DRAFTING THE FIRST AI CODE OF PRACTICE ANNI HELLMAN

MEMBER OF AI-DS WG
FORMER HEAD OF UNIT AT THE EUROPEAN
COMMISSION

FOR GENERAL-PURPOSE AI (GPAI) MODELS

Critical **co-regulatory tool** designed to bridge the gap until formal standards for GPAI models are fully developed and apply (expected by August 2027 or later).

THE PROCESS

- The Al Office facilitated its drafting through four working groups involving nearly 1000 stakeholders (businesses, Member States, civil society, experts).
- The third draft of the Code of Practice was published on March 11, 2025.
- The final version of the Code of Practice was to be published May 2025.
- Last information for the working groups is that the publication will take place at latest in August 2025,

CODE OF PRACTICE FOR GENERALPURPOSE AI (GPAI) MODELS

THE CONTENT

Transparency and Copyright rules for all GPAI providers: documenting models, information to downstream providers, policies for complying with EU copyright law

Safety and Security measures for systemic-risk GPAI models (GPAISR), covering risk identification, assessment, mitigation, and internal governance.

Emphasis on **proportionality to risks** and **continuous updates** to reflect technological advancements.

DEFINITIONS

General Purpose AI (GPAI)

an Al model, including where such an Al model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks

Systemic Risk

- specific to the high-impact capabilities of general-purpose Al models, with significant impact on the Union market due to their reach, or actual or reasonably foreseeable negative effects on...
- public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain;

RECITAL (117)

By means of implementing acts, the Commission may decide to approve a code of practice and give it a general validity within the Union

THE AI OFFICE

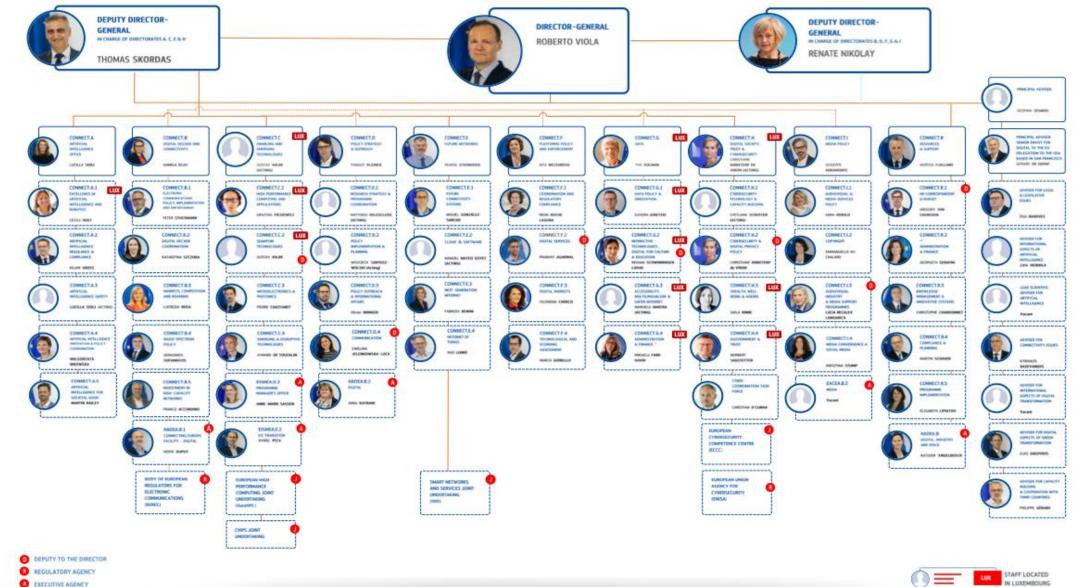
Article 50

Transparency obligations for providers and deployers of certain AI systems

• • •

7. The **Al Office** shall encourage and facilitate the drawing up of codes of practice at Union level to facilitate the effective implementation of the obligations regarding the detection and labelling of artificially generated or manipulated content.





JOINT UNDERTAKING



DEPUTY DIRECTOR-GENERAL

IN CHARGE OF DIRECTORATES A, C. E & H

THOMAS SKORDAS



DIRECTOR-GENERAL ROBERTO VIOLA

CONNECTS

PLATFORMS HOLD'S

NAME AND CONCERNATION.

INTA WEST-MOSTO

CONNECT F.1

RESIDENTIA

COMPLIANCE

mint wince

S. Print Print

.......

......

CONNECTES

DESTRUCTION OF THE PERSON NAMED IN

PERSONAL CHARGES

CONNECTE2

INGITAL SERVICES

PRAINA? AGAINMA.

COMEWATION AND

,-----





CONNECTA MARKET LANGE BY

DECKLA MOU





COOKE MADE



MEDITORIAL PROPERTY. HEALTHAND IN COMPLIANCE





COMMECTA 3: MATERIAL PROPERTY.

SHEKLA SHIKE OCTROL



COMMECTA & ACCRECATE MARTINGS SHARING A POLICY. Company Com-

MINISTERNAL STATE WHITWING



COMMETTAS WITH DEPOSIT FOR 100014-0008 DESCRIPTION BASISTY.



DESTRUMENTS ATTEMPOR Marinet Liberary LCY, MYLDMONTHINE IN DRIVING SHARKS

TER-ETHERMORE ---------

MMECT 8.2 IFM, DECKER DESCRIPTION OF

DEMOCT 8.3

ACCRECATION CO.

KORDA MOSA

DAMPHICT BLA

HOME

STAMPACES.

MARCTAL

OFFICE AND ADDRESS.

DE-CHENCES

INCO ACTIVIDADO

--------ADEA R.L.

ARRECTIVE DURING

CALIFFE - BROOKS

Families 5

MANUFACTURE OF SE

MAKETS, TEMPETTHON

INDEXES SECTION



WEIGHTLECTRONICS & deal County

COMMECTIC

Devision (MI)

TELEMONORS

BUTTER WATER

CONNECTES

(THEO, POINT, AME

APPLICATIONS.

TELEMONORES.

DETAY ANDRE

HICK PERSONALINES

SALTHA PESATNET

COMMECTIC 2 LUX

SERVICES.

and Trees.

PERSONAL CONTINUES.

~~~~~~~~~~~~~~~~~~~~~~ 



EFFMEAD 2

MARKAGET'S SPRING

ARREST SASSESS

**PRODUCTION** 





CONNECTO PRINTY SPAINERY & SUPPLIANT.

THRESE BLONCK

COMMECTICAL

THE REAL PROPERTY.

and Tombs

PLANNES.

•

ATTURN

-----

COST, INC.

.......

DOMESTIC:

CONNECT 0.2

IMPLEMENTATION IS

WOLDSON SMANSE

WOUTH Definal

CONNECTES

PERSONAL PROPERTY.

& SCHOOL COME.

Christ Belleting

CONNECT.D.4

COMMERCATION

HILDWOWSKA GROX.

MERCHANIST &

WITHOUT WILLIAMS

......



COMMECTE 1 Printer. A Description of the last of t 100046

CONNECTE

**PURPLE METATORIS** 

PLANC ENDMANDE

WORLD WHEREIT MACRO .......



GOLD & SOFTWARE

PRODUCTS, MARKED BUTTET (ACTIVAL)



OCT LINES AMBRICA DEDMIN

......

.......



CONNECTE.4 WITHMET DR Terrent la THE RESERVE



,-----COMMECT F.A. TECHNOLOGICAL AND ACCUMENT NAMED SHOWING

·----



COMMITTE.



THE VOLUME

CONNECT 6.1

BATE POLICY &

DODBY AWKIEM

manufacturing to be pro-

-----

-----

messactive.

A STREET, STORY

-----

------

SAIDAR.

75 Deep 20075

GREATAL FOR DISCRIME

Name Of Street, or other Designation of the Owner, where the Publishers of the Owner, where the Owner, which is the Owner, where the Owner, which is t

CONNECT-6-3

SAFER INTERNET

METINGS

-------------------

DATES NAMED IN

CONNECT S.A

ACRES FACE FACE

HERATE FROM

A TRANCE "

mode.

\-----

WELL THE MISSION WAS IN

LHOU

COHNECT 6.3



CHECKET (DROTHROE) MINISTERF N WHERE OR THREE



(RESOURCE MADING Delitional Windowski

(A) TIME

·---------

CONNECTING. CHRISCOUNTY &

INGINAL PRIVACY

CHRISTIANI MINETE the Williams



BONG & AGENC

-----

SHOULD RESIDE

\_\_\_\_\_ ......



SASSINITION -----

COORDINATION TO PURCE

CHRISTIAN STORES

SURCESAN

CYBDISECURITY.

HADEA BLZ Shirtne

AND WATER



#### ABOUT THE PROCESS

- Focus on General Purpose AI, and in it specially when they have systemic risk
- A multi-stakeholder consultation with over 400 submissions
- Four specialised working groups led by Chairs and Vice-Chairs selected for their expertise by Al Office, around 1000 members
- Experience, independence, geographical and gender diversity
- Discussions and drafting sessions to be held between October 2024 and April 2025

### Working methodology

- Plenary for all participants (around 1000)
- Working Group I Transparency and copyright-related rules
- Working Group 2 Risk identification and assessment for systemic risk
- Working Group 3 Technical risk mitigation for systemic risk
- Working Group 4 Governance risk mitigation for systemic risk

#### CHAIRS OF THE COP PROCESS

Working Group 1, Transparency

Nuria Oliver (Co-Chair) & Rishi Bommasani (Vice-Chair)

Working Group 1, Copyright

Alexander Peukert (Co-Chair) & Céline Castets-Renard (Vice-Chair)

Working Group 2

Matthias Samwald (Chair), Marta Ziosi & Alexander Zacherl (Vice-Chairs)

Working Group 3

Yoshua Bengio (Chair), Daniel Privitera & Nitarshan Rajkumar (Vice-Chairs)

Working Group 4

Marietje Schaake (Chair), Anka Reuel & Markus Anderljung (Vice-Chairs)

# PUBLISHED TIMELINE MARCH 2025

#### **December 2024**

- Al Board debriefed on progress with the Code of Practice (CoP) vites Al Office, Chairs and Vice-Chairs
- Al Office publishes the second draft of the CoP (first draft)
- EU Survey launched simultaneously for participant feedback

#### January 2025

- Working Group Meetings
- provider workshops
- Al Board GPAI subgroup meeting with Chairs

- European Parliament
- Summary Plenary

#### March 2025

- Third draft of the Code of Practice
- EU Survey for participant feedback (coming shortly)
- March 2025 (dates to be confirmed)
- Working group meetings
- Fourth provider workshops
- Al Board GPAl subgroup meeting with Chairs (to be determined)
- Summary plenary

- March 2025 (date to be confirmed)
- Al Board full meeting
- From May 2025
- Final version of the first CoP to be presented in a Closing Plenary and published
- AI Office and AI Board assess the CoP and publish the assessment
- The Commission may approve CoP via an Implementing Act

#### NO NEWS SINCE MARCH 2025

New deadline "by August 2025"

#### THE OBJECTIVES

- Alignment with EU Principles and Values
- Alignment with the Al Act and International Approaches
- Proportionality to Risks
- Future-Proof
- Proportionality to the size of the general-purpose AI model provider
- Support and growth of the ecosystem for safe, human centric and trustworthy AI
- Innovation in AI governance and risk management

#### **GENERAL**

- I. Assisting providers of GPAI models to effectively comply with their obligations under the AI Act
- II. The Code should also enable the Al Office to assess compliance of providers who choose to rely on the Code to demonstrate compliance
- III. Assisting providers of general-purpose Al models to effectively **keep** up-to-date technical documentation of their models and
- IV. to effectively ensure a good understanding of general-purpose Al models along the **entire Al value chain**.

#### **SPECIFIC**

- V. Assisting providers of **GPAI** models to effectively **comply** with Union law on **copyright** and related rights and **increase transparency** on the data that is used in the pre-training and training of general-purpose AI models.
- VI. Assisting providers of GPAISRs to effectively and continuously assess and mitigate systemic risks, including their sources, that may stem from the development, the placing on the market, or the use of GPAISRs

#### **CONTENTS**

#### **GPAI**

- Transparency
- Copyright

#### **GPAISR**

Risk management

# TEMPLATES AND METHODS FOR TRANSPARENCY

- Measure I.I.2. Providing relevant information
- Measure I.1.3. Ensuring quality, integrity, and security of information
- A Model document form to be filled

**ALL GPAI** 

# TEMPLATES AND MEASURES FOR A COPYRIGHT POLICY

 Measure I.2.1. Draw up, keep up-to-date and implement a copyright policy

•

 Measure I.2.5. Mitigate the risk of production of copyrightinfringing output

**ALL GPAI** 

#### MEASURES FOR RISK MANAGEMENT

- I. Systemic risk acceptance criteria, systemic risk tiers, and forecasting
- 2. Systemic **risk assessment** and decision to proceed
- 3. Technical systemic risk mitigations
- 4. Governance risk mitigations

**GPAI WITH SYSTEMIC RISK** 

## HOW IS THIS RELEVANT FOR ACTUARIES?

- I. We know about Codes of Conduct and ethics, and we can be of assitance
- Actuarial tools may fall in the remit of the Al Act and the CoP requirements

#### **ACTUARIES AND CODES**

- Actuaries have decades if not centuries of experience with data and ethics...
- We also have a well established code of conduct.
- International and European standards of actuarial practice
- Actuaries have been dealing with data for as long as the profession has existed, with increasing amounts of data – we have had to establish rules for data ethics as we deal with personal and private data
- ACTUARIES ARE UNIQUELY POSITIONED TO BE OF ASSISTANCE

# HOW IS THE CODE RELEVANT FOR THE INSURANCE INDUSTRY?

CHAPTER III

#### **HIGH-RISK AI SYSTEMS**

SECTION I

Classification of AI systems as high-risk

Article 6

Classification rules for high-risk AI systems

•

2. In addition to the high-risk Al systems referred to in paragraph 1, Al systems referred to in Annex III shall be considered to be high-risk.

# HOW IS THE CODE RELEVANT FOR THE INSURANCE INDUSTRY?

#### **ANNEX III**

High-risk AI systems pursuant to Article 6(2) are the AI systems listed in any of the following areas:

- •
- 5.Access to and enjoyment of essential private services and essential public services and benefits:
- - (c) Al systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance;

#### **OBSERVATION**

Would some insurance companies with significant market share in life or health underwriting be within the scope, to be employing a GPAI for pricing and risk assessment?

GPAI definition →

- Not likely but not to be forgotten -

STATUS QUO

• So where are we now?

#### CHALLENGES

- How to ensure compliance without creating an impossible bureaucratic work load for the launchers of AI in Europe?
- First drafts were very prescriptive to the level of hindering operation in Europe
- The focus was not on proving compliance, but rather on documenting a process
- The KPIs and documentation requirements were very labor intensive, yet CSOs wanted more
- The CSOs insisted on access basing on transparency

#### However:

- The development from first to third draft was positive. But how is it reflected in the final version
- The requirements for GPAISRs are extremely heavy but only in rare cases required

### CONCLUSIONS ABOUT THE 3RD DRAFT

- Providers continue to perceive complying with the Code very labourintensive,
   whilst have seen definite positive development on less prescriptive measures.
- SCO continue to insist on more detail, transparency, access to the model and its
  data and inclusion of fundamental rights into the list of systemic risks.
- The part of the Al Act that is concerned with obligations for general-purpose Al models (GPAI) will become effective on **August 2, 2025**, except for some areas on **November 2025**.

#### IMPROVEMENTS IN THE 3RD DRAFT

#### The third draft of the CoP showed improvements

- Says only very few companies could be in the scope of its obligations: "5-15 companies"
- Focus on required outcomes more than on techniques to be used
- From "prescritive" to "descriptive" approach
- Flexible assessment timing: from every 6months to before deployment" and when there are "significant changes"
- Model evaluations to be carried out "at the most appropriate times"
- Model evaluation criteria emphasise model safety and security rather than allowing wide access to outsiders
- Introduces the term "Safety and Security Framework"

### PUBLISHED SAFETY AND SECURITY FRAMEWORKS:

Trend is towards more transparency and accountability from leading Al developers in this critical area

The Al Act CoP is bound to follow...

- Anthropic
- OpenAl
- Google DeepMind
- Microsoft
- Meta
- Cohere Magic
- NAVER
- **G42**
- Amazon
- xAI
- Nvidia

#### SINCE THEN

- On Tuesday 22 April 2025, the Commission launched a <u>targeted</u> <u>consultation</u> to inform upcoming Commission guidelines to explain concepts underlying the provisions in the Al Act on GPAI models.
- The Al Office invited contributions to further clarify the scope of GPAI rules within the Al Act, by May 22, 2025.
- All interested stakeholders, including providers of GPAI models, downstream providers of AI systems, civil society, academia, other experts, and public authorities, were invited to provide feedback via the survey.

### SINCE THEN FROM THE AI OFFICE

- April 25, 2025 The Al Office published preliminary guidelines clarifying the scope of obligations for providers of GPAI models.
- June 6, 2025 Commission launched a public consultation on the classification of AI systems as highrisk, running until July 18, 2025. This is a crucial step for clarity for businesses.
- The final General-Purpose Al Code of Practice and the Commission Guidelines on general-purpose Al is expected to be published ahead of August 2025.

### WAITING FOR THE FINAL CODE!