



European Center for  
Not-for-Profit Law

**moz://a**

# **Online Platforms' Risk Assessments in Theory and Practice**

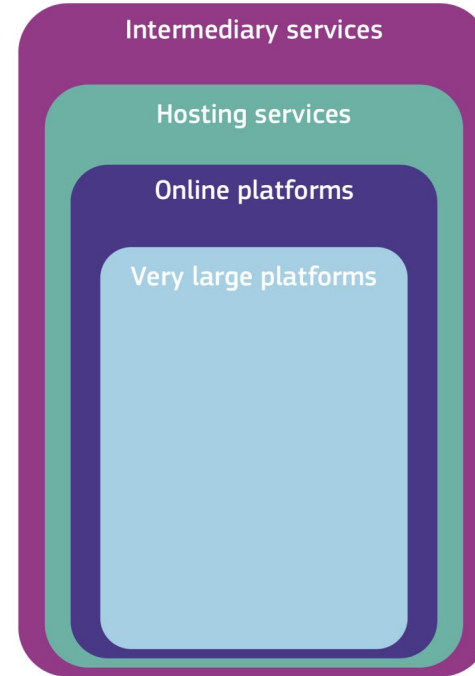
## **A presentation on the EU's Digital Services Act**

Online Platforms' Risk Assessments in Theory and Practice. A presentation on the EU's Digital Services Act © 2023 by ECNL and Mozilla is licensed under CC BY-SA 4.0. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

# The EU's Digital Services Act

## Key objectives:

- User empowerment
- Transparency and accountability
- Due process in content moderation
- Scrutiny by researchers
- Protection of minors



# Companies designated as VLOPs/VLOSEs as of October 2023

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

Bing  
Google Search

# The Obligation of Risk Assessments & Mitigation

- VLOPs/VLOSEs have to conduct yearly self-assessments into any “systemic risks” stemming from the design or the use of their service
- They must adopt “mitigation measures” to address these risks, with particular attention to fundamental rights impacts
- Consultation with users, impacted groups, independent experts, and civil society organisations is encouraged
- Risk assessments will be subject to external independent audits and they will also be subject to scrutiny by the European Commission

# What are “systemic risks”?

## **Systemic risks:**

- The dissemination of illegal content through their services
- Negative effects for the exercise of fundamental rights
- Negative effects on civic discourse and electoral processes, and public security
- Negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person’s physical and mental well-being

## **What needs to be taken into account when assessing systemic risks?**

- The design of the platforms’ recommender systems or other relevant algorithmic systems
- The design of content moderation systems
- Applicable terms and conditions and their enforcement
- Systems for selecting and presenting advertisements
- Data practices
- Specific regional or linguistic aspects
- Intentional manipulation of the service

# Involvement of external stakeholders

(90) Providers of very large online platforms and of very large online search engines should ensure that their approach to risk assessment and mitigation is based on the best available information and scientific insights and that they test their assumptions with the groups most impacted by the risks and the measures they take. To this end, they should, where appropriate, conduct their risk assessments and design their risk mitigation measures **with the involvement of representatives of the recipients of the service, representatives of groups potentially impacted by their services, independent experts and civil society organisations**. They should seek to **embed such consultations into their methodologies** for assessing the risks and designing mitigation measures, including, as appropriate, surveys, focus groups, round tables, and other consultation and design methods. In the assessment on whether a measure is reasonable, proportionate and effective, special consideration should be given to the right to freedom of expression.

# Enforcement Structure

“pan-European supervisory architecture”

- Digital Service Coordinators in each member state (by Feb 17, 2024)
- European Board for Digital Service Coordinators
- European Centre for Algorithmic Transparency (ECAT)
- European Commission DSA Enforcement Unit – this Unit will be primarily responsible for DSA enforcement vis-a-vis VLOPs/VLOSEs

# Scenario: Swapping.com

Swapping.com is a global company with its **European headquarters in the Netherlands**. It is a platform that connects individuals to one another so they can **exchange homes for short stays**. Users with accounts can promote their listings publicly on the platform and these listings are **recommended to other users by an algorithm**. Before listings are made publicly available, Swapping **uses an algorithm to verify** if the listing complies with terms and conditions and whether it does not constitute illegal content. Users can send **direct messages and they can create groups** around topics. Users can also **post reviews which are moderated** by Swapping before they are made visible to other users. Their platform allows users to **make payments using a third party payment platform**. The platform takes a small percentage of each exchange as part of their revenue model but most of their revenue comes **from advertisements** on the platform.

Swapping has **100 million active monthly users in the EU** and was recently designated as a “Very Large Online Platform” under the Digital Services Act. They are in the process of conducting their **first Risk Assessment**.



# Scenario: Swapping.com

In their Risk Assessment they are aiming to consider local/national risks. So far an internal team within the company has identified the following as systemic risks in the Netherlands.

1. **Illegal Content:** Child Sexual Abuse Material, Terrorist Content, Illegal Materials: the platform allows photo and video sharing in messages and groups and despite their use of the state of the art hashing tools, they know that their platform is abused by bad actors. There have been cases of purchasing or exchanging weapons and illegal drugs.
2. **Harmful Content:** scams, harassment, abuse: Users in the Netherlands frequently post fake or misleading listings and are able to trick others into sending payment. Less often but still frequently users will receive abusive messages or even coordinated harassment campaigns which might include sharing their personal details on other public fora (doxxing).
3. **Privacy and Security:** The platform collects personal information about users which it shares with a variety of third parties like advertisers. Users agree to terms and conditions and must agree to a new set of terms and conditions before using the payment platform, but the average user spends .3 seconds reading the lengthy contract. There is no version of the service without advertising. In addition, the platform is a target for hackers because of the data it collects and has suffered data breaches in the past compromising individuals' privacy and security.

# Scenario: Swapping.com

**4. Disinformation:** The Topics groups are often formed around topics like Covid-19 disinformation or other harmful conspiracy content. In the Netherlands the company only has moderation capacity in English and Dutch. However, content moderation algorithms work better for English and are not as effective for Dutch. The number of human moderators is also higher for English than Dutch and those speaking Dutch tend to be overwhelmed with their workload.

**5. Bias:** Some users report being ignored or denied swaps unfairly. Some also report that their listings are not promoted as much as similar listings. Racialised, gender-based, and sexual orientation bias is suspected on the part of other users and by the platform's recommendation system.

**6. Right to Housing:** The platform operates in cities suffering from housing shortages and there is research linking the rise in use of the platform with increased rent prices and housing precarity.

# Scenario Activity

When assessing the severity of these risks and forming mitigation measures, who should Swapping.com consult in this process and how? Be as specific as possible.

# Scenario Activity

**Scenario 1:** A Dutch climate justice advocacy group approaches the platform with evidence that the platform has contributed to increasing the pollution and CO2 emissions in the city of Amsterdam by encouraging short term tourists who fly, take cars, and produce trash in a disproportionate amount compared to locals and non SWAPPING tourists. The group argues that the platform poses a systemic risk to public health. SWAPPING feels this is not systemic. **How should they respond to the group's concerns? How should the advocacy group proceed with their concern?**

**Scenario 2:** An organization for persons with disabilities finds that the platform's recommendation system unfairly affects the appearance of accessible listings (listings desirable for persons with disabilities). SWAPPING is a signatory of a DSA code of conduct on accessibility but they feel it is not sufficient for this risk, and they want the European Commission to issue guidelines on best practices and mitigation measures for this problem (Art. 35.3). They take this complaint to the Dutch Digital Services Coordinator. **How should the coordinator proceed with this concern? Who should be involved?**

**Scenario 3:** The mayor of Amsterdam receives complaints from a local parents' organisation that the platform is affecting the safety of children because children are using the platform to rent out apartments and hosting dangerous parties with illegal activity. SWAPPING feels this problem is related to the third party payment system which the minors have somehow been able to access and use, and not their service. **What should the Mayor do with these concerns?**

# Discussion Questions

1. How should this risk assessment be published? Who should have access to it and when?
2. How do you think that DSA Risk Assessments can be made most effective based on experience from other fields (cyber security, environmental impact assessments, business & human rights, finance)
3. In particular considering the possibility of “risk assessment washing” and limited capacity of civil society, what consultation mechanisms could be useful? (ie: an advisory body)  
What frameworks could be provided?

# Civil Society Recommendations for Fundamental Rights Impact Assessments under the DSA

## 1. GOVERNANCE

FRIAs should be governed by the EU Charter of Fundamental Rights, which encompasses and elaborates on existing human rights standards and strengthens them with strong EU enforcement powers.

## 2. SCOPE

FRIAs must identify *all* negative effects on *all* fundamental rights and freedoms listed in the EU Charter that VLOPs and VLOSEs' products, services, or processes may cause, contribute to, or to which their services may be directly linked.

## 3. DETAIL

FRIAs should contain essential information about VLOPs/ VLOSEs' systems and processes, including details of the criteria and methodologies used to determine the most pressing negative effects on fundamental rights.

## 4. TRANSPARENCY

FRIAs must be transparent and publicly available for external stakeholders to scrutinise.

## 5. HARMONISATION

Every FRIA should be flexible and "fit for purpose," while following a harmonised structure, enforced by the European Commission to ensure accurate risk identification and facilitate future oversight and accountability.

## 6. CONSULTATION

FRIAs must be informed and shaped by regular input from external stakeholders, including civil society organisations and impacted communities.

Read ECNL and Access Now policy paper outlining key recommendations for meaningful fundamental rights impact assessments under the EU's Digital Services Act:

<https://ecnl.org/publications/human-right-s-impact-assessments-are-key-effective-dsa-enforcement>



European Center for  
Not-for-Profit Law



<https://ecnl.org/>



[@enablingNGOlaw](https://twitter.com/enablingNGOlaw)



[European Center for  
Not-for-Profit Law  
Stichting](https://www.linkedin.com/company/european-center-for-not-for-profit-law-stichting/)



<https://foundation.mozilla.org/en/>



[@mozilla](https://twitter.com/mozilla)



[@mozilla](https://www.instagram.com/mozilla)