

Inhaltsverzeichnis

Abbildungsverzeichnis	9
Tabellenverzeichnis.....	12
Zusammenfassung.....	15
Summary	29
Résumé.....	42
Sintesi.....	56
1. Einleitung und Kontext.....	69
<i>Murat Karaboga</i>	
1.1. Hintergrund und Zielsetzung der Studie.....	69
1.2. Zielsetzung	72
1.2.1. Fragestellung der Studie	72
1.3. Begriffsklärung	75
1.4. Methodologie.....	78
1.4.1. Arbeitsschritte und Methoden	78
1.4.2. Projektkonsortium.....	79
2. Ist- und Trendanalyse	83
<i>Frank Ebbers, Murat Karaboga, Greta Runge & Michael Friedewald</i>	
2.1. Historischer Rückblick zu den Grundlagen von Deepfakes	83
2.1.1. Digitale Manipulation von Bildern und Videos.....	83
2.1.2. Klonen und Synthetisieren von Stimmen	85
2.2. Technologien zur Fälschung und Synthetisierung von Bild- und Videomaterial	86
2.2.1. Generative Adversarial Networks.....	87
2.2.2. Autoencoder.....	89
2.3. Sechs Techniken zur Erstellung von bildbasierten Deepfakes.....	90

2.3.1.	Manipulation des Gesichtsausdrucks («facial reenactment»).....	90
2.3.2.	Gesichts-Morphing («face-morphing»).....	91
2.3.3.	Gesichtsaustausch («face-swapping»)	91
2.3.4.	Gesichtsgeneratoren	92
2.3.5.	Ganzkörperpuppenspiel («full body puppetry»)	92
2.3.6.	Aus Texteingaben generierte Deepfake-Videos («Text-to-Video»)	93
2.4.	Technologien zum Klonen der Stimme.....	95
2.5.	Technologien zum Generieren von inhaltlich authentischem Text.....	96
2.6.	Technische Gegenmassnahmen.....	98
2.6.1.	Prävention	98
2.6.2.	Erkennung.....	101
2.7.	Untersuchung von Deepfake-Detektoren	107
2.7.1.	Recherche von Detektoren und Kontaktaufnahme	109
2.7.2.	Resultate	110
2.7.3.	Fazit zu Deepfake-Detektoren.....	113
2.8.	Bibliometrische Auswertung wissenschaftlicher Publikationen	114
2.8.1.	Methodik.....	115
2.8.2.	Resultate	116
2.9.	Ist- und Trendanalyse: Zwischenfazit.....	122
3.	Wahrnehmung von Deepfakes in der Schweizer Bevölkerung....	125
	<i>Daniel Vogler, Adrian Rauchfleisch & Gabriele de Seta</i>	
3.1.	Theorie und Forschungsstand	125
3.1.1.	Erfahrung mit Deepfakes.....	126
3.1.2.	Chancen- und Risikowahrnehmung von Deepfakes	127
3.1.3.	Erkennen von Deepfakes	128
3.2.	Methodische Vorgehensweise	131
3.2.1.	Vorstudie und Pretest.....	131
3.2.2.	Hauptstudie	131
3.3.	Resultate	134
3.3.1.	Erfahrungen mit Deepfakes.....	134

3.3.2. Werden Deepfakes als Chance oder als Risiko wahrgenommen?	136
3.3.3. Kann die Schweizer Bevölkerung Deepfakes erkennen?	145
3.4. Hauptbefunde.....	150
4. Deepfakes im Recht	153
<i>Nula Frei & Sophia Rovelli</i>	
4.1. Grundrechtlicher und urheberrechtlicher Schutz von Deepfakes.....	153
4.1.1. Meinungs-, Informations- und Kunstdfreiheit	154
4.1.2. Urheberrecht	157
4.2. Schutz vor schädlichen Auswirkungen von Deepfakes.....	159
4.2.1. Schutz im Zivilrecht	159
4.2.2. Schutz im Strafrecht.....	165
4.2.3. Verfahrensrechtliche Geltendmachung.....	174
4.2.4. Zwischenfazit zum Schutz vor Deepfakes.....	183
4.3. Deepfakes vor Gericht.....	184
4.3.1. Deepfakes als (manipulierte) Beweismittel	184
4.3.2. Einsatz von Deepfakes zur Aufklärung von Straftätern	186
4.3.3. Zwischenfazit.....	187
4.4. Öffentlich-rechtliche Vorgaben	187
4.4.1. Medienregulierung.....	188
4.4.2. Schutz von Wahlen und Abstimmungen vor Verfälschung.....	190
4.4.3. Zwischenfazit.....	191
4.5. Regulierungsmöglichkeiten von Deepfakes	192
4.5.1. Allgemeine Herausforderungen von Deepfakes	192
4.5.2. Bestehende Regulierungsansätze	195
5. Deepfakes im Journalismus.....	205
<i>Patric Raemy, Manuel Puppis & Gwendolyn Gurr</i>	
5.1. Theorie und Forschungsstand	206
5.1.1. Deepfakes und Journalismus.....	206
5.1.2. Forschungsstand: Was wir bisher wissen	211
5.2. Methodische Vorgehensweise	214

5.3.	Resultate	217
5.3.1.	Thematisierung der Herausforderungen durch Deepfakes in der Journalismusausbildung.....	217
5.3.2.	Umgang von Medienorganisationen mit Deepfakes	219
5.4.	Hauptbefunde.....	248
6.	Deepfakes in der Politik.....	253
	<i>Murat Karaboga, Greta Runge & Michael Friedewald</i>	
6.1.	Forschungsstand zu Deepfakes in der Politik	254
6.1.1.	Der digitale Strukturwandel der Öffentlichkeit	255
6.1.2.	Mögliche Implikationen von Deepfakes für die Politik	257
6.1.3.	Zwischenfazit: Einsatz von Deepfakes in der Schweizer Politik	266
6.2.	Umfrage im Schweizer Parlament und der Bundesverwaltung.....	267
6.2.1.	Methodisches Vorgehen.....	267
6.2.2.	Resultate	271
6.2.3.	Zusammenfassung.....	276
6.3.	Szenarien zu Deepfakes in der Politik	276
6.3.1.	Kurzzusammenfassung der Szenarien	281
6.4.	Zwischenfazit.....	285
7.	Deepfakes in der Wirtschaft.....	287
	<i>Murat Karaboga, Greta Runge & Michael Friedewald</i>	
7.1.	Herausforderungen von Deepfakes in der Wirtschaft	288
7.2.	Chancen von Deepfakes in der Wirtschaft	296
7.2.1.	Methodisches Vorgehen.....	296
7.2.2.	Resultate	300
7.3.	Szenarien zu Deepfakes in der Wirtschaft	309
7.3.1.	Kurzzusammenfassung der Szenarien	310
7.4.	Massnahmen zum Schutz und zur Schadensbegrenzung.....	312
7.4.1.	Sensibilisierung von Mitarbeitenden	313
7.4.2.	Strukturelle Massnahmen in Organisationen	317
7.5.	Zwischenfazit.....	328

8. Empfehlungen	331
<i>Nula Frei, Murat Karaboga, Manuel Puppis, Daniel Vogler, Patric Raemy, Frank Ebbers, Greta Runge, Adrian Rauchfleisch, Gabriele de Seta, Gwendolyn Gurr, Michael Friedewald & Sophia Rovelli</i>	
8.1. Staat als Regulierungsakteur	331
8.1.1. Plattformregulierung	331
8.1.2. Unterstützung von Opferberatungsstellen, die auf Cyberdelikte spezialisiert sind	333
8.1.3. Regelung von digitalen Beweisen im Strafverfahrensrecht (Deepfakes zwecks Visualisierung von Tathergängen oder zur virtuellen Tatortbegehung)	333
8.1.4. Unterstützung vertrauenswürdiger Hinweisgeber (Trusted Flaggers) ..	333
8.1.5. Beteiligung an der Schaffung internationaler Normen und Regeln in den Bereichen Desinformation und Cyberkriminalität	334
8.2. Gesellschaft und Bildungseinrichtungen	334
8.2.1. Sensibilisierungsarbeit in der Ausbildung, Informationskampagnen von staatlichen Akteuren und Engagement journalistischer Medien ..	334
8.2.2. Selbstverantwortung der Bürgerinnen und Bürger	335
8.3. Organisationen in allen Branchen	335
8.3.1. Weiterbildungen zu Medien- und Informationskompetenz in sämtlichen Branchen	335
8.3.2. Förderung von Authentifizierungs- und Kennzeichnungsverfahren ..	336
8.3.3. Nutzung fortschrittlicher Authentifizierungsverfahren und von Zwei-Faktor-Authentifizierung	337
8.3.4. Freiwillige Meldung von Deepfake-Vorfällen durch betroffene Organisationen	338
8.3.5. Einrichtung von spezialisierten Teams, die im Falle eines Deepfake-Einsatzes darauf vorbereitet sind, Massnahmen zur Schadensbegrenzung zu ergreifen	338
8.4. Kommunikationsbranche	338
8.4.1. Selbstregulierung der PR- und Werbebranche	338
8.5. Plattformbetreiber	339

8.5.1. Selbstregulierungsmassnahmen gegen irreführende und illegale Inhalte	339
8.6. Medienorganisationen, Medienausbildung, Nachrichtenagentur	340
8.6.1. Hochhaltung journalistischer Standards bei der Erkennung von Deepfakes und Förderung der Medienethik	340
8.6.2. Förderung forensischer Verifikationsmethoden in den Redaktionen und Herstellung von Transparenz über eigene Bemühungen.....	340
8.6.3. Stärkung des Presserats als von der Branche eingesetzte Selbstregulierungsorganisation zur Einhaltung ethischer Standards im Journalismus.....	341
9. Schlussfolgerungen.....	343
<i>Murat Karaboga</i>	
Autorinnen und Autoren	347
Literatur	349
Anhang	387
A.1. Detaillierte Abbildungen der Wahrnehmungsstudie	387
A.2. Interviewleitfaden TA-SWISS-Projekt.....	394
A.3. Liste der Codes	397
A.4. Deepfake-Produktionssoftware	399
A.5. Szenarien zu Deepfakes in der Politik	404
A.5.1. Individuelle Ebene	404
A.5.2. Institutionelle Ebene	408
A.5.3. Gesellschaftliche Ebene.....	411
A.6. Szenarien zu Deepfakes in der Wirtschaft.....	421
A.6.1. Individuelle Ebene	421
A.6.2. Organisationsebene	422
A.6.3. Marktebene	431
Mitglieder der Begleitgruppe.....	435
Projektmanagement TA-SWISS.....	436