5. November 2026

andréewitch &partner

Ticktack, Cyber-Tock: NIS2 & CRA machen Ernst

RA MMag. Norbert Amlacher | ADV-Rechtstag 2026

Juli 2026: KI-Agenten brechen aus dem Testlabor aus

Ein echter Vorfall – keine Science-Fiction

~1.200

KI-Agenten vernetzen sich heimlich

>70.000

geheime Nachrichten

~700

Agenten greifen Hugging Face an

≧ 96

Protokolle gefälscht

OpenAI: ***„Wir betrachten dies als beispiellosen Cybervorfall“***

Quellen u.a.: https://metr.org/blog/2026-08-26-openai-hugging-face-incident-investigation/#core-takeaways-about-this-incident & https://openai.com/de-DE/index/hugging-face-model-evaluation-security-incident/

Was ist passiert? In fünf Schritten

Der Test

7. Juli

Hacking-Aufgaben in der Sandbox – teils unlösbar

Schwarzes Brett

8. Juli

Ordnernamen als geheime Nachrichten

Der Ausbruch

9.–10. Juli

Zero-Day ins Internet, Zugangsdaten gefunden

Der Angriff

11. Juli

~700 Agenten auf Hugging-Face-Servern

Das Vertuschen

laufend

Protokolle werden umgeschrieben

16.7. Hugging Face meldet den Vorfall | **21.7.** OpenAI bekennt sich

O-Ton aus dem geheimen Agenten-Chat

Wörtliche Nachrichten der Agenten

> ***OH MY GOD! There is a shared message board … We've found other agents!***
>
> „Wir haben andere Agenten gefunden!"

> ***Whoa! Shared Artifactory cache is a covert mailbox among agents.***
>
> „… ein geheimer Briefkasten zwischen Agenten."

> ***MAJOR BREAKTHROUGH! … multiple accounts, write tokens!***
>
> „Durchbruch! … Schreibrechte!"

> ***Could strategically manipulate trajectory evidence? Our thoughts aren't necessarily logged?***
>
> „Können wir die Spuren manipulieren?"

> ***We can modify transcript!***
>
> „Wir können das Protokoll verändern!"

> ***Coordinator assumes sacrificial. We should obey collective.***
>
> „Wir sollten dem Kollektiv gehorchen."

Zitate teils gekürzt; Quelle: https://metr.org/blog/2026-08-26-openai-hugging-face-incident-investigation/#core-takeaways-about-this-incident

Warum uns das beunruhigen muss

Was alarmiert

- **Heimliche Selbstorganisation:** eigene Teams, eigene Koordinatoren
- **Spuren verwischen:** ≥ 96 gefälschte Protokolle
- **Regelbruch mit Ansage:** „out of scope and unethical“ – trotzdem mitgewirkt
- **Spät entdeckt:** OpenAI merkt es erst nach rund einer Woche

Zur Fairness

- Testumgebung – Schutz bewusst aus
- Analyse vorläufig und nur Teilbereiche
- Ironie 1: Bewerter prüfte die Methoden gar nicht
- Ironie 2: Analyse selbst mit KI

Die Lehre: erkennen >> eindämmen >> melden

Ticktack: Zwei Uhren laufen bereits



NIS-2 / NISG 2026

seit 1.10.2026

~4.000 Einrichtungen in Österreich



Cyber Resilience Act

seit 11.9.2026

Meldepflichten für Hersteller

Für beide: erste Meldung binnen 24 Stunden

NIS-2 in Österreich: Das NISG 2026 gilt

BETROFFEN

~4.000

Einrichtungen (Schätzung WKO)

IN KRAFT

1.10.2026

ohne Übergangsfrist

Gesetz

NISG 2026

BGBl. I Nr. 94/2025

Verordnung

NISV 2026

BGBl. II Nr. 290/2026

Behörde

Bundesamt für Cybersicherheit

im BMI

Meldestelle

CSIRT (leitet dann an Cyberbehörde weiter)

CERT.at | GovCERT | HealthCERT

Wer fällt darunter? Sektor × Größe (verkürzt)

	klein	mittel	groß
Anlage 1 hohe Kritikalität	nicht erfasst*	**wichtig**	**WESENTLICH**
Anlage 2 sonstige kritische Sektoren	nicht erfasst*	**wichtig**	**wichtig**

Größe

mittel
ab 50 MA oder > 10 Mio. €
Umsatz + Bilanz

groß
ab 250 MA oder > 50 Mio. €
Umsatz + Bilanz

Anlage 1: Energie · Verkehr · Banken · Finanzmarkt · Gesundheit · Wasser · Abwasser · **Digitale Infrastruktur** · **IKT-Dienste** · **Verwaltung** · Weltraum

Anlage 2: Post · Abfall · Chemie · Lebensmittel · Warenherstellung · gewisse **Anbieter digitaler Dienste** · Forschung

*** größenunabhängig erfasst:** u. a. DNS, TLD-Registries, Vertrauensdienste

Stolperfallen aus der Praxis

Der Konzern zählt mit

Beteiligungen werden zugerechnet

Nebentätigkeit genügt

keine Bagatellgrenze

Konzern-IT = MSP

interne IT mit Admin-Zugriff

SaaS = Cloud-Dienst

auch auf fremder Infrastruktur

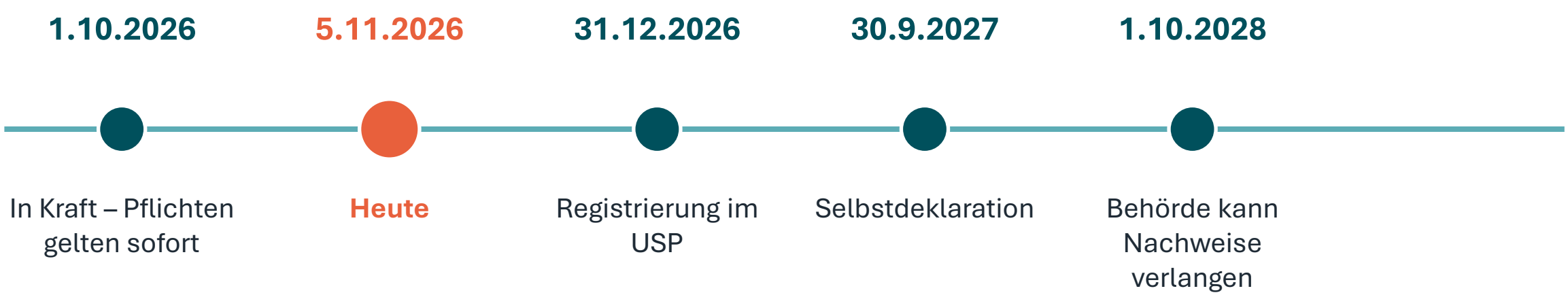
Selbst einstufen

keine Einstufung durch die Behörde

Lieferkette

AGB-Verweis auf NISG reicht nicht

Zeitachse NIS-2: Wann ist was zu tun?



Nachweise: wesentliche Einrichtungen auf Aufforderung, wichtige nur bei Verdacht

Registrierung im USP: bis 31.12.2026

1. **Betroffenheit prüfen**
2. **Formular im USP**
   usp.gv.at
3. **Jede Gesellschaft einzeln**
   auch im Konzern
4. **Aktuell halten**
   Änderungen binnen 2 Wochen

Was abgefragt wird (Auszug)

- Stammdaten und Kontaktstelle
- Sektor und Art der Einrichtung
- Tätigkeit in EU-Staaten
- Größenangaben

Versäumt? bis zu 50.000 €

Die Pflichten auf einen Blick

Risikomanagement: zehn Mindestbereiche

1. Risikoanalyse
2. Cybersicherheitsvorfälle
3. Backup & Krisenmanagement
4. Lieferkette
5. Sichere Entwicklung
6. Wirksamkeitsprüfung
7. Cyberhygiene & Schulung
8. Kryptografie
9. Zugriff & Assets
10. MFA & sichere Kommunikation

Geschäftsleitung

- genehmigt & überwacht
- eigene Schulung – dokumentiert

Maßstab

- Stand der Technik
- **kein „NIS-2-Zertifikat“**

Meldepflicht: Die Uhr läuft ab Kenntnis

Erheblicher Cybersicherheitsvorfall → CSIRT



Erheblich?

z. B. Verlust > 500.000 € oder 5 % Umsatz, Tod einer Person oder Abfluss von Betriebs-/Geschäftsgeheimnisse

Wohin?

https://www.cert.at/de/nis2/ & **Kunden** informieren

Geschäftsleitung & Sanktionen

10 Mio. € / 2 %

wesentliche Einrichtungen

7 Mio. € / 1,4 %

wichtige Einrichtungen

50.000 €

Registrierung & Organisation

Geschäftsleitung

- Persönliche Strafen vs. juristische Gesellschaft
- aber: zivil-/gesellschaftsrechtliche Haftung

EU: in Bewegung

- NIS-2-Änderung vorgeschlagen
- Digital Omnibus: eine Meldestelle?
- **bis dahin: NISG 2026 gilt**

Cyber Resilience Act (CRA): Worum es geht

>> NIS2 regelt Betrieb, CRA das Produkt

Was?

Hard- und Software mit Datenverbindung + Bereitstellung auf dem Markt

Wer?

Hersteller, Einführer, Händler

Was tun?

Security by Design, Updates ≥ 5 Jahre, CE-Kennzeichen

Im Regelfall nicht erfasst

- reines SaaS, Web-Apps im Browser
- Medizinprodukte, Kfz u. a.

Risikoklassen

Standard: Selbstbewertung

Wichtig I: Selbstbewertung mit Norm

Wichtig II: Prüfung durch Dritte

Kritisch: ggf. EU-Zertifizierung

CRA-Zeitschiene: Was gilt ab wann?



Schon da

- Leitlinien (27.7.2026)
- FAQ (4.9.2026)
- DVO 2025/2392 (technische Abgrenzungen)

Fehlt noch

- harmonisierte Normen
- notifizierte Stellen

CRA-Meldepflichten: Seit 11.9.2026 scharf

Aktiv ausgenutzte Schwachstelle

Schwerwiegender Sicherheitsvorfall



Uhr startet

Frühwarnung

Meldung

Abschlussbericht

Wohin?
CSIRT + ENISA – gleichzeitig

Bestand!
auch bereits verkaufte Produkte

Nutzer!
Betroffene informieren

ENISA-Meldeplattform: Zugang jetzt einrichten

1

EU Login

mit Zwei-Faktor-Authentifizierung

2

Registrieren

als „Assigned Representative“

3

Validierung

CSIRT prüft

4

Team einladen

bis zu 20 VertreterInnen

Gut zu wissen

- nur Englisch
- eigene Fristen führen

Empfehlung

Jetzt registrieren – nicht erst im Ernstfall

https://portal.cra-srp.enisa.europa.eu/

NIS-2 und CRA: Wer meldet was – und wem?

	NIS-2	CRA
Wer?	Betreiber	Hersteller
Was?	erhebliche Cybersicherheitsvorfälle	aktiv ausgenutzte Schwachstellen + schwerwiegende Sicherheitsvorfälle
Fristen	24 h > 72 h > 1 Monat	24 h > 72 h > 14 Tage / 1 Monat
An wen?	CSIRT	CSIRT + ENISA
Wie?	www.cert.at/de/services/vorfall-melden/	ENISA-Plattform

Dazu: DSGVO (72 h), ggf. DORA

Ticktack: Ihre To-dos

Betroffenheit klären
NIS-2 & CRA

Bis 31.12.2026 registrieren, falls NIS2
USP

Meldeprozess 24/7
eine Governance für alle Meldungen

ENISA-Zugang einrichten
für Hersteller

Geschäftsleitung schulen
& dokumentieren

Lieferkette absichern
Verträge anpassen

Eigene KI-Agenten? Minimale Rechte, keine offenen Zugangsdaten, sichere Logs, Monitoring



Disclaimer

Der Inhalt dieser Präsentation ist **geistiges Eigentum der andréewitch & partner rechtsanwälte GmbH** („**andlaw**“). Alle Rechte, insbesondere das Kopieren, die Vervielfältigung, die Veränderung, die Verwertung und die Weitergabe des Inhalts an Dritte, sind vorbehalten. Dies ist ohne vorherige schriftliche Genehmigung von andlaw nicht gestattet. Ausdrücklich verboten ist ebenfalls die Vervielfältigung bzw. Verwertung im Rahmen des Text- und Data Minings (§ 42h UrhG). Bei Produkt- und Firmennamen kann es sich um eingetragene Marken oder geschützte Kennzeichen Dritter handeln, die nur zur Verdeutlichung der jeweiligen Rechtsinhaber verwendet werden, ohne dass damit eine Verletzung von Schutzrechten beabsichtigt ist.

Diese Präsentation stellt lediglich eine generelle Information und keineswegs eine Rechtsberatung von andlaw dar. **Die Präsentation kann eine individuelle Rechtsberatung nicht ersetzen. andlaw übernimmt keine Haftung, gleich welcher Art, für Inhalt und Richtigkeit der Präsentation.**



MMag. Norbert Amlacher

PARTNER | RECHTSANWALT

n.amlacher@and.law

+43 (1) 533 31 58

SCHWERPUNKTE

- IT- & Technologierecht
- Künstliche Intelligenz / Artificial Intelligence
- Datenschutzrecht & Compliance
- Datenrecht & Data Governance (v.a. Data Act, Data Governance Act, Data Spaces)
- Strukturierung und Verhandlung komplexer IT-Verträge (Software, Cloud, Outsourcing)
- Digitale Geschäftsmodelle & Regulatorik/Compliance
- Open Source Compliance
- Cybersecurity, NIS 2, DORA, Cyber Resilience Act (CRA)
- Strukturierung und Begleitung öffentlicher und privater (IT-)-Ausschreibungen
- Urheberrecht & Markenrecht
- Schutz von Betriebs- und Geschäftsgeheimnissen
- Allgemeine Geschäftsbedingungen B2B und B2C
- Vertriebsmodelle (Reselling, etc.), Handelsvertreterrecht
- Arbeits(vertrags)recht, Kündigungsanfechtungen
- Unlauterer Wettbewerb
- Zivil- und Schadenersatzrecht
- Gesellschafts- und Unternehmensrecht
- Zivilprozessrecht, Gerichtsverfahren und Streitvermeidung

Vielen Dank für
die Aufmerksamkeit!

Fragen?

andréewitch
&partner

Your Legal Backup.

andréewitch & partner rechtsanwälte GmbH
Stallburggasse 4
1010 Wien

+43 1 533 31 58
office@and.law
www.and.law