erschwert eine KI-basierte Extrahierung?

*Welche weiteren Aspekte erschweren die automatisierte Extrahierung von Anforderungen?*





**Vielen Dank für Ihre  
Aufmerksamkeit**

# Kontakt



**M.Sc.  
Daniel Preuß**

+49 5251 / 60 62 59  
daniel.preuss@hni.upb.de

 @HNI\_PB

 @Heinz Nixdorf Institut



**Dr.-Ing.  
Jens Pottebaum**

+49 5251 / 60 62 58  
jens.pottebaum@hni.upb.de

Heinz Nixdorf Institut  
Universität Paderborn

Produktentstehung

Fürstenallee 11  
33102 Paderborn

Telefon: +49 5251 / 60 62 75  
Fax: +49 5251 / 60 62 80  
iris.graessler@hni.upb.de  
www.hni.uni-paderborn.de/pe

**Univ.-Prof. Dr.-Ing. Iris Gräßler**

**HEINZ NIXDORF INSTITUT** | **PE**  
UNIVERSITÄT PADERBORN