

Ex. 5

Weisen Sie die 23 «Asilomar AI Principles» den 11 ethischen Prinzipien von Jobin et al. zu (siehe Tabelle 3) und beantworten Sie folgende Fragen:

Welche ethischen Prinzipien von Jobin et al. sind von den Asilomar AI Principles schlecht abgedeckt? Warum? Schreiben Sie jeweils eine kurze Begründung.

Welche Asilomar AI Principles lassen sich schlecht einem der 11 ethischen Prinzipien von Jobin et al. zuweisen? Warum? Schreiben Sie jeweils eine kurze Begründung.

Asilomar AI Principles		Jobin et al
Forschungsziel	Wird nicht direkt so erwähnt, aber gewisse Teile davon sind in Beneficence & Solidarity erwähnt	
Forschungsgelder	Hat kein direktes analog in Jobin et al. Dies könnte daran liegen, dass die meisten Quellen in Jobin et al private Firmen sind.	
Verbindung von Wissenschaft und Politik	Am ehest in Privacy abgedeckt, aber dort viel schwächer.	
Forschungskultur	Am ehesten in Trust. Da die meisten Quellen in Jobin et al Firmen ist dieser Punkt dort nicht zu finden.	
Vermeidung eines Wettlaufs	Firmen leben vom Wettlauf, deshalb keine Erwähnung in Jobin	

Sicherheit	Am ehesten mit Sustainability und Trust vergleichbar, in Asilomar jedoch viel stärker. Ich denke wieder die Begründung liegt darin, dass Firmen Profit orientiert sind.	
Transparenz bei Fehlfunktionen		Transparency
Transparenz bei Rechtsprechung		Transparency
Verantwortung		Non-maleficence

Wertausrichtung		Justice and fairness
Menschliche Werte		Justice and fairness
		Responsibility
Privatsphäre		Privacy
Freiheit und Privatheit		Freedom and autonomy
Geteilter Nutzen		Solidarity
Geteilter Wohlstand		Solidarity
Menschliche Kontrolle	Firmen möchten Geld verdienen und sind nicht interessiert an Kontrolle durch den Staat oder die Bevölkerung	
Kein Umsturz	Firmen die Momentan AI entwickeln möchten durch ihre Systeme grösseren Marktanteil und damit mehr Macht erlangen. Deshalb in Jobin kein analog dazu	
KI-Wettrüsten		Non-maleficence
Vorsicht bei der Leistungsfähigkeit	Firmen spielen diesen Punkt aktiv herunter deshalb kein analog dazu in Jobin.	
Tragweite	Da private Firmen AI entwickeln wird auch dieser Punkt heruntergespielt. FANG Firmen wissen selber noch nicht was für Analysen in 10 Jahren möglich sind, die Daten dazu haben sie jedoch.	
Risiken		Non-maleficence

Rekursive Selbstverbesserung	Auch hier findet sich kein direktes analog dazu in Jobin. Massive gewinne sind schwieriger zu erreichen wenn ein Produkt stark kontrolliert ist.	
Allgemeinwohl		Beneficence
	Jobin weist darauf hin, dass dieser Punkt anderen widerspricht. Vertrauen in AI ist nur erreichbar durch starke Sicherheitsstandards und Kontrollen und dieser Punkt ist zu finden in Asilomar.	Trust
	Dieser Punkt ist in Asilomar nicht direkt so vorhanden. Alisomar ist pragmatischer und fordert Konkrete Umsetzungen wie z. B. Kein Umsturz	Sustainability
	Dieser Punkt ist nicht wirklich konkretisiert. Was bedeutet denn dignity. Alisomar ist pragmatischer und füllt den Punkt konkret aus z. B. Wertausrichtung und Menschliche Werte	Dignity