

# Whose bias is it, anyway? The need for a four-eyes principle in AIdriven cartel screening

Jerome De Cooman jerome.decooman@uliege.be 24 March 2023, Toledo (Spain)









## I. Introduction



- AI reshuffles the deck of this pursuit and evasion scenario
- On the one hand, AI allows the development of new strategy that favours Pac-Man
  - > "algorithm-based technological solutions" = "structural competition problem" (Vestager 2020)
- On the other hand, AI allows the development of new strategy that favours the Ghosts

# I. Introduction



- Amongst these tools lies AI-driven cartel screening (Huber and Imhof 2019)
  - > Flags unusual patterns that triggers the need for further investigation
- AI-driven cartel screening works
- Yet, there are at least three challenges to overcome:
  - > Data challenge: availability and quality
  - Algorithmic challenge: explicability
  - > Human challenge: cognitive biases
- AI is a pharmakon.

### Structure of the presentation



- I. Introduction
- II. Cartel Screening
  - A. Strengthening Competition Law Enforcement
  - B. Definition and Promises
- III. The Risks of Cartel Screening
  - A. Data Challenge
  - B. Algorithmic Challenge
  - C. Human Challenge
- IV. Conclusion



LIÈGE université

Cité

A. Strengthening Competition Law Enforcement

ENDVLCIS PATRIÆ SPENLAVRVCINGAT VTIPSE

B. Definition and promises



- Despite leniency programmes and screening methods, the probability of cartel detection is still around 15% (Combe 2020)
- There is however an almost ten year long political initiative to update competition law enforcement in light of digital technology
  - > 2014: "We must take much better use of the great opportunities offered by digital technologies" (Juncker)
  - > 2017: The Commission "will continue to monitor the opportunities and challenges brought by artificial intelligence solutions"
  - > 2017: Consultancy tender aimed at gathering "informed knowledge (...) about existing AI solutions" for law enforcement and particularly how AI "could potentially improve DG Competition's processes of evidence management, legal drafting and market intelligence gathering"



- > 2019: von der Leyen asked Vestager to make sure that "competition policy is fit for the modern economy"
- > 2020: White Paper on AI: how to equip law enforcement authorities, including competition law authorities
- > 2020: European Court of Auditors: scissors effect: reduction of market surveillance capacity and increasing of case's complexity
- > 2021: the Single Market Programme proposed to improve competition enforcement (art. 3(2)(a)(i)) for instance with data-gathering and analysis tools (att. 8(2)(d)(i))



- The probability of cartel detection is not exogenous and depends on competition authorities' choices (Combe 2020)
  - > The EC has finite resources
    - » The EC is entitled to give different priority degrees to complaints received (Automec)
    - » The EC is free to focus "*its enforcement resources on cases where it appears likely that an infringement may be found*." (EC Best Notice 2011)



- In light of priority and resources allocation, AI systems help the competition law authorities initiate the "*right investigation*" (von Bonin and Malhi 2020)
  - Refinement of Regulation 1/2003 ambition of "freeing up resources to focus on serious infringement" (§ 36).
  - AI systems draw the sketch of suspicious businesses by identifying cartelists' recurring characteristics or patterns (Sanchez-Graells 2019)
- "Algorithmic shift in the fight against cartels" (de Marcellis-Warin, Marty and Warin 2022)
  - > Process data quicker and more efficiently
    - $\rightarrow$  Sooner identification of market deficiencies
    - $\rightarrow$  Shift from reactive claim to proactive investigations
    - $\rightarrow$  Increases the probability of detection that increases the efficiency of leniency programmes

#### B. Cartel Screening: Definition and Promises



- How does it works?
- There is "conventional wisdom on collusion" that permits the identification of "factors that are supposed to hinder or facilitate" collusive behaviours (Tirole 1988)
  - > Structural screens: analysis of market structure
  - > Behavioural screens: analysis of the collusive methods or outcome of collusion

#### Structural screens



Structural screens		High probability of cartelisation	
Structural factors	Number of firms (concentration)	Low (high)	
	Entry barriers	High	
	Undertakings' interaction	Frequent	
	Transparency	Low demand side, high supply side	
Supply-side factors	Vertical product differentiation	Homogeneous product	
	Innovation	Low-innovative markets	
	Advertisement	Low-advertising industries	
Demand-side factors	Demand	Stable	
	Buyer bargaining power	Low	
	Horizontal product differentiation	Low differentiation	

#### Behavioural screens



Collusive markers		Collusive behaviour		
Price	Price evolution	Low variance Sharp increase in high price-cost margin Sharp decline of price followed by sharp increase		
	Product price and quality	Homogenisation through increased product standardisation and pricing formula		
	Prices across customers	Decrease of customer-specific prices		
Market shares	Sales quotas	Distribution of market shares seems more stable under collusion		
	Exclusive territories	Price increase in the home-market, export decreases		
	Customer allocation	Stable customer base		
Enforcement	Buy-back	In time t a firm A sells above its historical market share while a firm B sells below its historical market share; in t+1, A buys products from B		
	Compensation	In time t a firm A sells above its historical market share while a firm B sells below its historical market share; in t+1 the sale levels are inverted		

#### B. Cartel Screening: Definition and Promises



- Screens identify and flag "unusual patterns" (Cocciolo et al., 2022)
- Screens do not "prove collusion or manipulation" (Abrantes-Metz et al., 2012)
- Screening raises red flags that trigger the need for, e.g., dawn raids (Harrington and Imhof 2022).
- From a procedural perspective, competition law is a three-stages process
  - > Triage to identify cases worthy of close scrutiny
  - > Verification through investigation
  - > Sanction

# B. Cartel Screening: Definition and Promises



Studies demonstrate (AI-driven) cartel screening works

- > Detection of illegal agreements (Coglianese and Lai 2022)
- > Detection of corruption (e.g., in bid-rigging)
- Faster assessment of merger control (Casey and Niblett 2021)
- However, AI-driven cartel screening "still has sceptics" (Abrantes-Metz 2014)
- This solution faces three challenges







#### A. Data Challenge

Availability – Quality

### A. Data Challenge – Availability



- All digital solutions are information-dependent and are therefore "significantly affected by problems in the availability (...) of the information they rely on." (Sanchez-Graells 2021)
- Detect collusion in dataset T, trained on dataset W (same market as dataset T) or on dataset Z (comparable market) if W does not exist
- Upshot? "no data, no fun" (Sanchez-Graells 2021)
- Caveat: Problems may still arise with available large dataset (e.g., UK Screening for Cartels Tools):
  - > Training on 100 tenders involving 500 bids
  - > Distributed model



#### The Elephant Tale (Sanchez-Graells 2019)

- Implications for AI:
  - > Training on the "entire universe of data" or on statistically representative data
  - > The quality of training dataset is critical: "dirty data, bad predictions" (Richardson et al 2019)
  - > The elephant tale warns against the "distortionary effects likely to result from policy developed on the basis of poor data that misrepresents reality." (Sanchez-Graells 2019)



#### Risk of Type II error – non detection of cartel

- > "With oligopoly, everything is possible" (Stiegler 1964)
  - » Animal Feed Phosphate Cartel
  - » "None of the collusive markers identified are universal, and each must be used with caution" (Harrington 2008)
  - » "One target, one rule" (Tinbergen 1952)
- > Interaction between collusive markers
  - » In principle, demand fluctuations hinder collusion
  - » But an increase in demand fosters collusion when entry barriers are sufficiently high
- > Selection bias
  - » Are discovered cartel statistically representative of the whole population of cartel?



#### Risk of Type I error – mistakenly detecting a cartel

- > Erroneously condemn competitive behaviour
- > Waste of time and resources
- > Structural screens are particularly at risk (Harrington 2008)
  - » This is not because the market presents the structural characteristics of collusion that there is collusion
- > Screening does not distinguish between parallelism and anticompetitive behaviour
- // with an ultrasound of the thyroid that triggers the need for a biopsy?
  (Abrantes-Metz 2013)



#### Screening is a resource and data-intensive activity

- > Data obtained from undertakings: Reliable but impossible to access them without tipping them off
- > Publicly available or aggregated data: far less trustworthy (OECD 2013)
- Upshot? It is "neither productive nor efficient" to "implement screens in every market and at every moment in time" (Abrantes-Metz 2013)
- Even if reliable data available, the extensive cost of developing and implementing screens might be burdensome for certain competition authorities (Kovacic 2013)





# B. Algorithmic Challenge

Good Administration – Duty to state reasons – Explicability – Human oversight – Transparency

## B. Algorithmic Challenge – Good administration



- "Every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions and bodies of the Union." (art. 41 EUCFR)
- This applies to administrative body that has to "gather, in a diligent manner, the factual elements necessary for the exercise of its broad discretion" (Claire Staelen v European Ombudsman)
- "The official shall take into consideration the relevant factors and give each of them its proper weight in the decision." (art. 9 CGAB)
- The duty of care "although not explicitly listed in Article 41 CFR, is generally understood as a key component of good administration" (Forrester 2009)



- "The obligation of the administration to give reasons for its decisions." (art. 41(2)(c) CFR based on art. 296 TFEU)
- No decision can be made on brief or vague grounds; or without an individual reasoning
- A decision should contain the facts, the law, and the facts-to-law leap (Fink and Finck 2022)
- Rationale
  - Protect against arbitrariness (Nehl 2009)
  - > Ensure effective judicial review (art. 47 CFR)



The EC has to respect the duty to state reason (Martinair)

- > During preliminary investigations (e.g., Hoechts, Roquette Frères, Deutsche Ban)
- > And administrative procedures (Shell International; Cimentaries; Schindler)
- What about cartel screening?
  - > Useful to trigger dawn raid
  - > The duty to state reasons applies to dawn raid to some extent
  - > To be in possession of "information and evidence providing reasonable grounds for suspecting infringement of the competition rules by the undertaking concerned" (Roquette Frères)
  - > Is the cartel screening's recommendation a "reasonable ground"?



- Is red flags raised by cartel screening a reasonable grounds for suspicion? (Roquette frères)
- Is the statement of reasons "excessively succinct, vague and generic"? (Heidelberger Cement)
- It depends:
  - > Hypo 1: "The AI system said so" is definitely "excessively succinct, vague and generic" (Fink and Finck 2022)
  - > Hypo 2: if human officer is able to disclose how the different parameters were weighted and to what extent the recommendation was decisive in the final decision, then the duty to state reasons will not be infringed (Yeung 2019)



Hypo 1 or hypo 2 depends on the AI system's opacity

- > Public officer's illiteracy (Fink and Finck 2022)
- > Intrinsic opacity (Pasquale 2019)
- When neither the factors nor the weight are known, then a reference to the AI system's recommendation
  - > cannot fulfil the duty to state reasons
  - > Cannot ensure an effective judicial review of the administrative decision

# B. Algorithmic Challenge – Explicability (?)



"The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate." (HLEG 2019)

#### Dawn raid are:

- > Highly intrusive and traumatic for staff (Aslam and Ramsden 2008)
- > Sometimes conducted without judicial warrant
- The degree of explicability is expected to be high

#### B. Algorithmic Challenge – Human Oversight



- Drawing inspiration from the AI Act: ensuring human autonomy through human agency and oversight
  - > Human agency means human must be able to make informed choice
  - Human oversight means AI system does not undermine human autonomy because there is still a human-in(on)-the-loop (or in-command)
- The design of AI systems has to ensure "they can be effectively overseen by natural persons during the period in which the AI system is in use" (art. 14(1) AIA)
  - > to 'fully understand the capacities and limitations' of the AI systems (art. 14(4)(a))
  - > to be able to interpret the system's output (art. 14(4)(c) and 13(1))
  - > to be able to choose when (not) to use the AI system and when to disregard its output (14(4)(d)).

## B. Algorithmic Challenge – Transparency



- How? AI-driven cartel screenings should be designed "in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately" (art. 13 AIA).
- This requires public officers to receive "appropriate knowledge and tools to comprehend and interact with AI systems to a satisfactory degree and, where possible, be enabled to reasonably self-assess or challenge the system." (HLEG 2019)
- ▶ Users should receive "instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users" (art. 13(2) AIA)

# B. Algorithmic Challenge – Effective Human Oversight



- Public enforcers have the ability to exercise oversight in line with their mandate" (HLEG 2019)
  - > Cf. Art. 41 CFR
  - > human control on decision-making process has to remain effective (EGE 2018)
- AI systems should not "become the primary decision makers" that "take human decision making out of the process."
- Concrete human oversight prevents "the mindless rubberstamping of AI-generated proposed decision" (Sanchez Graells 2021)





## C. Human Challenge

Bias and noises - Discretion - Automation bias - Four-eyes principle



#### Debates on black box and explicability is paradoxical

- > Algorithms are criticised because opaque
- But both human beings (Thaler and Sunstein 2008) and administration are similar black boxes (Callon and Latour 2006)



- To err is human; but human also predictably err (Thaler and Sunstein 2008)
- Bias: "any systemic error that inclines people's judgements in a particular direction" (Sunstein 2022)
- Search satisfaction: stop searching once a first plausible explanation is found
- Anchoring: premature decision-making based on limited information initially available
  - > Confirmation bias: tendency to interpret information to fit preconceived opinion
  - Diagnostic momentum or hindsight bias: the pursuit of an action previously instigated by someone else without considering any new information and changing plan accordingly (a fortiori if hierarchical superior)



- Within EU competition law proceedings: Combination of investigative and decisionmaking powers
- **Prosecutorial bias:** "investigatory teams that have dedicated months finding enough evidence to support an infringement might suffer from the dreaded 'tunnel vision', which could cause them to adopt an unfair or biased decision." (Lachnit 2016)
  - biased investigation favouring information concluding to a collusive behaviour (confirmation bias) and discarding the others (hindsight bias and diagnosis momentum) (Wils 2004)
  - > Commitment bias: the unwillingness to adopt a decision that contradict what officials have done in the past due to the involvement of "both the Commission's human resources and reputation capital" (Teleki 2021)
  - Policy bias: high level of enforcement to keep-up with the statistics (Wils 2004) and because "promotion flow from taking decision" (Forrester 2013)



- Bias is systematic; noise is an "unwanted variability in judgements" (Sunstein 2022)
  - > Occasion noise: noise personal to a human: fatigue, blood sugar, local news or weather
  - > Level noise: a decision depends on the decisonmaker (lenient of severe)
  - > Pattern noise: leniency or severity depends on patterns of the decisionmaker

		Grounds			
		W	X	Y	Z
Decision-maker	Α	receptive	receptive	indifferent	indifferent
	В	receptive	receptive	indifferent	indifferent
	C	indifferent	indifferent	receptive	receptive
	D	indifferent	indifferent	receptive	receptive

### C. Human Challenge – The Exercise of Discretion



- Explicability goes beyond the algorithmic challenge.
- The duty to state reasons requires an explanation of the algorithmic operation <u>and</u> an explanation of "the influence that algorithm results have on (constraining) human decision-making" (Busuioc 2022).
- The weight of the recommendation should not be underestimated
- Going against the recommendation would require a well written reasoned decision that renders "the exercise of discretion costlier" (Petit 2018)
  - » "A hearing officer's belief that computer decisions are error-resistant increases the likelihood of inaccurate outcomes" (Citron 2008)
  - > "Computers also benefits from their traditional reputation of being intelligent and fair, making them seem credible sources of information and advice" (Fogg 2003)

#### C. Human Challenge – Automation bias



- Automation bias (or algorithmic dumbfounding), *i.e.* the irrational tendency to rely on automated decision even when the operator suspect malfunction (Goddard *et al.* 2012).
- The automation bias is the digital update of:
  - > Search satisfaction: stop searching once a first plausible explanation is found
  - > Anchoring: premature decision-making based on limited information initially available
  - > Confirmation bias: tendency to interpret information to fit the preconceived opinion
- The algorithmic recommendation is:
  - > A first plausible explanation...
  - > ... that tempts the officials to cease the scrutiny...
  - > ... and even if further investigation were to be conducted, the recommendation would serve as an anchor as any new information gathered would be interpreted as strengthening the preconceived opinion.



- This has been forecasted by the EU legislator
- "remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons" (art. 14(4)(b) AI Act)
- Effective (?) human oversight
- With this formulation, the AIA is not sufficient
- Proposal: four-eyes principle



- Four-eyes principle: A approves both the decision and statement of reasons of B
- Not alien to competition law
  - > Woodrow Wilson
  - > French Competition Authorities
    - » The Investigation Service (Le Service d'Instruction): opens investigation, gathers evidence
    - » The Board (Le Collège): takes the decision
  - > Belgian Competition Authorities
    - » The Investigation service (Auditorat)
    - » The Board (Le Collège de la concurrence)



- The bicephalic structure enhances procedural fairness (Lasserre 2009)
- The Authority is no more "the 'judge, jury, and executioner' of its own cases." (Lachnit 2016)
- Unbiasing decisionmaking: solve the commitment bias
- AI-driven cartel screening raises similar issue; calls for similar solution
- An independent team scrutinise the AI and its use: this mitigate the automation bias



- Easy to implement in France, Belgium, and other bicephalic institutions
- This will prolong the duration of case
- But that extra-time might is not wasted
- Complex to implement in "*all-in-one*" competition authority (*e.g.*, EC):
  - > No need to split DG Comp
  - > Extension of the Hearing Officer's role
    - » Already ensures the effective exercise of procedural rights
    - » Already an independent arbiter
  - > The Hearing Officer might well be the proper public overseer of AI-driven cartel screening.



# VII. Conclusion



- This paper is not a pamphlet against algorithmic screening
- The counterfactual scenario of not using these tools is full reliance on competition authorities' officials
- There is no evidence human decision is "significantly more accountable than AI" (Lim 2021)
- Both individual (Coglianese 2022) and groups (Callon and Latour 2006) are black boxes

# VII. Conclusion



- "When all possibilities (...) become probabilities, every possibility is the next thing to a certainty" (Melville, Moby Dick, 1851)
- Screening raises possibilities of collusion, nothing more (but also nothing less)
- Competition authorities have to remain aware of AIS' limitation.
  - > Data challenge
  - > Algorithmic challenge
  - > Human challenge
- If not, they might well be doomed to embody Ahab's fate, equating probabilities and certainties





Do algorithmic screening tools fall within the scope of application of the AI Act?

- "Without any prejudice to the application of Union competition law" (Explanatory Memorandum)
- > This is not a dead-end
- AI system is a software that generates either content, predictions or recommendations given a set of human-defined objectives (art. 3(1) AI Act)
- > High-risk AI system is
  - » Either covered by sectorial product legislation listed in Annex II and used as a product or a safety component (art. 6(1)(a) AI Act) for which a third-party conformity assessment is required (art. 6(1)(b) AI Act)
  - » Or not covered by sectorial product legislation but still considered as high-risk and as such listed in Annex III (arts. 6(2) and 7 AI Act).



#### Law enforcement activity mentioned in Annex III

- > The AI Act defines law enforcement authority as any public authority competent for law enforcement activities, *i.e.*, the prevention, investigation, detection, or prosecution of criminal offences (arts. 3(40) and 3(41) AI Act).
- Annex III submits to mandatory requirements AI systems used by law enforcement authorities "AI systems intended to be used by law enforcement authorities for predicting the occurrence or reoccurrence of an actual or potential criminal offence based on profiling of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 or assessing personality traits and characteristics or past criminal behaviour of natural persons or groups" (Annex III(6)(e) AI Act)
- This definition is quite close to the purpose of behavioural cartel screenings, and yet it is restricted to criminal offences



Reference to criminal offences is problematic, as competition law is qualified as criminal in some but not all Member States

- > Full criminalisation: Ireland, Estonia, Denmark, Greece and Slovenia
- > Criminalisation of specific competition law infringements in Luxembourg, Germany, Poland, Hungary, Austria, Italy, Belgium, Portugal and Croatia
- Criminalisation of competition law infringement in specific circumstances in France, Romania, Czechia and Slovakia (*caveat* Cyprus)
- No criminalisation in Bulgaria, Finland, Netherlands, Sweden, Malta, Lithuania, Latvia and EU competition law
- Upshot: The AI Act will only apply to competition law proceedings in legal orders that criminalise competition law
- How does this fit the objective of harmonisation?



Applying the AI Act to competition law through the backdoor? Criminal law if (ECtHR *Engel*)

- > Classification in domestic law as a starting point (ECtHR, Weber v. Switzerland 1990)
- > Nature of the offence
  - » Does the rule concern all citizens? (ECtHR, Bendenoun v. France 1994)
  - » Does the rule have a deterrent or punitive purpose or does it merely impose pecuniary compensation? (Ibid)
  - » Were the proceedings brought by a public authority under statutory powers of enforcement? (ECtHR, *Benham v. The United Kingdom* 1996)
  - » Does the rule at stake seek to protect general interests of society? (ECtHR, *Produkcija Plus Storitveni Podjetje D.O.O.* v. Slovenia 2018)
  - » Is the imposition of a penalty upon a finding of guilt? (ECtHR, Benham v. The United Kingdom 1996)
  - » Is the misconduct at stake classified as part of the criminal law in the vast majority of the Contracting States (ECtHR, Öztürk v. Germany 1984)
- Severity of the penalty (ECtHR, Campbell and Fell v. The United Kingdom 1984)
- Upshot? Competition law belongs to the criminal sphere (ECtHR, Société Stenuit v. France 1992; ECtHR, Lilly France S.A. v. France 2002; ECtHR, A. Menarini Diagnostic S.R.L. v. Italy 2011)



#### Within the EU

- > Competition law offences "shall not be of criminal nature" (Council Regulation 1/2003)
- > ECtHR Jussila: hard core vs peripheral criminal law
- > EU competition law is not hard core criminal law but belongs to its periphery (Bot 2010, Sharpston 2011, Kokott 2013, Wahl 2018, Bobek 2021)
- > Confirmed by the ECJ (recently, bpost 2022)
- > Competition law is "criministrative" law (Bailleux 2014)



- Q: Could the criministrative nature of competition law serve as a back-door to apply the AI Act to all competition law proceedings regardless of the domestic qualification?
- A: No
  - Contextual approach: the AI Act and SWD refer to "criminal matters" in a context of hard core criminal law
  - Coherence: "AI systems specifically intended to be used for administrative proceedings by tax and customs authorities should not be considered high-risk AI systems used by law enforcement authorities for the purposes of prevention, detection, investigation and prosecution of criminal offences" (Recital 38, *in fine*, AI Act)

#### **IV. Conclusion: Without Any Prejudice?**



- Annex III suggests AI systems intended to be used by law enforcement authorities in the course of detection, investigation and prosecution of criminal offences raise high-risk and are subject to mandatory requirements
  - > In legal orders that qualify competition law as criminal: algorithmic screening tools would have to comply with the AI Act
  - > In legal orders that do not qualify competition law as criminal: the AI Act does not apply
  - > The AI Act closes the door to an extension of its scope of application through peripheral criminal law
- As the AI Act is a harmonising regulation, keeping different standards of protection depending on national qualification makes no sense