

1 **Intelligente Persönliche Assistenten (IPA) mit Voice User**  
2 **Interfaces (VUI) als 'Beteiligte' in häuslicher Alltagsinter-**  
3 **aktion. Welchen Aufschluss geben die Protokolldaten der**  
4 **Assistenzsysteme?**

5 *Stephan Habscheid & Tim Moritz Hector & Christine Hrncal & David*  
6 *Waldecker*

7 **1. Einleitung: Phänomenbereich, Fragestellungen, Charakterisierung**  
8 **der Daten**

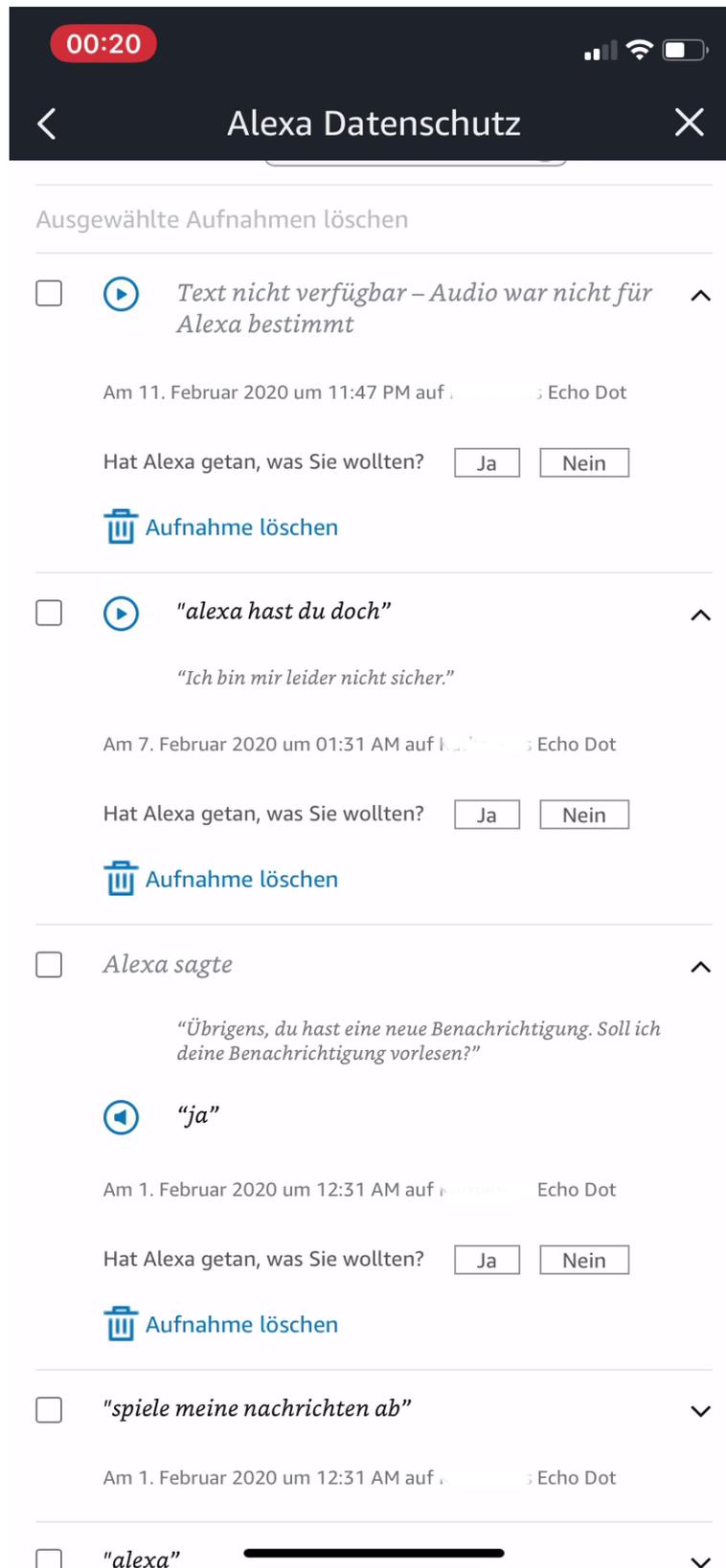
9 Mit der zunehmenden Verbreitung von *Smart Home*-Geräten  
10 halten Programme im Alltag Einzug, die in der Forschung  
11 generisch als „Intelligente Persönliche Assistenten“ (IPA) mit  
12 „Voice User Interfaces“ (VUI) bezeichnet werden (vgl. z.B.  
13 Porcheron et al. 2018). Konzipiert als Infrastrukturen „sozio-  
14 technische[r] Zukünfte“ (vgl. Strüver 2020: 2) sollen, so das  
15 Werbeversprechen, solche Systeme gesprochensprachlichen  
16 Input von Nutzer\_innen erkennen, im Fall einer Adressierung  
17 des Systems internetbasiert verarbeiten und in Verbindung  
18 mit synthetischen akustischen Sprachausgaben adäquat  
19 beantworten bzw. nach Sprachbefehlen elementare auto-  
20 matische Aufgaben (z.B. im Haushalt) erfüllen (vgl. Strüver  
21 2020: 1-10). Dazu werden in technischer Hinsicht Funktionen  
22 von *Speech Processing*, *Natural Language Processing* und  
23 *Information Retrieval* miteinander vereint (vgl. Natale 2020:  
24 5).

25 Zu den am weitesten verbreiteten Systemen dieser Art  
26 gehört „Alexa“ in Verbindung mit den vielfältigen „Echo“-  
27 Geräten des Unternehmens Amazon (vgl. Strüver 2020: 1-10):  
28 So werden beispielsweise *Smart Speaker* mit Hilfe der  
29 entsprechenden *Smartphone-App* über das Internet mit

30 einem Nutzerkonto bei Amazon verknüpft und ein *Wake*  
31 *Word* zur Adressierung des Systems wird zugewiesen.  
32 Anhand verschiedener Farben einer LED-Leuchte am Gerät  
33 ist erkennbar, ob die eingebauten Mikrofone in Betrieb sind  
34 und ob, sofern dies der Fall ist, das System lediglich die  
35 akustische Umgebung nach dem *Wake Word* absucht oder,  
36 nachdem das *Wake Word* technisch erkannt wurde bzw. das  
37 Gerät sich im „Aufmerksamkeitsmodus“ befindet, im Sinn  
38 maschineller Sprachverarbeitung „zuhört“, d.h. den  
39 akustischen Input aufzeichnet und in eine Cloud zur  
40 Verarbeitung weiterleitet. In diesem Zusammenhang haben,  
41 nicht zuletzt auch durch Medienberichte über technische  
42 Fehler, auch Fragen des Datenschutzes ein größeres  
43 öffentliches Interesse gefunden (vgl. Strüver 2020: 1-10;  
44 Wissenschaftliche Dienste des Deutschen Bundestages 2019).

45 Wer nun, sei es im Kontext der Nutzung, sei es in der  
46 Forschung, dem Innovationsversprechen bzw. den kritischen  
47 Fragen nachgehen und besser verstehen will, wie derartige  
48 Systeme in Verbindung mit Plattformen und Infrastrukturen  
49 operieren, wie die Dialogsysteme heute gestaltet sind, wie  
50 Nutzer\_innen die VUI bzw. IPA in ihre soziale Interaktion, in  
51 Alltagspraxis und -diskurs einbinden und wie sie dabei mit  
52 Belangen des Datenschutzes umgehen, wird vielleicht auch  
53 auf die Protokolldaten stoßen, die vom System selbst  
54 aufbereitet und den Nutzer\_innen in der *App* zur Verfügung  
55 gestellt werden. Mit diesen Daten, genauer: der Frage nach  
56 ihrer mehrschichtigen Verwendbarkeit in Alltags- und  
57 Forschungskontexten wollen wir uns im Folgenden näher  
58 beschäftigen. Wir bezeichnen diese Daten als „Protokoll-  
59 daten“, da sie nicht den primären Zweck der Gerätenutzung  
60 darstellen, sondern diesen Gebrauch dokumentieren und  
61 protokollieren. Dementsprechend sind im Fall der „Alexa“-  
62 *App* die Protokolldaten nicht vom Startbildschirm aus  
63 abzurufen, sondern an zwei hierarchisch tieferen Stellen in  
64 der *App*-Architektur einzusehen: Zum einen findet sich eine  
65 Darstellung der Protokolldaten in den „Einstellungen“ unter  
66 „Alexa-Datenschutz“ → „Sprachaufnahmenverlauf  
67 überprüfen“. Eine ähnliche Darstellung findet sich auch über  
68 den Eintrag „Aktivität“ im Menü der *App*. Die ausführlichere  
69 Darstellung unter dem Datenschutz-Menü ermöglicht explizit  
70 die Löschung einiger oder aller Einträge. Auf diese Variante

- 71 wollen wir uns in diesem Beitrag fokussieren. Screenshot 1  
72 zeigt die entsprechende Darstellung der Daten.



73

74

Screenshot 1

75 Die Daten werden schriftlich in Listenform dargestellt, wobei  
76 jeder Eintrag durch ein Antippen detaillierter angezeigt  
77 werden kann. In der eingeklappten Ansicht, in welcher hier  
78 der vierte Eintrag – „spiele meine nachrichten ab“ – zu sehen  
79 ist, können Datum und Uhrzeit der Aufzeichnung und das  
80 Gerät, das die Aufzeichnung erstellt hat, sowie der von  
81 „Alexa“ verstandene Text eingesehen werden. Zugleich kann  
82 jeder Eintrag mittels eines Kästchens am linken Rand  
83 ausgewählt und die Auswahl anschließend gelöscht werden.  
84 Beim Aufklappen des Eintrags werden weitere Aktionen zu  
85 jedem Eintrag angeboten. So lässt sich hier jede von „Alexa“  
86 aufgezeichnete Aufnahme auch mittels des blauen  
87 Pfeilsymbols akustisch abspielen. Außerdem wird die von  
88 dem IPA gesprochene Antwort in Textform wiedergegeben,  
89 sie kann selbst jedoch nicht angehört werden. Bei jedem  
90 Eintrag können die Nutzenden eine Rückmeldung geben, ob  
91 „Alexa“ den Anweisungen Folge geleistet hat. Auch lässt sich  
92 jede Aufnahme mittels des Tippens auf ein blaues Mülleimer-  
93 *Icon* auch ohne Vorauswahl löschen. An dem Screenshot  
94 zeigen sich außerdem einige besondere Ausprägungen von  
95 „IPA-Dialogen“ (wie wir den sprachlichen Umgang von  
96 Nutzenden mit den Geräten und Diensten bezeichnen  
97 wollen). Der erste Eintrag in dieser Liste zeigt eine Aufnahme,  
98 die als „nicht für Alexa bestimmt“ erkannt wurde; sie liegt  
99 nicht als transkribierter Text vor. An dieser Stelle wird für die  
100 Nutzenden unter anderem kontrollierbar, welche  
101 Interaktionen im Haushalt vom IPA aufgezeichnet wurden,  
102 ohne dass der IPA nutzerseitig dazu aufgefordert wurde. Der  
103 zweite Eintrag zeigt eine Interaktion, die zwar durch das  
104 *Wake Word* „Alexa“ ausgelöst wurde, jedoch vom IPA nicht  
105 verarbeitet werden konnte. Der dritte Eintrag wiederum zeigt  
106 eine jener (offenbar selteneren) Interaktionen, die nicht  
107 durch Nutzende, sondern durch den IPA initiiert wurden.  
108 Im Rahmen des Teilprojekts B06 „Un/erbetene Beobachtung  
109 in Interaktion: Intelligente Persönliche Assistenten“ unter der  
110 Leitung von Stephan Habscheid und Dagmar Hoffmann (seit  
111 2020)<sup>1</sup> im Sonderforschungsbereich (SFB) 1187 „Medien der

---

1 Das Vorläuferprojekt in der ersten Förderphase des SFBs (2016-2019) wurde von Wolfgang Ludwig-Mayerhofer geleitet.

112 Kooperation“<sup>2</sup> verfolgen wir als ein Teilziel, zu untersuchen  
113 und mithin zu explorieren, welchen Aufschluss die  
114 Protokoll Daten in verschiedenen Analysefeldern geben  
115 können. Zu diesem Zweck wird ein Protokoll Datenkorpus  
116 erstellt, das derzeit (Stand: Januar 2021) aus 244 IPA-Dialogen  
117 aus drei verschiedenen Haushalten besteht. Am oberen Rand  
118 des Screenshots 1 ist eine Sekundenanzahl auf rotem Grund  
119 zu sehen; diese ist den Umständen der Datenerhebung und  
120 der Einbettung der Daten in die Plattform von Amazon  
121 geschuldet. Die Audiodaten sind zwar über die *App*  
122 abspielbar und löscher, aber nicht ohne Weiteres zu  
123 exportieren. Eine Möglichkeit, die wir als Forschende  
124 gesehen haben, um sie in ihrem Nutzungskontext zu erheben,  
125 bestand in der Bitte an Datenspendende, ein Bildschirmvideo  
126 vom sukzessiven Abspielen der Audio-Daten zu erzeugen.  
127 Diese Videos wurden uns zur Verfügung gestellt. Aus diesen  
128 Videos wurden die Audiodaten extrahiert und gemäß den  
129 Konventionen des gesprächsanalytischen  
130 Transkriptionssystems GAT 2 (vgl. Selting et al. 2009) als  
131 Basistranskript verschriftlicht.<sup>3</sup> Dabei wurden weitere  
132 Angaben erfasst, die für die Auswertung der Aufzeichnungen  
133 relevant sein können, darunter Uhrzeit und Dauer der  
134 Aufnahmen, das verwendete Gerät sowie die Anzahl der auf  
135 der Aufnahme zu hörenden Sprecher\_innen. Zugeordnet  
136 wurde zudem die durch den Sprachassistenten zur weiteren  
137 Prozessierung verschriftete Form der Aufzeichnung, sofern  
138 eine solche vorlag.

139 In der Perspektive des SFBs 1187 sind Daten nicht als  
140 Objekte *sui generis* für die medienwissenschaftliche  
141 Forschung interessant, sondern als Teil von Praxis, in deren  
142 Kontext Daten jeweils kooperativ hervorgebracht bzw.  
143 genutzt werden. In diesem Sinn verweist der Begriff der  
144 ‚Datenpraktiken‘ im Kontext des SFBs darauf, dass mit Daten  
145 umgegangen werden muss, damit sie *als* Daten je situativ  
146 relevant werden (vgl. Abschnitt 3). Praxis wiederum muss sich  
147 in ihrem Vollzug nicht nur an anderen Beteiligten – den *Alter*  
148 des *Ego* in der klassischen Sozialtheorie – ausrichten,

---

2 Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – SFB-Geschäftszeichen Projektnummer 262513311 (SFB 1187 „Medien der Kooperation“).

3 Für Arbeiten an den Transkripten und für weitere Hinweise danken wir Viviane Börner, Franziska Petri und Franziska Niersberger-Gueye.

149 sondern auch an den materiellen Gegebenheiten einer  
150 Situation. An den Protokolldaten wird dies insofern deutlich,  
151 als diese nicht nur die Aufzeichnung aus einem Haushalt auf  
152 einer *App* abrufbar machen, sondern dabei in das komplexe  
153 und globale System der Kommunikationsinfrastruktur und in  
154 die Rechenkapazitäten der jeweiligen Betreiberfirmen und  
155 jener von Drittanbietern, die zusätzliche *Skills* anbieten,  
156 eingebunden sind. Solche Gegebenheiten sind in der  
157 Literatur unterschiedlich theoretisch konzeptualisiert  
158 worden: Während Latours (2007) Programm einer Akteur-  
159 Netzwerk-Theorie mit dem Begriff „Aktanten“ darauf abzielt,  
160 Menschen und Nicht-Menschen in Bezug auf ihre *Agency* auf  
161 die gleiche Stufe zu stellen, zählen ontologisch weniger  
162 radikale bzw. agnostische Theorien aus dem praxis-  
163 theoretischen Umfeld technische und andere Geräte zu den  
164 „materiellen Partizipanden des Tuns“ (Hirschauer 2004), da  
165 sie in der Praxis und für die Praxis „rekrutiert“ werden. In  
166 diesem Sinn ist auch unsere Bezeichnung der IPA als  
167 ‚Beteiligte‘ zu verstehen.

168 Vor diesem Hintergrund fragen wir im vorliegenden  
169 Beitrag danach, welchen Aufschluss die Protokolldaten der  
170 IPA den Nutzer\_innen bzw. den im Rahmen eines  
171 rekonstruktiven Ansatzes forschenden Wissen-  
172 schaftler\_innen darüber geben können,

- 173 • wie die IPA-Systeme durch ihre Verbindung mit
- 174 Plattformen und Infrastrukturen operieren (Abschnitt 3),
- 175 • wie die Dialogsysteme heute gestaltet sind und wie
- 176 Nutzer\_innen in Zwei- und Mehrparteien-
- 177 konstellationen diese in ihre alltägliche soziale Inter-
- 178 aktion einbinden (Abschnitt 4.1),
- 179 • in welche alltagspraktischen bzw. diskursiven Kontexte
- 180 die IPA systemseitig gestellt sind bzw. in der nutzungs-
- 181 seitigen sozialen Interaktion gestellt werden und wie auf
- 182 beiden Ebenen insbesondere Belange des Daten-
- 183 schutzes verhandelt werden (Abschnitt 4.2).

184 Um diese Fragen untersuchen zu können, sind die technisch-  
185 institutionelle und die kulturtheoretische Betrachtung von  
186 Medien, wie sie in der germanistischen Medienlinguistik  
187 weithin üblich sind, um eine elementare, sozialtheoretische

188 und interaktionistische Perspektive zu ergänzen. Hierum soll  
189 es im folgenden Abschnitt (2) gehen.

## 190 2. Medienlinguistische Grundlagen

191 In der Debatte über den Medienbegriff im Kontext der  
192 Germanistischen Linguistik stehen traditionell  
193 ‚Kommunikationsformen‘ im Mittelpunkt: Hierunter versteht  
194 man Strukturbedingungen von Kommunikation und Sprach-  
195 gebrauch, die durch die Verwendung technischer Artefakte  
196 (Medien i.e.S.), z.T. auch durch Medieninstitutionen geprägt  
197 sind (z.B. Infrastrukturen und Plattformen der kommerziellen  
198 IPA-Systeme); neben den an Genre-Konventionen (z.B.  
199 Wetterabfragen) gebundenen empirischen Ausprägungen  
200 können so auch (noch) nicht oder nur im Ausnahmefall, mehr  
201 oder weniger kreativ genutzte kommunikative *Potentiale* (z.B.  
202 Herausforderung oder Verspottung der IPA unter An-  
203 wesenden) ins Blickfeld kommen (vgl. Brock/Schildhauer  
204 2017, dazu und zum Folgenden auch Habscheid 2020). Zum  
205 anderen versteht man unter Medien (i.w.S.) kulturell ver-  
206 festigte Techniken oder Verfahren, von denen technische  
207 Medien ein Bestandteil sein können: Solche „medialen Ver-  
208 fahren“, wie z.B. durch Sprachassistenzsysteme vermittelte  
209 Kommunikation mit digitalen Plattformen, bilden demnach  
210 die materiale und prozedurale Seite des Gebrauchs von  
211 Zeichen, dieser ist zudem eingebettet in konventionelle  
212 (kommunikative) Praktiken (z.B. Wissensrecherche; Internet-  
213 Shopping) und hängt außerdem von der individuellen  
214 Kompetenz der (Zeichen-)Verwender ab (vgl. Schneider 2017:  
215 45).

216 Nicht im Blickfeld liegt in beiden Fällen, wie Medien –  
217 jenseits ihrer Potentiale und der kreativen Beteiligung von  
218 Individuen – als *soziale* Instanzen zustande kommen und auf  
219 der Grundlage ihrer Materialität geprägt werden (vgl. Meiler  
220 2019), dynamisch variieren und sich in der Zeit wandeln  
221 können. Um diese Fragen untersuchen zu können, sind die  
222 technische und die kulturtheoretische Betrachtung von  
223 Medien, wie sie in der germanistischen Medienlinguistik  
224 weithin üblich sind, um eine elementare, sozialtheoretische  
225 Perspektive zu ergänzen. In einer praxeologischen  
226 Perspektive, wie sie im SFB 1187 „Medien der Kooperation“

227 an der Universität Siegen verfolgt wird, werden so Medien  
228 verstanden als „kooperativ erarbeitete Kooperations-  
229 bedingungen“ oder, kurz gesagt, als „Medien der  
230 Kooperation“ (Schüttpelz 2016: 5). Diese Kooperation wird  
231 als den Zeichen, deren materialen und prozeduralen  
232 Strukturbedingungen und ggf. konventionellen Nutzungs-  
233 weisen, einschließlich der Herausbildung und Veränderung  
234 von Symbolsystemen, *logisch* vorgelagert aufgefasst (vgl.  
235 Meyer/Ayaß 2012: 14f.). Auch *Daten*, wie sie im vorliegenden  
236 Beitrag fokussiert werden, sind nur in Zusammenhängen  
237 kooperativer sozialer Praxis zu verstehen (vgl. Abschnitt 3).

238 Ein solcher sozialtheoretischer Ansatz, der den Begriff der  
239 Praxis zum Dreh- und Angelpunkt der Theoriebildung macht,  
240 fordert kulturalistische Medientheorien und das Konzept der  
241 Kommunikationsformen gleichermaßen heraus. Diese  
242 Position, wie sie in der konversationsanalytischen Tradition  
243 von Charles Goodwin (2018) bzw. im Grenzbereich von  
244 Konversationsanalyse und Medientheorie durch Erhard  
245 Schüttpelz und Christian Meyer (2017) entwickelt wurde,  
246 blendet kulturelle Verfestigungen und Übereinkünfte (wie  
247 Techniken, kommunikative Gattungen oder Symbole) keines-  
248 wegs aus, geht jedoch auf einer elementareren Ebene der  
249 Ontologie nicht von diesen aus, sondern stellt stattdessen das  
250 Konzept der wechselseitigen ‚Praxis‘ (Schüttpelz/Meyer) bzw.  
251 der ‚Co-operative Action‘<sup>4</sup> (Goodwin) in den Mittelpunkt:  
252 Beteiligte an der Herstellung von Sinn verfertigen demnach  
253 wechselseitig Abläufe, indem sie jeweils die von ihren  
254 Vorgängern ins Spiel gebrachten, zeichenförmig geprägten  
255 materiellen Ressourcen (nicht zuletzt indexikalische  
256 Verweise)<sup>5</sup> partiell aufgreifen und in der Verfolgung ihrer  
257 Ziele transformierend wiederverwenden. Auf dieser Grund-  
258 lage können sich mit dem Effekt einer Vereinfachung der  
259 Verständigung Konventionen herausbilden und verändern,  
260 Medien selbst – ohne die Kommunikation nicht denkbar ist –  
261 werden jedoch elementarer gefasst (wir kommen darauf  
262 zurück). Theoriebildungen, die Medien konstitutiv an  
263 Konventionen binden, greifen insoweit zu kurz.

---

4 Wobei der Bindestrich den operativen, inkrementellen Aspekt der durch  
situieren Zeichengebrauch vermittelten kooperativen Verständigung markiert.

5 Für eine praxeologische Beschreibung verschiedener sprachlicher  
Zeichentypen, die den Status von Symbolen und Repräsentation relativiert, vgl.  
Meiler 2019: 68-72.

264 Auch Brock und Schildhauer (2017) gehen von der Über-  
265 legung aus, dass Kommunikationsformen aufgrund ihres auch  
266 potentiellen Charakters stets offen sind für unterschiedlichste  
267 kommunikativ-funktionale Nutzungen, die sich auf längere  
268 Sicht zu Genres verfestigen *können*.<sup>6</sup> Dieser Aspekt mag  
269 kulturwissenschaftlich unbedeutend erscheinen, insofern das  
270 *Potential* einer Kommunikationsform nicht sinnvoll zu  
271 ergründen ist: Der Kreativität der Nutzer\_innen und der  
272 Eigendynamik interaktionaler Ordnungsbildungen in  
273 situierten Kommunikationsprozessen sind keine Grenzen  
274 gesetzt. Gleichwohl lassen sich aus dem Aspekt der  
275 Potentialität von Kommunikationsformen zwei wichtige  
276 Einsichten ableiten: Zum einen betonen Brock und  
277 Schildhauer, dass im Blick auf die konkrete Nutzung von  
278 Kommunikationsformen im Rahmen von Genres zwischen  
279 *notwendigen* und nur *typischen* Eigenschaften systematisch  
280 unterschieden werden sollte. Zum anderen beugen Brock und  
281 Schildhauer mit ihrer Betonung des Potentials und der  
282 Offenheit von Kommunikationsformen einem  
283 konventionalistisch verengten Blick auf kommunikative und  
284 sprachliche Praxis vor, sei es in Bezug auf Genres, sei es  
285 hinsichtlich der Medien / Kommunikationsformen. Hier  
286 wiederum kann eine praxeologische Perspektive unmittelbar  
287 anschließen.

288 Interaktionale, soziale und kognitive Prozesse liegen nach  
289 Goodwin (vgl. 2018: 1) vielfältigsten Dimensionen und  
290 Aspekten des (in vielerlei Hinsicht einzigartigen) mensch-  
291 lichen (Zusammen-)Lebens zugrunde (vgl. zum Folgenden  
292 auch Schüttpelz/Meyer 2018): Dazu zählen gleichsam klein-  
293 teilige Aspekte wie die multimodale und sequenzielle Inter-  
294 aktion, aber auch Sequenzen und Schichten von Handlungen  
295 und letztlich die funktionale und historische Aus-  
296 differenzierung von Gesellschaften. Goodwin zeigt auf, wie  
297 neue Handlungen systematisch ko-operativ hervorgebracht  
298 werden, also jeweils zustande kommen und verständlich  
299 werden durch sinnhaft strukturierende und geschichtete  
300 Operationen über dem, was andere an Ressourcen und  
301 Lösungen zuvor geschaffen haben (vgl. Goodwin 2018: 431;  
302 vgl. auch Schüttpelz/Meyer 2018: 179f.). Auf diese Weise

---

6 Diese sind selbst vielfach (wenn nicht immer) polyfunktional (vgl. Brock/Schildhauer 2017, 20f.; vgl. dazu auch Hausendorf u.a. 2017, Kap. 9).

303 können nach Goodwin Erträge menschlichen Handelns  
304 transformiert, akkumuliert und tradiert werden, und  
305 Akteur\_innen können nur vor diesem Hintergrund ihre  
306 situative bzw. situationsübergreifende Handlungsmacht  
307 erlangen (vgl. Goodwin 2018: 440).<sup>7</sup>

308 Dem Konzept „co-operative action“ bei Goodwin  
309 entspricht – bei aller Kritik an der Theoriearchitektur und  
310 Begriffsbildung im Einzelnen (vgl. Schüttpelz/Meyer 2018) –  
311 bei Schüttpelz und Meyer in etwa der Begriff der „Praxis“,  
312 hier verstanden als „das in einer wechselseitigen Verfertigung  
313 befindliche Geschehen“ (Schüttpelz/Meyer 2017: 158). Im  
314 Rahmen dieser Theorie ist die Praxis „allen anderen sozialen  
315 Größen vorzuordnen“: „Kooperation, Interaktion, Praktiken,  
316 Handlungen, Routinen, Techniken, technische Medien  
317 werden ‚in der Praxis‘ hervorgebracht“ – und das heißt eben:  
318 „in einem sich in wechselseitiger Verfertigung befindenden  
319 Geschehen“ (Schüttpelz/Meyer 2017: 159). Demnach können  
320 durch Praktiken und Handlungen Routinen und Techniken  
321 ausgebildet werden, die jedoch nicht mit den Praktiken zur  
322 Deckung kommen, vielmehr beruhen Praktiken auch auf  
323 einer „wechselseitigen Improvisation“ (Schüttpelz/Meyer  
324 2017: 156).

325 Wichtig in unserem Kontext erscheint, dass das grund-  
326 legende Prinzip der wechselseitigen Verfertigung – wie  
327 übrigens auch bei Goodwin – nicht nur für die Inhalte,  
328 sondern auch für die Mittel der Verständigung gilt, also auch  
329 für an Medien gebundene Zeichen (um an Schneider 2017  
330 anzuknüpfen). Auch diese Mittel liegen nicht einfach bereits  
331 vor, sondern sie müssen stets aufs Neue situativ und  
332 kooperativ hervorgebracht werden. Vor diesem Hintergrund  
333 werden, wie bereits erwähnt, bei Schüttpelz Medien all-  
334 gemein gefasst als „kooperativ erarbeitete Kooperations-  
335 bedingungen“ (Schüttpelz 2016: 5). Nach Goodwin (vgl. 2018:  
336 445) gehört es zu den Grundprinzipien von Praxis, dass sie  
337 semiotisch „opportunistisch“ und „gefräßig“ ist: Alle Arten  
338 von wahrnehmbaren Materialien können zur lokalen  
339 Konstruktion von Handlungen einbezogen und damit zu  
340 sinnhaft strukturierenden und strukturierten Medien werden

---

7 Auf einen ähnlich umfassenden Gegenstandsbereich richtet sich die *Nexus Analysis* von Scollon/Scollon (2004) (Hinweis J. Androutsopoulos). Eine vergleichende Befassung mit diesem Ansatz ist im Rahmen des vorliegenden Beitrags nicht möglich.

341 (vgl. Goodwin 2018: 445), einschließlich aller Elemente der  
342 Systemebenen von Sprache in der Interaktion (vgl. auch  
343 Schüttpelz/Meyer 2018: 179f., 182).

344 An dieser Stelle muss betont werden, dass eine derartige  
345 Kooperation *Wechselseitigkeit* voraussetzt, aber nicht  
346 unbedingt *Gemeinsamkeit* im Sinne bereits geteilter Ziele,  
347 Werte, Bräuche, Zeichensysteme usw. Schüttpelz und Meyer  
348 (vgl. 2017: 6) verweisen an anderer Stelle auf die in der  
349 englischsprachigen Forschungsliteratur weithin übliche  
350 Unterscheidung von *mutual* („wechselseitig“), wie in *mutual*  
351 *constitution, assistance, repair* usw., und *common* oder auch  
352 *joint* oder auch *shared* („gemeinsam“), (wie in *common goals,*  
353 *means, actions* usw.). Derartige Gemeinsamkeiten stellen das  
354 Ergebnis und eine Erleichterung von Kooperation dar, nicht  
355 jedoch deren notwendige Voraussetzung.

356 Damit positioniert sich der Ansatz nicht zuletzt auch in  
357 einer sozialtheoretischen Tradition. Wenn Menschen im  
358 Alltag miteinander handeln, stellt die bewusste, rational  
359 reflektierte Verständigung über Ziele, Mittel usw. stets nur  
360 einen kleinen Ausschnitt – sozusagen die Spitze des Eisbergs  
361 (vgl. Garfinkel 2012: 56) – dessen dar, was für die  
362 Verständigung tatsächlich relevant ist. Weite Teile dessen,  
363 was in der Situation relevant ist, müssen als fraglos gegeben  
364 erachtet werden (Garfinkel 2012: 56 f.). Mit anderen Worten:  
365 Wann immer in der Kommunikation ein Konsens erreicht  
366 wird, muss, wie Meyer und Ayaß es formulieren, eine  
367 „außer-konsensuale“ — man könnte auch sagen: vor-  
368 vertragliche — „Grundlage“ bereits gegeben sein, „aus  
369 welcher der Konsens überhaupt erst entstehen und  
370 begründet werden kann“ (Meyer/Ayaß 2012: 14; vgl. dazu  
371 auch Durkheim 1992). Umgekehrt ist in der Perspektive der  
372 Praxistheorie aber auch eine Begründung von Kooperation  
373 durch gesellschaftlich-kulturelle Konventionen für sich  
374 genommen unbefriedigend. So ging etwa Harold Garfinkel in  
375 kritischer Distanz zur Theorie seines Lehrers Talcott Parsons  
376 davon aus,

377 „dass Normen keine externen, abstrakten Leitlinien sein  
378 können, sondern von den Akteuren selbst auf der Grundlage  
379 einer verkörperten und reflexiven Sozialität lokal  
380 hervorgebracht, verwaltet und situationsspezifisch im Hier

381 und Jetzt angewendet werden müssen.“ (Meyer/Ayaß 2012:  
382 14).

383 An die zentrale Stelle der Sozialtheorie tritt damit ein Begriff  
384 von ‚Praxis‘ im Sinn von Schüttpelz und Meyer bzw. von ‚Ko-  
385 operativität‘ im Sinn Goodwins.

### 386 3. Protokolldaten, Infrastrukturen, Plattformen

#### 387 3.1 ‚Daten‘ in praxeologischer Perspektive

388 Wenn nun in der Theorie der Praxis von Schüttpelz und  
389 Meyer Medien nicht als feste Rahmen für kommunikative  
390 Formen gelten, sondern selbst als Teil dieser Formen  
391 interaktiv hervorgebracht werden müssen, kann man sich  
392 fragen, was dies für die Betrachtung von Protokolldaten und  
393 überhaupt für ein entsprechendes sozialtheoretisches und  
394 medienlinguistisches Verständnis von *Daten* heißt. Daten sind  
395 aus dieser Perspektive nichts rein Gegebenes, worauf die  
396 Etymologie des Begriffs schließen ließe (lat.: *dare* = geben),  
397 sondern zum einen etwas *Hergestelltes*, zum anderen etwas,  
398 das situativ als Datum *relevant* gemacht werden muss, um als  
399 solches Verwendung finden zu können.

400 Die Produktion von Daten wurde sozialwissenschaftlich  
401 vor allem in der Wissenschaftssoziologie untersucht (vgl.  
402 Pickering 1993) – wegweisend waren Studien Bruno Latours,  
403 beispielsweise zur Hormonforschung (vgl. Latour/Woolgar  
404 1986). Daten, so Latours These, liegen nicht einfach vor,  
405 sondern sind als Teil eines Netzwerks aus Personen, Geräten,  
406 Institutionen, Orten und Gelegenheiten aus diesem heraus  
407 entstanden und verständlich.<sup>8</sup> In einem anderen Kontext  
408 wären sie entweder unverständlich oder würden ihre  
409 Bedeutung verändern. In diesem Sinn kann man mit Latour  
410 Daten als „immutable mobiles“ bezeichnen, als „*unver-*  
411 *änderliche und kombinierbare mobile Elemente*“ (Latour  
412 2009: 129). Mit Latour ist dabei auch hervorzuheben, dass die  
413 Daten selbst in einer konkreten materiellen Form vorliegen

---

8 In diesem Text beziehen sich Latour und Woolgar zwar auf die *construction of scientific facts* durch die Erzeugung von „literary inscription[s]“ (1986: 87); das Argument lässt sich aber ebenso auf diesen Fakten zugrundeliegenden Daten ausweiten.

414 und für eine Darstellung entsprechend aufbereitet werden  
415 müssen (vgl. z.B. Latour/Woolgar 1986: 50).

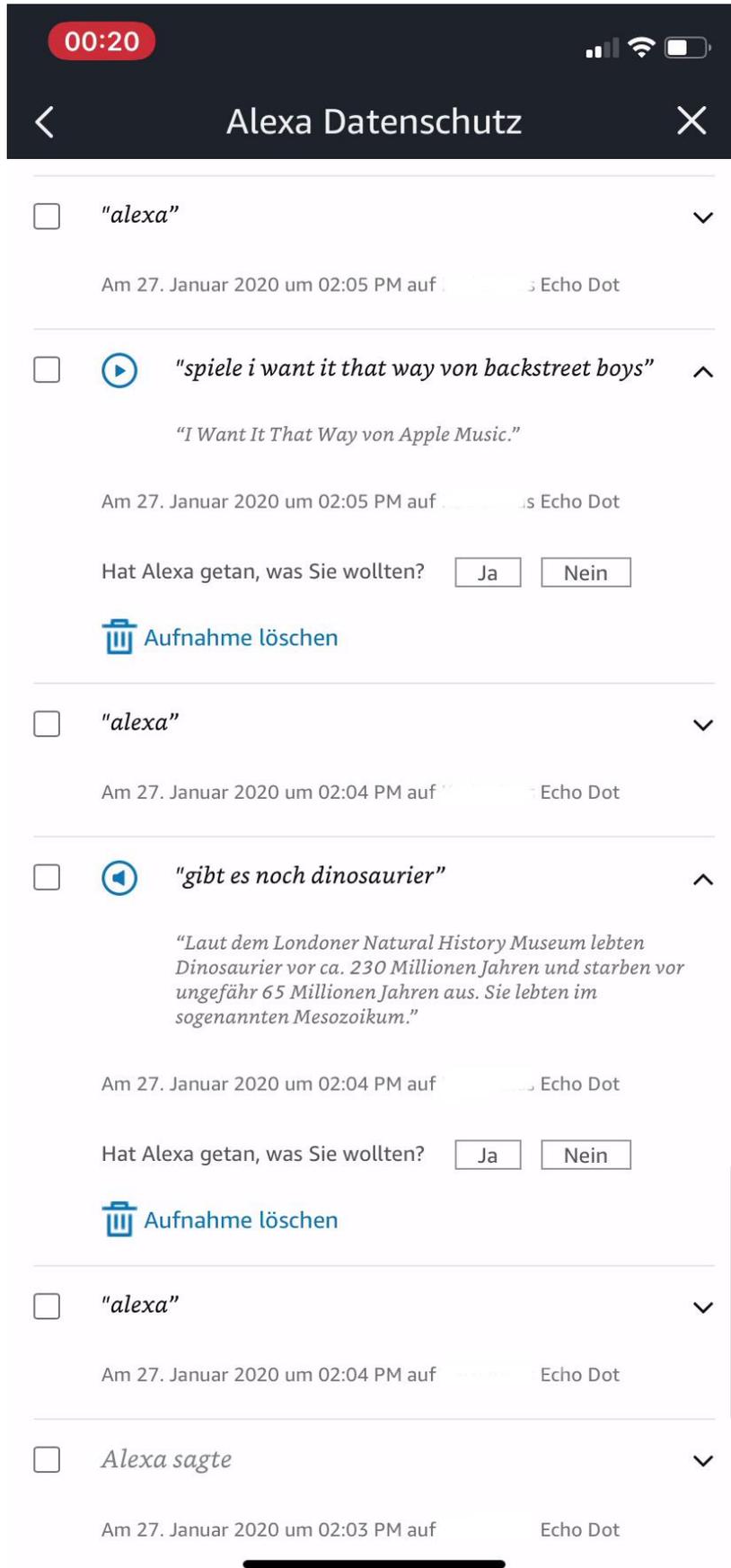
416 Die These von der wechselseitigen Verfasstheit von Daten  
417 wird anhand des Labors besonders anschaulich, da diese  
418 Verfassung hier auf händischer Arbeit beruht. Im Fall der  
419 Protokolldaten ist sie weit weniger sichtbar, da die  
420 Herstellung weitgehend automatisch erfolgt. Zugleich zeigt  
421 jedoch die Präsentation der Protokolldaten durch eine *App*,  
422 dass die Daten auch in diesem Fall in einer gewissen Form  
423 aufbereitet und für die Nutzenden als Daten präsentiert  
424 werden. In diesem Sinn ist der Begriff der „Datenpraktiken“  
425 zu verstehen, wie er im Forschungsprogramm des SFB 1187  
426 entwickelt wird.<sup>9</sup> Die Einbettung der Datenpraktiken in  
427 Infrastrukturen und Plattformen soll im Folgenden beleuchtet  
428 werden.

429 3.2 Die Verflechtung der Protokolldaten mit Infrastrukturen und  
430 Plattformen

431 Die Präsentation der Protokolldaten in der *App* ist für die  
432 Rekonstruktion der Nutzungspraktiken grundsätzlich  
433 aufschlussreich, da sie sowohl den Nutzenden wie den  
434 Forschenden einen chronologischen Einblick in die  
435 konkreten Interaktionen erlaubt und auch Fehlschläge  
436 sichtbar macht. Aber auch die Beschränkung in der Nutzung  
437 der Daten durch die Verhinderung eines direkten Exports der  
438 Daten aus der *App* ist insofern relevant, als sie auf die  
439 Einbettung der Protokolldaten in die Plattform von „Alexa“  
440 und damit der Firma Amazon verweist (vgl. Abschnitt 5). In  
441 der Darstellung der Interaktion mit der *App* wird des  
442 Weiteren ersichtlich, inwiefern der IPA auch auf andere  
443 Dienste zugreift:

---

<sup>9</sup> Siehe: <https://www.mediacoop.uni-siegen.de/de/forschungsprogramm/>



444

445

Screenshot 2

446 Diese Vermittlung der Dienste Dritter sowie die Rahmung der  
447 Protokoll Daten für die Nutzenden lassen den IPA sowohl als  
448 Teil einer Infrastruktur als auch einer Plattform erscheinen.  
449 Da beide Begriffe in der sozial- und kulturwissenschaftlichen  
450 Diskussion digitaler Medien relevant sind, sollen sie auch hier  
451 kurz in ihrer Relevanz für die IPAs geklärt werden. Im Kon-  
452 text der Germanistischen Linguistik wird das Konzept der  
453 Infrastruktur von Matthias Meiler herangezogen und mit dem  
454 Begriff der Kommunikationsform (vgl. Abschnitt 2) vermittelt  
455 (vgl. u.a. Meiler 2019: 73–76, am Beispiel von Weblogs).

456 Beide Begriffe, Plattform wie Infrastruktur, verweisen  
457 etymologisch auf Substrata, welche die Grundlage für andere  
458 Aktivitäten bilden und selbst meist unbemerkt bleiben. Insbe-  
459 sondere im Fall der Infrastruktur wird diese Grundlage sehr  
460 häufig dann im Alltag thematisiert, wenn sie ihren für selbst-  
461 verständlich genommenen Dienst versagt – der Strom fällt  
462 aus, der Abfluss ist verstopft etc. Aus einer solchen  
463 Perspektive können IPAs als Infrastruktur gelten, da sie die  
464 Grundlage für Informationsabfragen über die Dienste Dritter  
465 – Wetterdatenanbieter, Wikipedia, Nachrichtenportale – und  
466 die Steuerung von *Smart Home*-Geräten darstellen. Aber  
467 auch durch die Integration in die Steuerung der häuslichen  
468 Grundeigenschaften wie Wärme und Licht sind IPAs mit den  
469 klassischen Infrastrukturen verknüpft. Aus der Forschungs-  
470 perspektive der *Science and Technology Studies* (STS) lassen  
471 sich IPAs als Teil großtechnischer Systeme (vgl. Hughes 1987)  
472 und der aus ihnen gebildeten Netzwerke begreifen. Diese  
473 sind, wie Paul N. Edwards (vgl. 2003: 186) argumentiert,  
474 sowohl Produkt wie notwendige Voraussetzung moderner  
475 Gesellschaften.

476 In dieser Perspektive sind IPAs als ein weiterer Schritt in  
477 der technischen Anbindung des Haushalts an ein  
478 Versorgungssystem zu verstehen. IPAs basieren auf  
479 Infrastrukturen der Strom-, Telefon- und Internetverbindung  
480 von Haushalten. Des Weiteren wird für die Einrichtung der  
481 Geräte ein Smartphone vorausgesetzt. Der *Smart Speaker* im  
482 Wohnzimmer ist somit Ausläufer eines Netzes an  
483 Verbindungen zu auf der Welt verteilten Rechenzentren und  
484 den dort ablaufenden Rechenprozessen (vgl. Eggert/Kerpen  
485 2018). IPAs sind zum einen auf externe Dienstleister  
486 angewiesen und zum anderen auf eine Fülle an vernetzen

487 Geräten, die in Form „smarter“ Glühbirnen und Thermostate  
488 von den IPAs steuerbar sind.

489 In der Forschung zu digital-vernetzen Medien ist neben  
490 dem Begriff der Infrastruktur vor allem jener der Plattform  
491 zentral. Dieser Begriff der Plattform weist nun spezifisch auf  
492 die Eigentumsverhältnisse und organisatorische Einbettung  
493 der Technologien hin. Die „Plattform-Ökonomie“ (Srnicsek  
494 2016) ist zu einem populären Begriff geworden, um  
495 ökonomische Strategien nicht nur von *Social Media*-Firmen  
496 wie Facebook, sondern einen Trend des Outsourcing und der  
497 betrieblichen Flexibilisierung auf Basis der Analyse großer  
498 Datenmengen zu beschreiben. Wesentlich für Plattformen ist  
499 mithin die Erhebung und Verwertung vielfältiger Nutzungs-  
500 und Nutzer\_innen-Daten (vgl. Srnicsek 2016: 39–43 und  
501 Strüver 2020: 3).

502 Die ersten Medien, die als Plattform beschrieben wurden,  
503 und deren Thematisierung daher die *Platform Studies* be-  
504 gründete, waren Spielkonsolen wie jene der Firmen Atari,  
505 Sega oder Nintendo (vgl. Bogost/Montfort 2009). Sie  
506 standardisieren als Plattform technische Parameter der  
507 Computerspiele und können somit als „two-sided markets“  
508 (Rochet/Tirole 2003) beschrieben werden, die Anbietende  
509 und Kaufinteressierte in Bezug zueinander setzen. Dabei  
510 behalten die Plattformen die Kontrolle über die Regeln des  
511 Markt-Zugangs und der Transaktionen. Kirchner/Beyer  
512 (2016: 329) bezeichnen eben dies als „Plattformlogik“ (vgl.  
513 Staab 2019). Diese zeigt sich insbesondere bei den beiden  
514 großen Smartphone-Betriebssystemen „iOS“ und „Android“,  
515 die Software nur über die jeweiligen *App*-Stores auf den  
516 Endgeräten installierbar machen.

517 Langlois/Elmer (2019) weisen darauf hin, dass sich vormalig  
518 auf *einen* Plattformmarkt beschränkte Unternehmen zu-  
519 nehmend in anderen Bereichen engagieren und sich daher  
520 eine Bewegung von einer geschäftsmäßigen Organisation  
521 funktionspezifischer Plattformen hin zu Unternehmen  
522 beobachten lässt, die umfassende Infrastrukturen bereit-  
523 stellen oder sich in bestehende einschreiben. Dies lässt sich  
524 anhand der drei Unternehmen illustrieren, welche die auf  
525 dem westlichen Markt erfolgreichen IPAs anbieten. Diese  
526 haben jeweils ursprünglich eine bestimmte Plattform  
527 betrieben: Amazon im Versandhandel, Google die Such-  
528 maschine und Apple die Kombination aus Computer-

529 Hardware und Software. Inzwischen haben sie in ver-  
530 schiedene Bereiche expandiert, bei denen sie teilweise in  
531 Konkurrenz zueinander treten. Die jeweiligen IPAs sind in  
532 die Plattformen der Anbieter integriert, denn die Firmen  
533 stellen jeweils intelligente Lautsprecher mit herstellereigenen  
534 Betriebssystemen her, die als Vermittlungsinstanz auch bei  
535 der Kommunikation mit KI-Produkten anderer Anbieter zum  
536 Tragen kommt. Diese vertikale Integration und die „Platt-  
537 formlogik“ der durch die IPAs konstituierten Märkte werden  
538 zugleich von einer Einbettung in eine Infrastruktur von  
539 Diensten der Anbieter sowie Dritter komplementiert. In  
540 diesem Sinn kann man die Einbettung der IPAs in den  
541 häuslichen Alltag wohl mit jener Figur beschreiben, die  
542 Plantin et al. (2018: 306) auf digitale Medien insgesamt  
543 beziehen: als eine Gleichzeitigkeit der „platformization of  
544 infrastructures“ und der „infrastructuralization of platforms“.  
545 Wenn man nun die Protokolldaten und ihre Einbettung in  
546 die „Alexa“-App in Hinblick auf Amazon als Plattform  
547 betrachtet, fallen vor allem die Beschränkungen im Umgang  
548 mit den Daten auf. Zugleich fällt es aus dieser Perspektive  
549 leicht, die Nutzenden lediglich in Abhängigkeit von diesen  
550 Geräten und ihren Einbettungen in Wertschöpfungsketten,  
551 Interfaces und Voreinstellungen zu begreifen. So betreiben  
552 Plattformen laut Ulrich Dolata (2019: 195) über die konkrete  
553 technische Ausgestaltung eine „Kuratierung sozialer Verhält-  
554 nisse und sozialen Verhaltens“ – beispielsweise durch die  
555 mehr oder weniger große Möglichkeit, Beiträge oder Inhalte  
556 privat zu schalten. In der Wissenschafts- und Technik-  
557 soziologie finden sich jedoch primär Perspektiven, welche  
558 den scheinbar starren Begriff der Infrastruktur in vielerlei  
559 Hinsicht aufweichen. Zum einen wird auf die historische  
560 Genese, Pfadabhängigkeit und Wandelbarkeit von Infra-  
561 strukturen verwiesen, zum anderen wird die konstante Arbeit  
562 an der Aufrechterhaltung von Infrastrukturen betont. In  
563 Susan Leigh Stars Forschung (vgl. Star/Bowker 2006) wird  
564 aus der Infrastruktur daher auch ein Verb – *to infrastructure*  
565 –, das den Prozess der Infrastrukturierung beschreibbar  
566 macht. Infrastrukturen erscheinen aus dieser Perspektive als  
567 etwas, das der Pflege und Reparatur, Integration und  
568 Adaption bedarf. Des Weiteren weist die Forschung darauf  
569 hin, dass auch auf der Seite der Nutzenden, die nicht in die

570 professionelle Planung, Konstruktion und Pflege der Infra-  
571 struktur eingebunden sind, sich gleichwohl Umgangsweisen  
572 mit Infrastrukturen herausbilden, die nicht immer den  
573 Intentionen der Herstellenden und Betreibern der Infra-  
574 struktur entsprechen müssen. Beim IPA, wie schon beim  
575 internetfähigen Computer um die Jahrtausendwende (vgl.  
576 Röser et al. 2019), müssen die Mitglieder von Privathaushalten  
577 z.B. entscheiden, in welchem Raum sie das jeweilige  
578 Gerät aufstellen und wer dieses wann und wie nutzen kann  
579 bzw. darf.

580 Die von Schüttpelz (2016: 5) auch auf „infrastrukturelle  
581 Medien“ gemünzte Formulierung von den „kooperativ  
582 erarbeitete[n] Kooperationsbedingungen“ lässt so nicht nur  
583 die Fabriziertheit der Infrastruktur deutlich hervortreten,  
584 sondern auch die daran anschließende Kooperation auf Basis  
585 der Infrastruktur als Bedingung und Ermöglichung der  
586 Kooperation und damit Interaktion. Des Weiteren kann man  
587 mit dem Begriff der Infrastruktur technische Gerätschaften  
588 und Kommunikationsmedien einerseits, die Körper der  
589 Akteur\_innen (vgl. Star/Bowker 2006: 231) und soziale  
590 Aspekte wie die Gepflogenheiten der Interaktion (vgl.  
591 Schegloff 2012) andererseits in ein Kontinuum einreihen, um so  
592 ihre Relevanz für die Interaktion zu betonen. Die Frage wäre  
593 nun, wie sich der praktische Umgang mit den IPA und dessen  
594 Erforschung im Hinblick auf die Einbettung in Plattformen  
595 und Infrastrukturen darstellen.

#### 596 **4. Sprachliche Interaktion und soziale Kontexte**

##### 597 4.1 IPA-Dialoge und sprachliche Interaktion

598 Die Protokolldaten geben Einblicke sowohl in  
599 Charakteristika soziotechnischer IPA-Dialoge als auch – in  
600 deutlich eingeschränktem Maß – in soziale Interaktion unter  
601 anwesenden Nutzer\_innen in Situationen, in denen diese  
602 gemeinsam mit Sprachassistenzsystemen umgehen. IPA-  
603 Dialoge können mit Krummheuer (2010: 323–324)  
604 grundsätzlich als ein „hybrider“ bzw. „ambiger“ Austausch  
605 beschrieben werden, insofern Technisches und Soziales von  
606 den Beteiligten teilweise als Einheit, teilweise unter dem  
607 Aspekt der Differenz behandelt werden. In Verbindung mit

608 diesem soziotechnischen Austausch stoßen wir in den  
609 Protokoll Daten auch auf bruchstückhafte Dokumentationen  
610 sozialer Interaktion, die auf eine IPA-Nutzung in Zwei- oder  
611 Mehrparteienkonstellationen zurückzuführen ist (vgl.  
612 Porcheron et al. 2018; vgl. auch Pitsch et al. 2017).

613 Bei der Konzeptualisierung der sozialen Interaktion im  
614 Verhältnis zum hybriden Austausch mit dem Gerät stützen  
615 wir uns – wie auch Krummheuer (vgl. 2010: 13) – auf die  
616 Arbeiten Goffmans, der soziale Interaktion als „wechselseitige  
617 Wahrnehmung und Kommunikation von zwei  
618 körperlich anwesenden Personen“ versteht, „die wahrnehmen,  
619 dass sie wahrgenommen werden, und einen gemeinsamen  
620 Aufmerksamkeitsfokus teilen“. Der hybride Austausch weist,  
621 wie auch die folgenden Ausführungen zeigen werden, gewisse –  
622 und innerhalb grundsätzlicher Grenzen (vgl. Lotze 2016) tendenziell  
623 zunehmende – Ähnlichkeiten zur sozialen Interaktion auf. Diese  
624 ist aber in höherem Maße indexikalisch (vgl. Garfinkel 1967) und  
625 zeichnet sich durch ein komplexes Geflecht von kontextgebundenen  
626 Erwartungen und Folgeerwartungen aus, die im hybriden Austausch  
627 mit Assistenzsystemen nicht in gleichem Umfang produziert und  
628 eingelöst werden können.<sup>10</sup> Während mithin in den IPA-Dialogen  
629 trotz technischer Fortschritte die Anpassung der Nutzenden an die  
630 Dialogfähigkeit der Systeme in den Vordergrund tritt, zeigen sich  
631 bei der Einbettung der IPA in die soziale Interaktion deutlicher  
632 die kreativen Gestaltungsmöglichkeiten der Nutzenden im Umgang  
633 mit der Technik.

636 Das Verhältnis zwischen den IPA-Dialogen und sozialer  
637 Interaktion soll nachfolgend anhand zweier Beispiele aus dem oben  
638 beschriebenen Protokoll Datenkorpus erläutert werden. In Beispiel (1)  
639 („Pantheon“) zeigt sich die für unser Korpus typische Form eines  
640 zweizügigen IPA-Dialogs: Auf einen Sprachbefehl folgt eine  
641 entsprechende Reaktion des Sprachassistenten.

643 **Beispiel (1): Pantheon**

```
001 M1 definiere PANtheon.
```

---

<sup>10</sup> Dies gilt umso mehr, wenn Assistenzsysteme in arbeitsteilige soziotechnische Netzwerke in institutionellen Kontexten eingebettet sind, vgl. Amrhein/Cyra/Pitsch (2016).

644 In der *App* ist, ausschließlich schriftlich, die ursprünglich  
645 auditiv produzierte Reaktion des Sprachassistenten  
646 dokumentiert:

Das Wort „Pantheon“ hat folgende Bedeutungen: 1) ein antiker, allen Göttern geweihter Tempel und 2) die Gesamtheit der Götter einer (polytheistischen) Religion.

647 Lediglich wenige Ausnahmen in unserem Korpus weisen  
648 mehr als diese beiden Dialogzüge auf, was insofern –  
649 jedenfalls im Blick auf die Form der Dokumentation in den  
650 Protokolldaten – die Annahmen von Krummheuer (2010:  
651 320) zu bestätigen scheint, dass die „Verknüpfung der  
652 einzelnen Redebeiträge [...] im hybriden Austausch“ oft  
653 (noch) „nicht über drei, sondern über zwei Redezüge  
654 organisiert [ist]“ und „eine richtige Interpretation des ersten  
655 Redezugs unterstellt [wird]“. Allerdings scheint es inzwischen  
656 auch Abweichungen von diesem Muster zu geben: Im Zuge  
657 einer zunehmenden Annäherung an soziale Interaktion sollen  
658 auch dritte Dialog-Züge seitens der Nutzer\_innen (z.B.  
659 Reformulierungen von Fragen oder Befehlen) durch die  
660 Systeme ausgewertet werden, um die sequenzielle Angemes-  
661 senheit der technischen Reaktion zu kontrollieren, und es  
662 sollen künftig auch komplexere interaktive Abläufe projiziert  
663 und technisch unterstützt werden (vgl. Strüver 2020: 7–9).  
664 Während Beispiel (1) eine einfache Form des IPA-Dialogs  
665 bzw. von Mensch-Maschine-Interaktion illustriert, zeigen  
666 sich an dem nachfolgend präsentierten Beispiel (2)  
667 („Elefantentrompete“) zwei Phänomene, die einer genaueren  
668 Untersuchung wert sind. In diesem Beispiel versuchen die  
669 beteiligten Sprecherinnen W1 und W2 den *Smart Speaker* zur  
670 Imitation von Tierstimmen zu bewegen.

671 **Beispiel (2): „Elefantentrompete“**

001 W1 wie macht die KAtze,

002 W1 wie macht die KAtze,

003 W2 wie macht der eleFA:NT,

672 Schriftlich ist als Antwort auf die von W1 und W2 an den  
673 *Smart Speaker* gerichteten Fragen in der Smartphone-*App*  
674 folgende (ursprünglich vom Gerät auditiv wiedergegebene,  
675 aber nur schriftlich erfasste)<sup>11</sup> Ausgabe des IPA dokumentiert:

Der Elefant macht - Sie haben Elefantentrompete  
verwendet. Amazon empfiehlt Ihnen daher einen  
weiteren beliebten Bildung und Nachschlagewerke-  
Skill, Animal Sounds. Würden Sie ihn gern  
ausprobieren?

676 Die darauf folgenden mündlichen Äußerungen von W1 und  
677 W2 sind in der *App* wiederum auditiv abrufbar:

004 W1 nein.

005 W2 NE::IN,=

006 W1 =sehr UNgern würd ich dat AUsprob-

678 Beispiel (2) illustriert mehrere Phänomene: Erstens gibt es  
679 Aufschluss über mögliche Störungen und Reparaturen in IPA-  
680 Dialogen. Zweitens zeigt sich, dass auch der IPA bzw. das  
681 integrierte *Natural Language Processing* zweite Züge so  
682 erweitern können, dass Folgezüge relevant gesetzt werden:  
683 So produziert der IPA nicht nur den Austausch ab-  
684 schließende Turns, sondern auch sequenz-eröffnende Turn-  
685 Expansionen, die eine Fortsetzung erwartbar machen.

686 Wenden wir uns zunächst den Reparaturen zu: In den  
687 Zeilen 001 und 002 fällt auf, dass Sprecherin W1 zweimal  
688 hintereinander die gleiche Frage an das Gerät richtet („wie  
689 macht die KAtze“, Z. 001). Dieses wiederholte Einfordern  
690 einer Antwort ist ein Indiz dafür, dass die von W1 relevant  
691 gesetzte Antwort des IPA nicht geliefert wurde. Auch die  
692 über die *App* abrufbaren Protokolldaten zeigen an, dass keine  
693 Aktivität des Geräts erfasst wurde. Nachdem auch der zweite  
694 Versuch von W1 scheitert, den Sprachassistenten zu einer  
695 Reaktion zu bringen, wiederholt Sprecherin W2 die Frage,  
696 ersetzt bei gleichbleibender syntaktischer Struktur das Lexem

---

<sup>11</sup> Die Antwort wird für die Endnutzer\_innen nur auditiv ‚vorgetragen‘; die Produktion erfolgt jedoch als Teil des Natural Language Processings schriftlich und wird anschließend durch Speech Processing auditiv verbalisiert (siehe dazu Natale 2020).

697 „KATze“ jedoch durch das Lexem „eleFA:NT“ (Z. 003). An  
698 dieser Stelle wird deutlich, wie Sprecherin W2 im Rahmen  
699 des Testens der sogenannten *Skills* des Smart Speakers mit  
700 der von Sprecherin W1 relevant gesetzten, aber vom Gerät  
701 nicht realisierten Reaktion umgeht und die von W1 an das  
702 Gerät gestellte Frage entsprechend anpasst (vgl. Schegloff  
703 2012: 252–254). Solche von den beteiligten menschlichen  
704 Gesprächspartnern realisierten Lösungen können zudem  
705 Aufschluss darüber geben, was erstere als mögliche Ur-  
706 sache(n) für das Ausbleiben des relevant gesetzten zweiten  
707 Turns – hier des Adjazenzpaares Frage-Antwort – deuten: In  
708 diesem Fall, dass der IPA das zum Lexem „Katze“ erwartete  
709 Tiergeräusch nicht produzieren kann, aber andere Tierge-  
710 räusche gegebenenfalls verfügbar sind.

711 Betrachten wir mit Fokus auf die inkrementelle Er-  
712 weiterung des IPA-Dialogs nun den zweiten Teil der  
713 Äußerung des Sprachassistenten. Nach dem Abspielen des  
714 Elefantengeräuschs (im in der *App* hinterlegten Transkript  
715 durch einen Gedankenstrich visualisiert) wird der – hier im  
716 einfachen Fall hybrider Dialoge abgeschlossene – Zug seitens  
717 des IPA um eine Empfehlung expandiert: An die beant-  
718 wortete Frage anknüpfend wird eine Frage verbalisiert, die  
719 eine Antwort seitens des/der am Dialog beteiligten Menschen  
720 relevant setzt. Die vom Sprachassistenten in Bezug auf den  
721 Skill „Animal Sounds“ aufgeworfenen Frage „Würden Sie ihn  
722 gern ausprobieren?“ verneint W1 (Z. 004) und in unmittel-  
723 barem Anschluss auch ihre Gesprächspartnerin W2 mit  
724 einem paraverbal verstärkten „NE::IN,“ (Z. 005). Die von W1  
725 daraufhin geäußerte und die Sequenz abschließende Ex-  
726 pansion „sehr UNgern würd ich dat AUsprob-“ (Z. 006) zeigt,  
727 wie W1 auf der Ausgabe des Sprachassistenten operiert, in-  
728 dem sie das in der Ausgabe enthaltene sprachliche Material in  
729 ihrer Äußerung aufgreift, durch syntaktische Variation  
730 transformiert (vgl. Goodwin 2018: 431) und dialektal durch-  
731 fährt. Ob W1 ihre Äußerung nicht vollständig zu Ende führt  
732 oder ob die Aufnahme des IPA frühzeitig abbricht – z.B. nach  
733 dem erkannten Sprachbefehl „nein.“ (Z. 004) – lässt sich auf  
734 Basis des Ausschnitts nicht nachvollziehen; wahrscheinlich ist  
735 jedoch, dass der IPA nicht auf die Rezeption einer zweiten  
736 Sprecherin ausgerichtet ist und bereits nach dem Erkennen  
737 der ersten, turnabschließenden Äußerung („nein.“, Z. 004)  
738 nicht mehr im „Listening Mode“ ist, sodass die Aufnahme

739 abbricht, bevor W1 ihre inkrementelle Erweiterung (Z. 006)  
740 vollendet hat (siehe dazu auch Pitsch et al. 2017: 396).

741 Die Äußerung von W1 (Z. 006) ist doppelt funktionalisiert:  
742 Einerseits bekräftigt sie den Abschluss des IPA-Dialogs,  
743 gleichzeitig nutzt die Sprecherin die Äußerung, um in der  
744 sozialen Interaktion mit der ko-präsenten Sprecherin W2 die  
745 Äußerungen des IPA als Ressource für die weitere Inter-  
746 aktion nutzbar zu machen: Ihre spöttische Zurückweisung  
747 des Vorschlags verweist indexikalisch auf eine durch  
748 Geringschätzung des IPA-Skills geprägte Haltung und den  
749 Kontext des gemeinsamen unterhaltsamen Erprobens mehr  
750 oder weniger skurriler und überflüssiger technischer Funk-  
751 tionen (vgl. Abschnitt 4.2). Ein weiteres, bereits für Gruppen-  
752 interaktionen, in die ein technisches Artefakt eingebettet ist,  
753 beobachtetes Phänomen (vgl. Reeves/Porcheron/Fischer  
754 2019; Habscheid et al. 2020) scheint sich in diesem Beispiel  
755 zu manifestieren: Die scheinbar von W1 an das Gerät ge-  
756 richtete Äußerung adressiert (primär) ihre ko-präsente  
757 Gesprächspartnerin W2 – diesen Schluss lassen die dialektale  
758 Durchfärbung und die für den IPA-Dialog irrelevante Er-  
759 gänzung um die Bewertung „sehr UNgern“ (Z. 006) zu. Die  
760 Äußerung erfüllt somit auch und sogar primär Zwecke für die  
761 soziale Situation, denn „what is said *to* the device is  
762 necessarily often said *around* others“ (Reeves/Porcheron/  
763 Fischer 2019: 49). Dabei nutzt sie Bestandteile des IPA-  
764 Dialogs und bindet diesen so in die soziale Interaktion ein; es  
765 entsteht eine Verflechtung von hybridem Austausch mit dem  
766 Gerät und sozialer Interaktion mit W2 vor dem Gerät.

767 Da die von Amazon erhobenen Mitschnitte von Sprachein-  
768 und Sprachausgaben als verschriftete isolierte Einzel-  
769 elemente ohne die Situation, in der sie realisiert werden,  
770 gespeichert werden, lässt sich über die genauere Einbettung  
771 sowie die prosodische Realisierung der Äußerungen des  
772 Sprachassistenten im vorliegenden Beispiel (2) nur  
773 spekulieren. Da diese Informationen aber erforderlich für  
774 eine verlässliche Rekonstruktion der sprachlichen Interaktion  
775 sind, lassen sich über die mit dem IPA-Dialog verbundene  
776 soziale Interaktion *vor* und *mit* dem Gerät anhand der Proto-  
777 kolldaten nur in recht geringem Umfang Rückschlüsse ziehen.  
778 Multimodale Interaktionsanalysen, die den Kontext, die  
779 sozialräumliche Umgebung und insofern auch die materiale  
780 Beschaffenheit der Interaktionssituation mit dokumentieren,

781 sind für eine umfassende gesprächsanalytische Betrachtung  
782 zwischenmenschlicher Interaktion im Umgang mit dem IPA  
783 unerlässlich.

784 Dagegen kann für die Beschäftigung mit IPA-Dialogen die  
785 Analyse der Protokolldaten aufschlussreicher sein. IPA-  
786 Dialoge sind deutlich von sozialer, zwischenmenschlicher  
787 Interaktion zu unterscheiden, in der die sequenzielle  
788 Organisation, das Turn-Taking und die Gleichzeitigkeit von  
789 Produktion und Rezeption sowie die gegenseitige Wahr-  
790 nehmungswahrnehmung (vgl. Goffman 1983: 2) generische  
791 Merkmale der Interaktion und entscheidend für die Lösung  
792 von kommunikativen Problemen sind (vgl. Schegloff 2012:  
793 246; siehe auch Auer 2000). IPA-Dialoge sind in diesen Hin-  
794 sichten grundlegend anders strukturiert. So lässt sich etwa im  
795 zweizügigen Fall nur mit großen Einschränkungen überhaupt  
796 von einer sequenziellen Organisation sprechen. Allerdings  
797 scheinen sich die Merkmale – in Grenzen – anzunähern,  
798 wenn Antworten des IPA durch Expansionen Folge-  
799 äusserungen erwartbar machen, die in die Position einer  
800 Verständnissicherung eintreten.

801 4.2 Produkt- und nutzungsseitige sprachliche Herstellung von  
802 Alltagskontexten

803 Nicht zuletzt geben die Protokolldaten auch einen Aufschluss  
804 darüber, in welche alltagspraktischen – kommunikativen und  
805 über Kommunikation hinausreichenden – bzw. diskursiven  
806 Kontexte die IPA systemseitig gestellt sind bzw. in der  
807 nutzungsseitigen sozialen Interaktion gestellt werden (vgl.  
808 dazu Habscheid 2016). So zeigt etwa Beispiel 1 eine Wissens-  
809 recherche, die an der Bedeutung (alltagssprachlich:  
810 „Definition“) eines sprachlichen Ausdrucks (*Pantheon*)  
811 festgemacht wird. In Beispiel 2, in dem neben dem IPA-  
812 Dialog mehr oder weniger zufällig auch ein Auszug aus der  
813 sozialen Interaktion „vor“ dem Gerät in den Protokolldaten  
814 dokumentiert ist, wird die Praktik erkennbar, in einer ge-  
815 selligen Aktivität das z.T. skurrile Funktions- und Leistungs-  
816 spektrum des IPA zu erkunden; außerdem zeigt sich bruch-  
817 stückhaft der in Alltagspraxis situierte Diskurs von  
818 Nutzer\_innen *über* das Gerät, hier eine mokante Bewertung  
819 (vgl. Abschnitt 4.1).

820 Protokolldaten wie Beispiel 1 dokumentieren Nutzungs-  
821 situationen, die systemseitig angelegt sind, sie sind damit  
822 sowohl Ausdruck der Nutzung als auch einer gewissen  
823 Formation der Nutzenden durch den IPA. So lässt sich an-  
824 hand der Protokolldaten erschließen, dass routinierte  
825 Nutzungsweisen des Geräts über weite Strecken das übliche  
826 Repertoire betreffen (vgl. auch Ammari et al. 2019), u.a. das  
827 Abrufen von Musiktiteln, Playlists und Filmen, Erzählwitzen,  
828 Tierlauten und Naturgeräuschen; Wissensrecherchen und  
829 Wetterabfragen; Rechenaufgaben; Befehle zu Kommuni-  
830 kationsfunktionen („spIEle meine ↑↑NACHrichten ab;“; „was  
831 steht in meinem kaLENDER?“) und *Smart home*-Geräten („te  
832 VAU an;“). Seltener werden nach unseren Protolldaten  
833 Dialogsequenzen in solchen Kontexten auch von Seiten des  
834 IPA initiiert:

835 **Beispiel (3):** „Benachrichtigung“

Übrigens, du hast eine neue Benachrichtigung. Soll  
ich deine Benachrichtigung vorlesen?

001 M1 äh JA:,

836 Im Gegensatz zu Standardsituationen zeigt Beispiel 2 die  
837 kreativere Herstellung einer Nutzungspraxis durch zwei  
838 Nutzer\_innen, wobei die in der *App* angelegten Möglich-  
839 keiten aufgegriffen und in den alltäglichen Diskurs einge-  
840 bettet werden. Dass allerdings auch spielerische Nutzungs-  
841 weisen produktseitig antizipiert sind, zeigt die Kate-  
842 gorisierung eines ähnlichen Skills (in „Neuheiten und  
843 Humor“) in einer sequenzinitiiierenden Erweiterung des IPA  
844 in Beispiel 4 (vgl. zur Interaktionsstruktur Abschnitt 4.1):

845 **Beispiel (4):** „Katzensimulator“

Sie haben Katzensimulator verwendet. Amazon empfiehlt  
Ihnen daher einen weiteren beliebten Neuheiten und  
Humor-Skill, Ein Kompliment. Test gefällig?

846 Auf spielerische Nutzungskontexte verweisen auch be-  
847 ziehungssensitive Äußerungen (Beleidigungen, Auf-  
848 forderungen zu Komplimenten etc.), mit denen der  
849 humanoide Charakter des Geräts testend herausgefordert  
850 wird, wie in Beispiel 5:

851 **Beispiel (5): „Liebe“**

001 M1 LIEBST du mich?

852 Hier kann, nebenbei bemerkt, die paradoxe Antwort des IPA  
853 durchaus zum Weiterdenken anregen:

Ich bin gerne bei dir, aber zu echter Liebe bin ich  
leider gar nicht in der Lage.

854 Zu den systemseitig angelegten, subsidiären Nutzungs-  
855 möglichkeiten gehört, wie – neben entsprechenden  
856 Nutzer\_innen-Äußerungen („schalte das Mikrofon aus.“) –  
857 auch die Protokolldaten als solche zeigen, nicht zuletzt auch  
858 die Regulierung des Datenschutzes in Bezug auf die erstellten  
859 Aufzeichnungen, die damit in den Verantwortungsbereich der  
860 Nutzer\_innen geschoben wird. In einem von Niklas Strüver  
861 (vgl. 2020: 1) beschriebenen Amazon-Werbevideo, in dem  
862 humorvoll-selbstironisch ein Bezug des IPA zu historischen  
863 Vorläufern hergestellt wird, tritt dementsprechend neben  
864 Dienstmädchen, Hofnarr, Zeitungsjunge und anderen auch  
865 eine Sekretärin auf, die Tonbandaufzeichnungen löschen soll.

866 **5. Fazit und Ausblick**

867 Als Teil der *App* geben die Protokolldaten und ihre  
868 Inszenierung durch die *App* einen Eindruck davon, wie sie in  
869 die Formierung eines bestimmten Gebrauchs durch die  
870 Plattform und Infrastruktur eingebunden sind. Unsere  
871 Erhebung der Protokolldaten über Bildschirmvideos zeigt  
872 darüber hinaus die im Programm angelegte Interaktion der  
873 Nutzenden mit den Protokolldaten. Während der Zweck der  
874 Darstellung unter „Aktivität“ erst einmal unklar bleibt, wird  
875 die Darstellung der Daten im anderen Fall deutlich als  
876 Möglichkeit gerahmt, als Nutzer\_in Einfluss auf die von  
877 Amazon gespeicherten Daten zu nehmen. Die Möglichkeit  
878 der Datenlöschung kann dabei als Reaktion von Amazon auf  
879 die öffentliche Debatte bezüglich der Verwendung von  
880 personenbezogenen Daten durch Internetfirmen verstanden  
881 werden. Aktuell ist die Löschung aller aufgezeichneten  
882 Befehle möglich, nicht nur in der *App*, sondern auch per

883 Sprachbefehl (vgl. Herbig 2020). Wie bereits dargestellt, ist in  
884 der *App* die Löschung der Daten, nicht aber ihr Export  
885 möglich. Die Daten können hier also lediglich der Kontrolle  
886 durch Amazon entzogen werden, sollen jedoch entsprechend  
887 dem Interface-Design nicht ohne weitere Umstände für  
888 weitere Zwecke verwendet werden.

889 Man kann die Protokolldaten als „natürliche Daten“  
890 (Salheiser 2019) betrachten und dabei von der oben be-  
891 schriebenen Intervention der Forschenden absehen. Denn  
892 die Protokolldaten selbst sind ohne die Forschenden zu-  
893 stande gekommen, lediglich die Erhebung ist das Ergebnis  
894 eines Forschungsprozesses.

895 Auf dieser Basis geben diese Daten einen Einblick in den  
896 tatsächlichen Vollzug der Beobachtung der Nutzenden durch  
897 die Infrastruktur. Da hier nicht nur die erfolgreichen  
898 Sprachbefehle, sondern auch die Missverständnisse proto-  
899 kolliert werden, werden auch jene gescheiterten Anläufe des  
900 IPA studierbar, die vielleicht von den Nutzenden nicht be-  
901 merkt worden wären. Zugleich lässt sich nachvollziehen,  
902 wann und wie der IPA genutzt wurde, es entsteht somit ein  
903 Protokoll der Nutzung, das um einiges genauer ist als in-  
904 direkte Erhebungsmethoden (vgl. Tietze/Roßbach 1991): Wie  
905 erwartet, erlauben die Protokolldaten gewisse Aufschlüsse  
906 darüber, wie derartige Systeme in Verbindung mit Platt-  
907 formen und Infrastrukturen operieren, wie die Dialogsysteme  
908 heute gestaltet sind und welche Nutzungskontexte sie  
909 nahelegen.

910 Da im Regelfall nur der unmittelbare Sprachbefehl nach  
911 dem *Wake Word* erfasst wird, wird die Einbettung in soziale  
912 Interaktion und deren praktische bzw. diskursive Kontexte  
913 dagegen nur im Ausnahmefall und unvollständig erfasst. Diese  
914 Unvollständigkeit betrifft sowohl die simultan ablaufende  
915 Interaktion in anderen als den dokumentierten Modalitäten  
916 als auch die Einbettung in größere sequenzielle Abläufe. Die  
917 Daten können u.U. begrenzt Aufschluss über soziale Inter-  
918 aktion und Alltagspraxis unter Einbindung der IPA geben,  
919 bedürfen aber in dieser Hinsicht unbedingt der Ergänzung  
920 durch Beobachtungs- bzw. audiovisuelle Aufzeichnungs-  
921 daten.

922 Dagegen zeigt unsere Erhebung der Protokolldaten über  
923 Bildschirmvideos die im Programm angelegte Interaktion der  
924 Nutzenden mit den Protokolldaten selbst. Die *App* und damit

925 Amazon scheinen jedenfalls daran interessiert zu sein, durch  
926 die Möglichkeit der Datenlöschung einen gewissen Daten-  
927 schutz, und zwar als Aufgabe des Nutzers, zu offerieren.  
928 Dabei bleibt es jedoch eine empirische Frage, inwiefern diese  
929 Möglichkeit der Datenlöschung von den Nutzenden tatsäch-  
930 lich verwendet wird und wie diese eingeschätzt wird. Studien  
931 zum Gebrauch von IPAs aus den Sozialwissenschaften und  
932 HCI (vgl. bspw. Malkin et al. 2019) deuten darauf hin, dass  
933 dieses Wissen insgesamt nicht sonderlich verbreitet ist. Selbst  
934 dann, wenn die Möglichkeit der Dateneinsicht und Löschung  
935 bekannt ist, wird sie anscheinend selten genutzt. Vor diesem  
936 Hintergrund kann unsere Erhebung als Krisenexperiment  
937 (vgl. Garfinkel 1967: 37) verstanden werden, da sie die  
938 Nutzenden in vielen Fällen zum ersten Mal mit der  
939 Möglichkeit konfrontiert, ihren Datengebrauch zu  
940 kontrollieren.

#### 941 **Literatur**

- 942 Ammari, Tawfiq/Kaye, Jofish/Tsai, Janice Y./Bentley, Frank  
943 (2019): Music, Search, and IoT: How People (Really) Use  
944 Voice Assistants. In: *ACM Transactions on Computer-  
945 Human Interaction (TOCHI)* 26, 17:1–17:28.
- 946 Auer, Peter (2000): *On line*-Syntax – Oder: was es bedeuten  
947 könnte, die Zeitlichkeit der mündlichen Sprache ernst zu  
948 nehmen. In: *Sprache und Literatur* 31, 43–56.
- 949 Amrhein, Antje/Cyra, Katharina/Pitsch, Karola (2016):  
950 Processes of Reminding and Requesting in Supporting  
951 People with Special Needs: Human Practices as Basis for  
952 Modeling a Virtual Assistant? In: *EDIA 2016. Proceedings of  
953 the 1st Workshop on Ethics in the Design of Intelligent  
954 Agents. In conjunction with the 22th European Conference  
955 on Artificial Intelligence – ECAI 2016*. The Hague,  
956 Holland, August 30, 2016, 14-19. URL: [http://ceur-  
957 ws.org/Vol-1668/paper3.pdf](http://ceur-<br/>957 ws.org/Vol-1668/paper3.pdf), letzter Zugriff am 02.01.2021.
- 958 Ayaß, Ruth/Meyer, Christian (Hg.) (2012): *Sozialität in Slow  
959 Motion. Theoretische und empirische Perspektiven*.  
960 Festschrift für Jörg Bergmann. Wiesbaden: Springer VS.
- 961 Bogost, Ian/Montfort, Nick (2009): Platform Studies:  
962 Frequently Questioned Answers. In: *Proceedings of the  
963 Digital Arts and Culture Conference 2009, After Media:*

- 964 *Embodiment and Context*. University of California, Irvine.  
965 URL: <https://escholarship.org/uc/item/01r0k9br>, letzter  
966 Zugriff am 22.6.2020.
- 967 Brock, Alexander/Peter Schildhauer (Hg.) (2017):  
968 *Communication Forms and Communicative Practices.*  
969 *New Perspectives on Communication Forms, Affordances*  
970 *and What Users Make of Them*. Frankfurt am Main u.a.:  
971 Peter Lang (Language and Text Studies, Volume 15).
- 972 Dieter, Michael/Gerlitz, Carolin/Helmond, Anne/Tkacz,  
973 Nathaniel/Vlist, Fernando van der/Weltevrede, Esther  
974 (2018): *Store, interface, package, connection: methods and*  
975 *propositions for multi-situated app studies*. Siegen: SFB  
976 Medien der Kooperation. URL: [https://dspace.ub.uni-](https://dspace.ub.uni-siegen.de/handle/ubsi/1341)  
977 [siegen.de/handle/ubsi/1341](https://dspace.ub.uni-siegen.de/handle/ubsi/1341), letzter Zugriff am 09.10.2020.
- 978 Dolata, Ulrich (2019): Plattform-Regulierung. Koordination  
979 von Märkten und Kuratierung von Sozialität im Internet.  
980 In: *Berliner Journal für Soziologie* 29, 179–206.
- 981 Durkheim, Emile (1893/1992): *Über soziale Arbeitsteilung.*  
982 *Studie über die Organisation höherer Gesellschaften.*  
983 Frankfurt a. M.: Suhrkamp (= stw 1005). (Original 1893).
- 984 Edwards, Paul N. (2003): Infrastructure and Modernity.  
985 Force, Time and Social Organization in the History of  
986 Socio-Technological Systems. In: Misa, Thomas J./Brey,  
987 Philip/Feenberg, Andrew (Hg.): *Modernity and Technology.*  
988 Cambridge, MA; London: MIT Press, 185–225.
- 989 Eggert, Michael/Kerpen, Daniel (2018): Wer Datengesellschaft  
990 sagt, muss auch Cloud-Computing sagen. Die Cloud als  
991 zentrale Infrastruktur der datafizierten Gesellschaft. In:  
992 Houben, Daniel/Prietl, Bianca (Hg.): *Datengesellschaft.*  
993 *Einsichten in die Datafizierung des Sozialen*. Bielefeld:  
994 transcript, 155–177.
- 995 Garfinkel, Harold (1960/2012): Die rationalen Eigenschaften  
996 von wissenschaftlichen und Alltagsaktivitäten. In: Ayaß,  
997 Ruth/Meyer, Christian (Hg.): *Sozialität in Slow Motion.*  
998 *Theoretische und empirische Perspektiven*. Festschrift für  
999 Jörg Bergmann. Wiesbaden: Springer VS, 41–57. (Original  
1000 1960).
- 1001 Garfinkel, Harold (1967): *Studies in Ethnomethodology.*  
1002 Cambridge: Polity.
- 1003 Goffman, E. (1983). The Interaction Order. In: *American*  
1004 *Sociological Review* 48, 1–17.

- 1005 Goodwin, Charles (2018): *Co-Operative Action*. New York:  
1006 Cambridge University Press.
- 1007 Habscheid, Stephan (2016): Handeln in Praxis. Hinter- und  
1008 Untergründe situierter sprachlicher  
1009 Bedeutungskonstitution. In: Deppermann, Arnulf/Feilke,  
1010 Helmuth/Linke, Angelika (Hg.): *Sprachliche und*  
1011 *kommunikative Praktiken*. Berlin/New York: de Gruyter  
1012 (IDS Jahrbuch 2015), 127–151.
- 1013 Habscheid, Stephan (2020): Rezension zu: Brock, Alexander  
1014 & Peter Schildhauer (Hg.) (2017): Communication Forms  
1015 and Communicative Practices. New Perspectives on  
1016 Communication Forms, Affordances and What Users Make  
1017 of Them. Frankfurt am Main u.a.: Peter Lang (Language  
1018 and Text Studies, Volume 15). In: *ZRS* 12 (1-2), 150-155.
- 1019 Habscheid, Stephan/Hrncal, Christine/Carros, Felix/Lüssem,  
1020 Jens (2020): Professionelle Emotionalität und humanoide  
1021 Robotik in der institutionellen Kommunikation. In: Gruber,  
1022 Helmut/Spitzmüller, Jürgen/de Cillia, Rudolf (Hg.):  
1023 *Institutionelle und organisationale Kommunikation.*  
1024 *Theorie, Methodologie, Empirie und Kritik*. Wien: V&R  
1025 Unipress, 169–188.
- 1026 Hausendorf, Heiko/Kesselheim, Wolfgang/Kato,  
1027 Hiloko/Breitholz, Martina (2017): *Textkommunikation: ein*  
1028 *textlinguistischer Neuansatz zur Theorie und Empirie der*  
1029 *Kommunikation mit und durch Schrift*. Berlin: de Gruyter  
1030 (Reihe Germanistische Linguistik, 308).
- 1031 Herbig, Daniel (2020): Amazon Alexa: Nutzer können  
1032 Speichern von Sprachaufnahmen abschalten. In: *heise.de*,  
1033 24.09.2020. URL: [https://www.heise.de/news/Amazon-](https://www.heise.de/news/Amazon-Alexa-Nutzer-koennen-Speichern-von-Sprachaufnahmen-abschalten-4912086.html?view=print)  
1034 [Alexa-Nutzer-koennen-Speichern-von-](https://www.heise.de/news/Amazon-Alexa-Nutzer-koennen-Speichern-von-Sprachaufnahmen-abschalten-4912086.html?view=print)  
1035 [Sprachaufnahmen-abschalten-4912086.html?view=print](https://www.heise.de/news/Amazon-Alexa-Nutzer-koennen-Speichern-von-Sprachaufnahmen-abschalten-4912086.html?view=print),  
1036 letzter Zugriff am 05.10.2020.
- 1037 Hirschauer, Stefan (2004): Praktiken und ihre Körper. Über  
1038 die materiellen Partizipanden des Tuns. In: Hörning, Karl  
1039 H./Reuter, Julia (Hg.): *Doing Culture. Neue Positionen zum*  
1040 *Verhältnis von Kultur und sozialer Praxis*. Bielefeld:  
1041 Transcript, 73–91.
- 1042 Hughes, Thomas P. (1987): The Evolution of Large Technical  
1043 Systems In: Bijker, Wiebe E./Hughes, Thomas P./Pinch,  
1044 Trevor (Hg.): *The Social Construction of Technological*  
1045 *Systems: New Directions in the Sociology and History of*  
1046 *Technology*. Cambridge, MA: MIT Press, 51–82.

- 1047 Kirchner, Stefan/Beyer, Jürgen (2016): Die Plattformlogik als  
1048 digitale Marktordnung. In: *Zeitschrift für Soziologie* 45,  
1049 324–339.
- 1050 Krummheuer, Antonia (2010): *Interaktion mit virtuellen*  
1051 *Agenten? Zur Aneignung eines ungewohnten Artefakts.*  
1052 Stuttgart: Lucius&Lucius (Qualitative Soziologie 11).
- 1053 Langlois, Ganaele/Elmer, Greg (2019): Impersonal  
1054 subjectivation from platforms to infrastructures. In: *Media,*  
1055 *Culture & Society* 41, 236–251.
- 1056 Latour, Bruno/Woolgar, Steve (1986): *Laboratory Life. The*  
1057 *Construction of Scientific Facts* 2. Aufl., Princeton, NJ:  
1058 Princeton University Press.
- 1059 Latour, Bruno (2007): *Eine neue Soziologie für eine neue*  
1060 *Gesellschaft. Einführung in die Akteur-Netzwerk-Theorie.*  
1061 Berlin: Suhrkamp.
- 1062 Latour, Bruno (2009): Die Logistik der immutable mobiles. In:  
1063 Döring, Jörg/Thielmann, Tristan (Hg.): *Mediengeographie.*  
1064 *Theorie – Analyse – Diskussion.* Bielefeld: Transcript, 111–  
1065 144.
- 1066 Lotze, Netaya (2016): *Chatbots. Eine linguistische Analyse.*  
1067 Frankfurt u.a.: Peter Lang.
- 1068 Malkin, Nathan/Deatrick, Joe/Tong, Allen/Wijesekera,  
1069 Primal/Egelman, Serge/Wagner, David (2019): Privacy  
1070 Attitudes of Smart Speaker Users. In: *Proceedings on*  
1071 *Privacy Enhancing Technologies* 2019, 250–271.
- 1072 Meiler, Matthias (2019): Zur praxeologischen Verhältnis-  
1073 bestimmung von Materialität, Medialität und Mentalität  
1074 oder: Medien als Praxis. In: *Zeitschrift für Semiotik* 41 (1–  
1075 2), 63–88.
- 1076 Meyer, Christian/Ayaß, Ruth (2012): Einleitung. In: Ayaß,  
1077 Ruth/Meyer, Christian (Hg.): *Sozialität in Slow Motion.*  
1078 *Theoretische und empirische Perspektiven.* Festschrift für  
1079 Jörg Bergmann. Wiesbaden: Springer VS, 11–18.
- 1080 Natale, Simone (2020): To believe in Siri: A critical analysis of  
1081 AI voice assistants. In: *Communicative Figurations*  
1082 *Working Paper* 32. URL: [https://www.kommunikative-  
1083 figurationen.de/en/publications/working-papers/](https://www.kommunikative-figurationen.de/en/publications/working-papers/).
- 1084 Pickering, Andrew (1993): The Mangle of Practice: Agency  
1085 and Emergence in the Sociology of Science. In: *American*  
1086 *Journal of Sociology* 99, 559–589.
- 1087 Pitsch, Karola/Gehle, Raphaela/Dankert, Timo/Wrede,  
1088 Sebastian (2017): Interactional Dynamics in User Groups.

- 1089 In: Wrede, Britta (Hg.): *Proceedings of the 5th International*  
1090 *Conference on Human Agent Interaction*. Bielefeld,  
1091 Germany, 10/17/2017 - 10/20/2017. New York: ACM Press,  
1092 393–397.
- 1093 Plantin, Jean-Christophe/Lagoze, Carl/Edwards, Paul  
1094 N./Sandvig, Christian (2018): Infrastructure studies meet  
1095 platform studies in the age of Google and Facebook. In:  
1096 *New Media & Society* 20, 293–210.
- 1097 Porcheron, Martin/Fischer, Joel E./Reeves, Stuart/Sharples,  
1098 Sarah (2018): Voice Interfaces in Everyday Life. In: *CHI*  
1099 *2018*, April 21–26, 2018, Montreal, QC, Canada. URL:  
1100 [http://www.cs.nott.ac.uk/~pszsr/files/porcheron-2018-](http://www.cs.nott.ac.uk/~pszsr/files/porcheron-2018-voice-interfaces-in-everyday-life.pdf)  
1101 [voice-interfaces-in-everyday-life.pdf](http://www.cs.nott.ac.uk/~pszsr/files/porcheron-2018-voice-interfaces-in-everyday-life.pdf), letzter Zugriff am  
1102 26.07.2019.
- 1103 Reeves, Stuart/Porcheron, Martin/Fischer, Joel (2019): “This  
1104 is not what we wanted”: Designing for Conversation with  
1105 Voice Interfaces. In: *Interactions* 26 (1), 46–51. DOI:  
1106 <https://doi.org/10.1145/3296699>.
- 1107 Rochet, Jean-Charles/Tirole, Jean (2003): Platform  
1108 Competition in Two-Sided Markets. In: *Journal of the*  
1109 *European Economic Association* 1, 990–1029.
- 1110 Röser, Jutta/Müller, Kathrin Friederike/Niemand,  
1111 Stephan/Roth, Ulrike (2019): *Das mediatisierte Zuhause im*  
1112 *Wandel. Eine qualitative Panelstudie zur Verhäuslichung*  
1113 *des Internets*. Wiesbaden: Springer VS.
- 1114 Salheiser, Axel (2019): Natürliche Daten: Dokumente. In:  
1115 Baur, Nina/Blasius, Jörg (Hg.): *Handbuch Methoden der*  
1116 *empirischen Sozialforschung*. Wiesbaden: Springer, 1119–  
1117 1134. URL: [https://doi.org/10.1007/978-3-658-21308-4\\_80](https://doi.org/10.1007/978-3-658-21308-4_80),  
1118 letzter Zugriff am 09.10.2020.
- 1119 Schegloff, Emanuel A. (2006/2012): Interaktion: Infrastruktur  
1120 für soziale Institutionen, natürliche ökologische Nische der  
1121 Sprache und Arena, in der Kultur aufgeführt wird. In:  
1122 Ayaß, Ruth /Meyer, Christian (Hg.): *Sozialität in Slow*  
1123 *Motion. Theoretische und empirische Perspektiven*.  
1124 Festschrift für Jörg Bergmann. Wiesbaden: Springer VS,  
1125 245–268. (Original 2006).
- 1126 Schneider, Jan Georg (2017): Medien als Verfahren der  
1127 Zeichenprozessierung: Grundsätzliche Überlegungen zum  
1128 Medienbegriff und ihre Relevanz für die Gesprächs-  
1129 forschung. In: *Gesprächsforschung* 18, 34–55.

- 1130 Schüttpelz, Erhard (2016): Infrastrukturelle Medien und  
1131 öffentliche Medien. In: *Media in Action* 0, 1–21 (Pre-  
1132 Publication). URL: [https://www.mediacoop.uni-](https://www.mediacoop.uni-siegen.de/wp-content/uploads/2016/06/schuettpelz_infrastrukturelle_medien.pdf)  
1133 [siegen.de/wp-](https://www.mediacoop.uni-siegen.de/wp-content/uploads/2016/06/schuettpelz_infrastrukturelle_medien.pdf)  
1134 [content/uploads/2016/06/schuettpelz\\_infrastrukturelle\\_m](https://www.mediacoop.uni-siegen.de/wp-content/uploads/2016/06/schuettpelz_infrastrukturelle_medien.pdf)  
1135 [edien.pdf](https://www.mediacoop.uni-siegen.de/wp-content/uploads/2016/06/schuettpelz_infrastrukturelle_medien.pdf), letzter Zugriff am 20.01.2021.
- 1136 Schüttpelz, Erhard/Meyer, Christian (2017): Ein Glossar zur  
1137 Praxistheorie. „Siegener Version“ (Frühjahr 2017). In:  
1138 *Navigationen* 17 (1), 155–163.
- 1139 Schüttpelz, Erhard/Meyer, Christian (2018): Charles  
1140 Goodwin’s Co-Operative Action: The Idea and the  
1141 Argument. In: *Media in Action* 1, 171–188. URL:  
1142 [https://www001.zimt.uni-](https://www001.zimt.uni-siegen.de/ojs/index.php/mia/article/view/37)  
1143 [siegen.de/ojs/index.php/mia/article/view/37](https://www001.zimt.uni-siegen.de/ojs/index.php/mia/article/view/37), letzter Zugriff  
1144 am 20.01.2021.
- 1145 Schütz, Alfred (1943/2010): Das Problem der Rationalität in  
1146 der Sozialwelt. In: Eberle, Thomas Samuel/Dreher,  
1147 Jochen/Sebal, Gerd (Hg.): *Zur Methodologie der*  
1148 *Sozialwissenschaften* (Werkausgabe Band 4). Konstanz:  
1149 UVK, 203–233.
- 1150 Scollon, Ron/Scollon, Suzie Wong (2004): *Nexus Analysis:*  
1151 *Discourse and the Emerging Internet*. London: Routledge.
- 1152 Selting, Margret/Auer, Peter/Barth-Weingarten,  
1153 Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner,  
1154 Karin/Couper-Kuhlen, Elizabeth/Deppermann,  
1155 Arnulf/Gilles, Peter/Günthner, Susanne/Hartung,  
1156 Martin/Kern, Friederike/Mertzluft, Christine/Meyer,  
1157 Christian/Morek, Miriam/Oberzaucher, Frank/Peters,  
1158 Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock,  
1159 Anja/Uhmann, Susanne (2009): Gesprächsanalytisches  
1160 Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung*  
1161 10, 353–402.
- 1162 Srnicek, Nick (2016): *Platform Capitalism*. Cambridge, MA:  
1163 MIT Press.
- 1164 Staab, Philipp (2019): *Digitaler Kapitalismus. Macht und*  
1165 *Herrschaft in der Ökonomie der Unknappheit*. Berlin:  
1166 Suhrkamp.
- 1167 Star, Susan Leigh/Bowker, Geoffrey C. (2006): How to  
1168 Infrastructure. In: Lievrouw, Leah A./Livingston, Sonia M.  
1169 (Hg.): *Handbook of New Media*. London: Sage, 151–162.
- 1170 Strüver, Niklas (2020): *Amazons Alexa als Infrastruktur der*  
1171 *Zukunft? Eine qualitative Betrachtung von*

- 1172      *Nutzungspraktiken mit Sprachassistenten im Kontext*  
1173      *digitaler Plattformtechnologien*. Masterarbeit. RWTH  
1174      Aachen.
- 1175      Tietze, Wolfgang/Roßbach, Hans-Günther (Hg.) (1991):  
1176      *Mediennutzung und Zeitbudget. Ansätze, Methoden,*  
1177      *Probleme*. Wiesbaden: Deutscher Universitätsverlag.
- 1178      Wissenschaftliche Dienste des Deutschen Bundestags. (2019)  
1179      *Zulässigkeit der Transkribierung und Auswertung von*  
1180      *Mitschnitten der Sprachsoftware „Alexa“ durch Amazon*.  
1181      Berlin. URL:  
1182      [https://www.bundestag.de/resource/blob/650728/3f72e6ab](https://www.bundestag.de/resource/blob/650728/3f72e6abc1c524961e5809002fe20f21/WD-10-032-19-pdf-data.pdf)  
1183      [c1c524961e5809002fe20f21/WD-10-032-19-pdf-data.pdf](https://www.bundestag.de/resource/blob/650728/3f72e6abc1c524961e5809002fe20f21/WD-10-032-19-pdf-data.pdf),  
1184      letzter Zugriff am 02.01.2020.