DIGITAL SERVICES ACT : The European content moderation framework

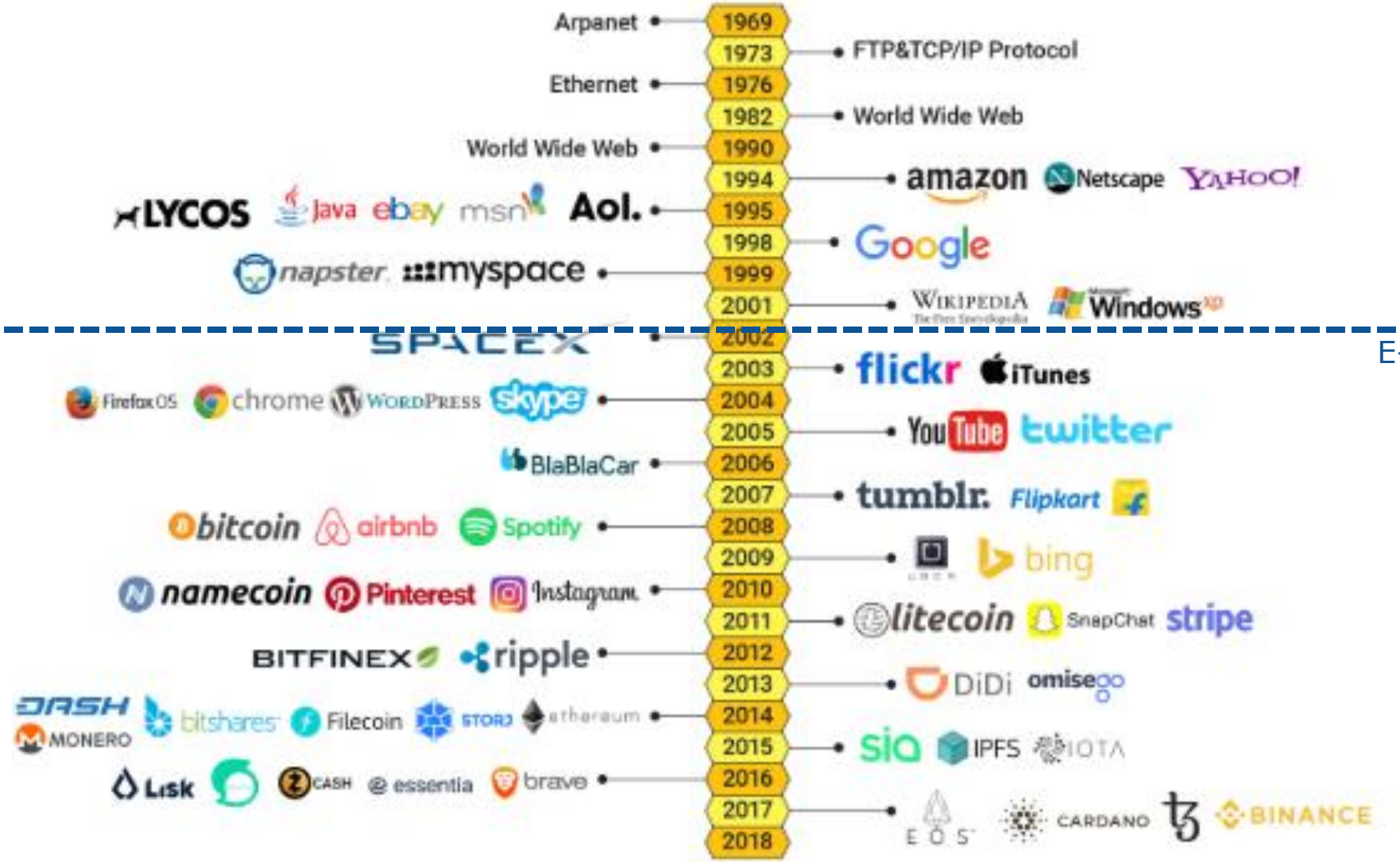
Training for the DSA team

The E-commerce Directive

The e-commerce Directive is the cornerstone legal framework for all digital services. The key rules on intermediaries liability:

- Article 14 safe harbour, and
- Article 15 no ex ante monitoring.

Case C-324/09 L'Oréal – Ebay. Article 14(1) of the E-commerce Directive must be interpreted as applying to the operator of an online marketplace where that operator has not played an active role allowing it to have knowledge or control of the data stored.



E-Commerce Directive adopted in the EU

European Commission

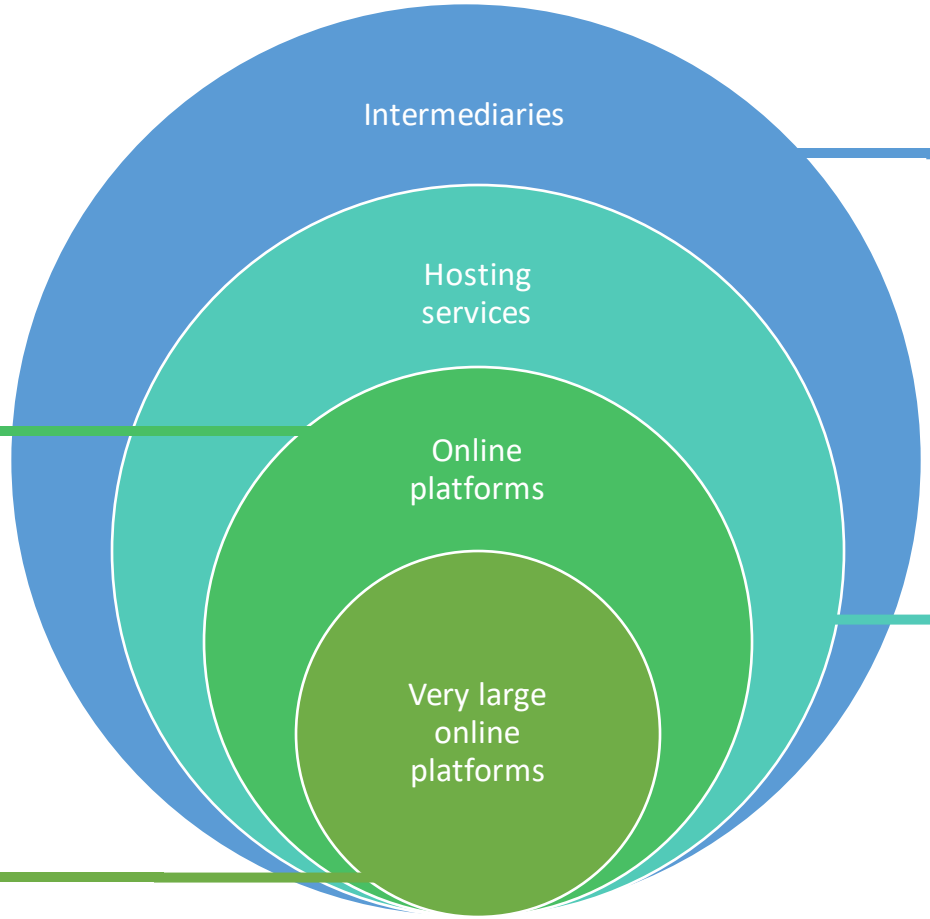
The DSA complements and updates the e-commerce Directive

1. to create a **safer digital space** in which the **fundamental rights** of all users of digital services are protected

2. to establish a **level playing field** to foster **innovation, growth, and competitiveness**, both in the European Single Market and globally

European Commission

Intermediary services:



Online marketplaces, app stores, collaborative economy platforms, social networks

Online platforms with over 45 million users in the EU

Internet access providers, domain name registries

Cloud services, webhosting

European Commission

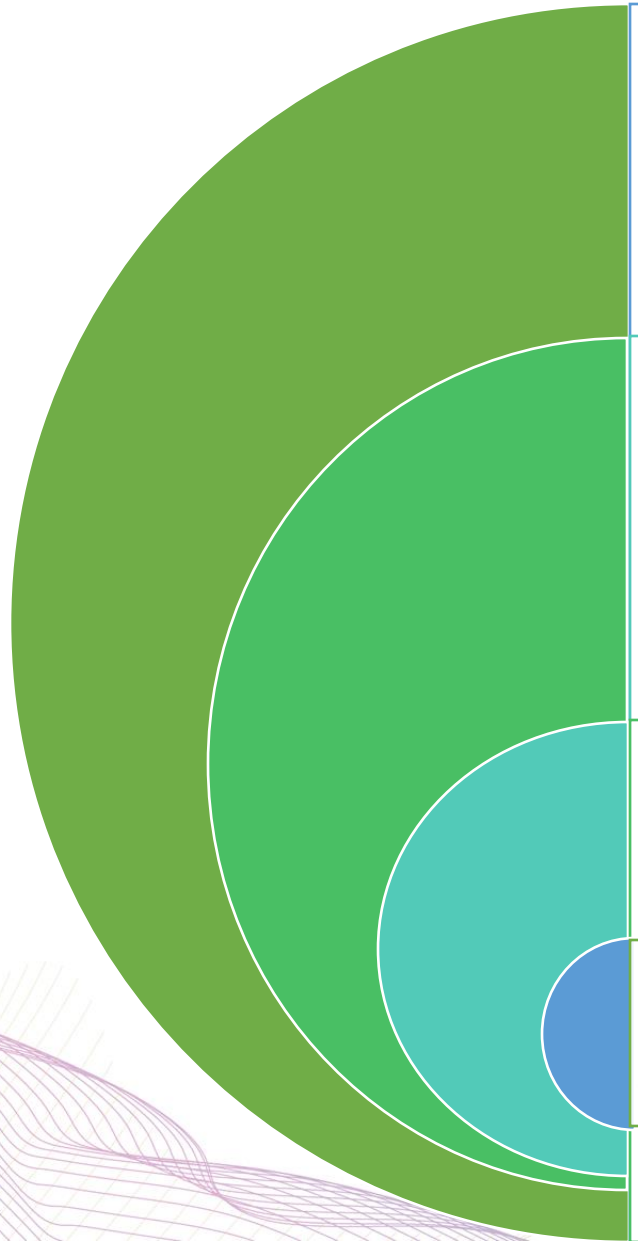
Digital Services Act risk-based design

Asymmetric obligations crafted according to size and impact of digital service provider.

Basic obligations applicable to all intermediaries.

Stricter obligations for online platforms and hosting services providers.

Proportionate, risk-based approach for the DSA:



Very large online platforms

- **Risk management, crisis response & audits**
- Recommender systems: choices
- Ad repositories
- Data access for researchers and supervisory authorities
- Compliance officer
- Further transparency reporting

Online platforms

- Internal & out of court complaint systems
- Trusted flaggers
- Limiting misuse
- Obligations for marketplaces
- Advertising transparency and bans on certain targeted ads
- Transparency of recommender systems
- Child protection measures
- Bans on 'dark patterns'
- Enhanced transparency reporting

Hosting services

- Notice & action
- Information to notice-providers
- Information to content provider
- Suspicious criminal evidence

All intermediaries

- Points of contact & legal representatives
- Clear terms and conditions & diligent, objective, proportionate enforcement
- Transparency reporting

Risk management obligations for VLOPs and VLOSEs

VLOPs and VLOSEs designation

On 25 April 2023 the Commission designated 17 VLOPs and 2 VLOSEs that reach at least 45 million monthly active users. These are:

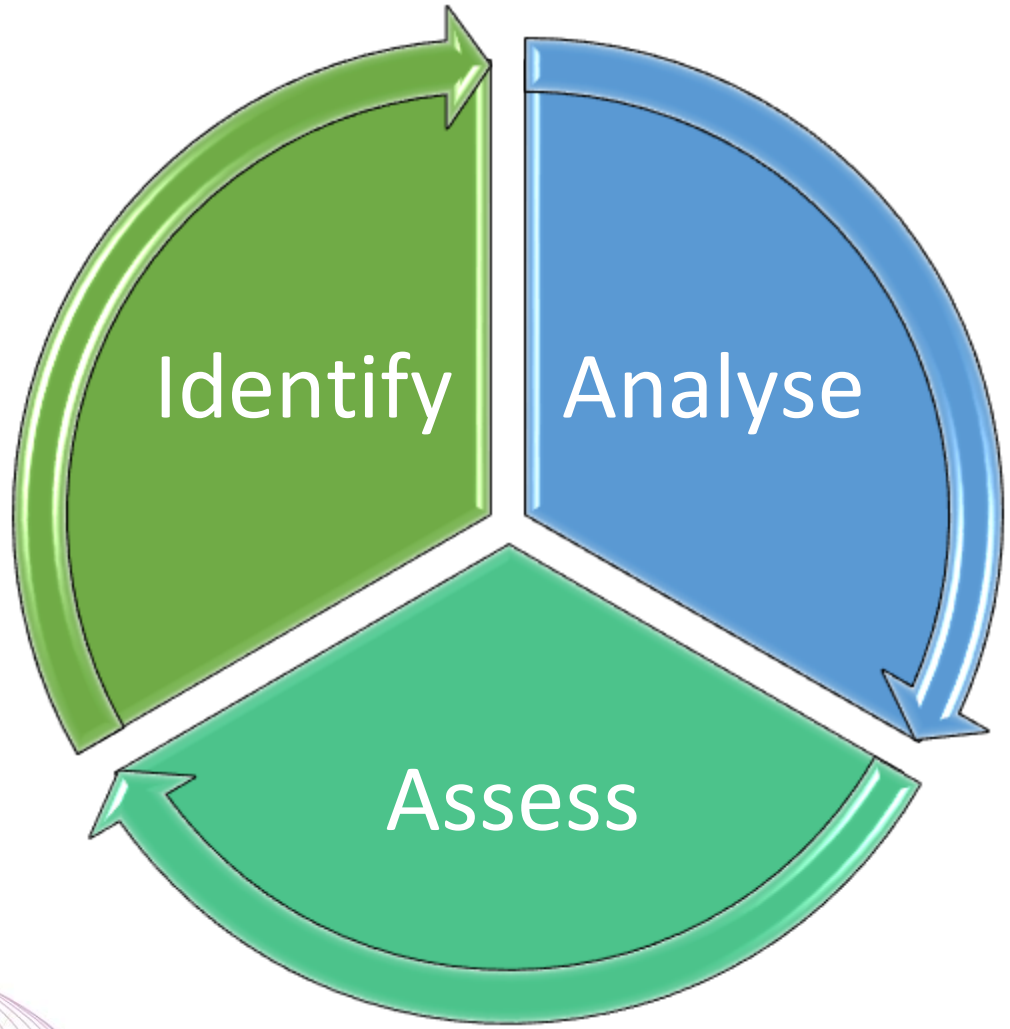
Very Large Online Platforms:

Alibaba, AliExpress, Amazon Store, Apple, AppStore, Booking.com, Facebook, Google Play, Google Maps, Google Shopping, Instagram, LinkedIn, Pinterest, Snapchat, TikTok, Twitter, Wikipedia, YouTube, Zalando.

Very Large Online Search Engines:

Bing and Google Search

3 key steps:



1. **<u>Comprehensive</u>**: systemic risks stemming from the design, functioning and use of the service, at least once per year
2. **<u>More specific</u>**: prior to deploiying functionalities that are likely to have a critical impact on the risks identified

European Commission

a) Dissemination of illegal content

- All types of illegal content, **as defined in EU and national law**. E.g.:
  - the sale of illegal goods, such as dangerous goods, unsafe toys, illegal medicines, counterfeits, scams and other consumer protection infringing practices, or even wildlife trafficking, illegal sale of protected species, etc.;
  - the dissemination of illegal content such as child sexual abuse material, terrorist content, illegal hate speech and illegal ads targeting individuals, IPR infringing content, etc.;
  - the provision of illegal services such as non-compliant accommodation services on short-term rental platforms, illegal marketing services, services infringing consumer protection provisions, etc.

b) Actual or foreseeable negative effects on the exercise of fundamental rights

- Some examples:
  - **Human dignity**: through online behaviours such as as cyberbullying, hate speech, online harassment, doxing, revenge porn, privacy violations, or other form of behaviour that undermines an individual's dignity.
  - **Freedom of expression and freedom to conduct a business**: risks from abusive notices used to silence speech or affect competing businesses.
  - **Other risks to freedom of expression and media freedom**: through content moderation practices, delisting and demotion, suspension of accounts, etc.
  - **Non-discrimination**: design of recommender systems and targeted advertising systems discriminating vulnerable groups
  - **High level of consumer protection**: use of dark patterns to mislead consumers
- 'Cumulative' risk assessments: checks and balances in the assessment and mitigation of measures, with due regard to fundamental rights

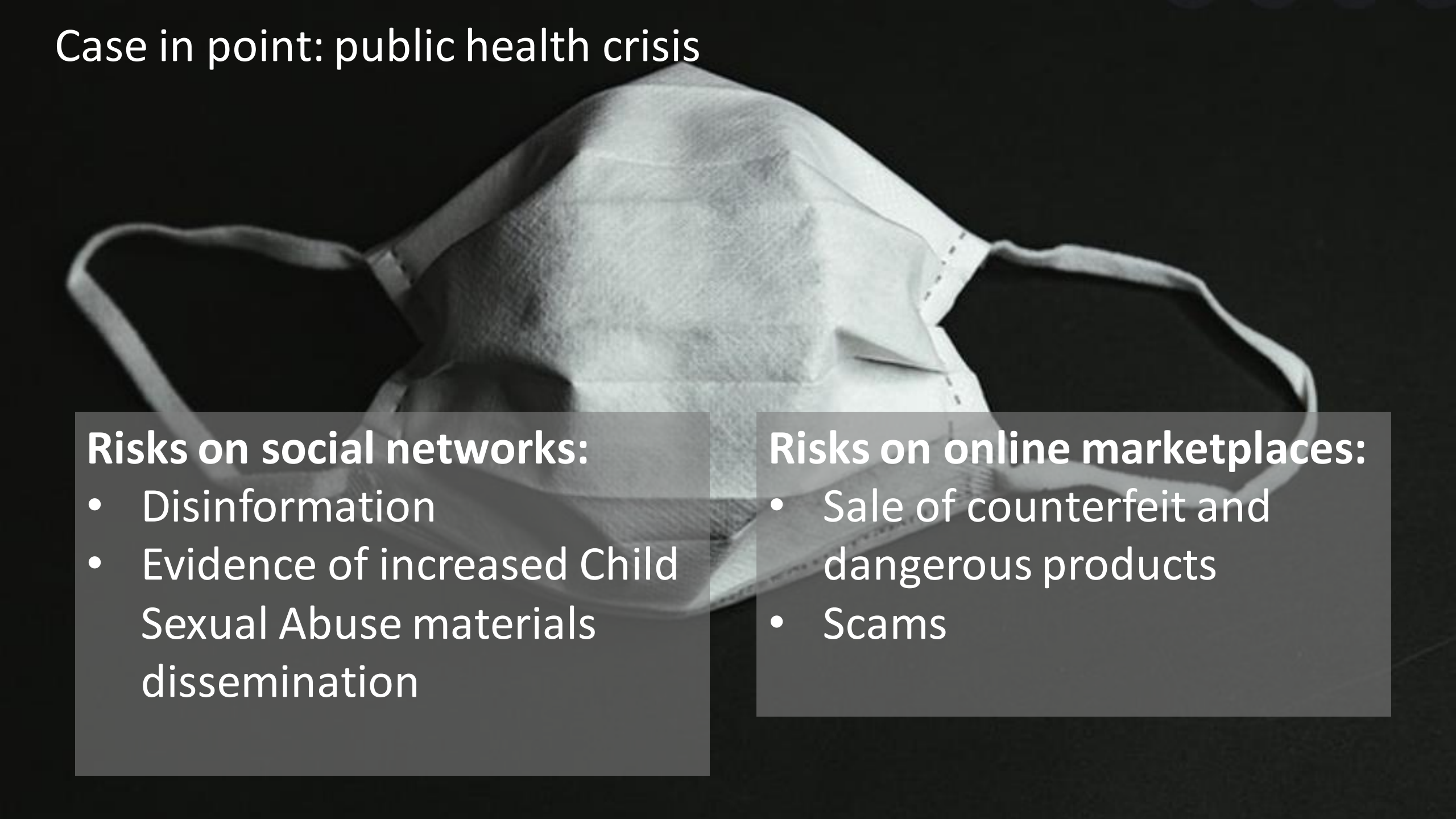
c) Actual or foreseeable negative effects on civic discourse and electoral processes, and public security

- Key provision to tackle:
  1. **misinformation** - false or misleading content shared without harmful intent though the effects can still be harmful, e.g. when people share false information with friends and family in good faith;
  2. **disinformation** - false or misleading content that is spread with an intention to deceive or secure economic or political gain and which may cause public harm;
  3. **information influence operation** - to coordinated efforts by either domestic or foreign actors to influence a target audience using a range of deceptive means, including suppressing independent information sources in combination with disinformation; and
  4. **foreign interference in the information space -** often carried out as part of a broader hybrid operation, can be understood as coercive and deceptive efforts to disrupt the free formation and expression of individuals' political will by a foreign state actor or its agents.
- Public security
  - e.g. malicious use of the services to perpetuate serious theats to public security, for example through violent extremism, terrorism, organised crimes, armed conflicts, etc.

d) Actual or foreseeable negative effects in relation to gender-based violence, public health, minors, serious negative consequences on physical and mental well-being

- Wide-range of issues, with an evolving evidence-base
- Key provision for the protection of children, together with Art 28 provisions
  - Is the service used by minors? How does the design, functioning and (mis)use of the service puts minors at risk?
  - Can minors understand the design and the functioning of the service and controls around it? Can they be exposed to content that may impair their health, physical and mental development?
  - Addictive behaviours, eating disorders, other mental and physical health issues?
- Gender-based violence
  - Do recommender systems amplify public incitement to violence and hatred based on gender?
  - Is their platform use for sharing non-consensual intimate materials or synthetic materials? Cyber-stalking and harassment?

Case in point: public health crisis



Risks on social networks:

- Disinformation
- Evidence of increased Child Sexual Abuse materials dissemination

Risks on online marketplaces:

- Sale of counterfeit and dangerous products
- Scams

Risk mitigation obligations

- Obligation to ‘put in place risk mitigation measures that are:

- Tailored to the specific systemic risks identified
- Examples of (categories of) risk mitigation measures included in Article 35

Examples of risk mitigation measures

- Design, features, functioning of the service, e.g.:
  - redesign interfaces to exclude dark patterns and offer informed choices
- Adapting terms and conditions and their enforcement, e.g.:
  - Address gender-based violence or child protection
  - Readjust resources allocated to the enforcement of T&C
- Adapting content moderation processes, including the speed and quality of processing notices, as well as decision-making processes and dedicated resources e.g.:
  - training of content moderators, correcting internal processes that optimise speed but reduce the diligence in the analysis of notices, monitoring controls for effectiveness and supplementing resources, where needed

Competences

- COM has sole competence to enforce Art 34, 35, 41, 42 (and the entire Section 5 – specific obligations for VLOPs and VLOSEs)
- DSCs may request COM to assess specific matters (Art 65(2))
- COM may use further support and expertise from DSCs: shared expertise and capability (Art 64)
- Full investigatory and enforcement powers:
  - Requests for information, inspections, access to data and algorithms
  - Interim measures – urgency due to the risk of serious damage for users
  - Commitments
  - Monitoring actions
  - Infringements and action plans + audits

Thank you!