Eric: So this week on the podcast, we had Gianclaudio Malgieri, Associate Professor of Law and Technology at the EDHEC Business School in little France. He recently wrote an excellent paper, entitled Algorithmic Impact Assessments Under the GDPR Producing Multi-layered Explanations, and that caught my attention.

He's also written a number of other articles and algorithmic decision-making AI under GDPR and federated learning. As I've said before on this podcast, the EU probably has the most comprehensive regulatory construct for privacy law in the world. It's undeniably a model, although not necessarily the model for countries, contemplating privacy policy and regularly.

Now the capabilities of algorithms processing personal information continues to get more sophisticated and it's logical for private and public entities to become more reliant on it. The challenge is that when they use personal data to generate decisions that are at odds with privacy rights, public policy, or regulation, how would we public policy makers, et cetera know? The more sophisticated the algorithms become, the more challenging it's going to be to try to reverse engineer or determine after the fact what it was doing. Even the creators of algorithms have not contemplated all the impacts. Understanding these impacts can have a multitude of benefits, including better algorithms and more aligned impacts.

Of course, if like anything, regulation can be a blunt tool with its own costs and impacts. So it's a careful balancing. Whether you believe more regulation or less regulation is needed or none, you're going to find this episode fascinating. But to take it a step further, and this is something I raised in an earlier episode with Diane Steward Mill impacts in creating or deploying algorithms and changes to them later, make sense at a business level, as well as at a policy level, no one wants biases in the outcomes of their algorithms.

And if you're not periodically assessing for those biases, particularly when you're making changes, the decisions you make from those algorithms will be. On the flip side, of course, that there is a risk that if you require a certain frameworks by regulation that might lead to the bureaucratization of them and possibly become a CYA or check the box exercise.

So again, a careful balancing act hope you enjoy this episode. And I also hope that if you really enjoyed it or other episodes that you're sharing it, that you're finding one person to share this and say, Hey, check out this this podcast. I like what I'm hearing. Welcome to The Encrypted Economy, a weekly podcast featuring discussions exploring the business laws, regulation, security, and technologies relating to digital assets and data.

I am Eric Hess, founder of Hess Legal Counsel. I've spent decades representing regulated exchanges, broker dealers, investment advisors, and all matter of FinTech companies for all things touching electronic trading with a focus on new and developing technologies.

Today, we're fortunate to have Gianclaudio Malgeiri on the podcast.

He is an associate professor of law and technology at HEC, uh, business school in France. He is a co-director of the Brussels privacy hub. He is an ethics expert for the European commission. He also, and he can correct me if I'm off on this. He conducts research on and teaches data protection law privacy, AI regulation, digital law, consumer protection in the digital market data sustainability and intellectual property law.

Excuse you got quite the full plate there. Gianclaudio, welcome. Thank you

Gianclaudio: very much, Eric. And thank you very much for having me here. I'm very happy to be here.

Eric: So I gave a little bit on your background, but do you want to give us maybe a little bit more on your path to get here, your interest specifically in you know, the GDPR and algorithmic you know, both AI and algorithmic issues raised by the GDPR.

Gianclaudio: Yeah, sure. Let's say that um, doing research in GDPR was quite amazing and exciting in the last, let's say five, 10 years in the European union, it was quite 10 obliged choice because when I started to dig into tech, low cyber law, and so on, the discussion was totally monopolized by data protection.

And I started to get fun, but data production to understand it more and more. And then yeah, basically, I've come across uh, automated decision making as the real new topic, new buzzword. Everyone was talking about explanation of AI, right? So I think that in the last years the most interesting part of my research.

Let's say for me, the most interesting parts were interred sectoral and interdisciplinary research we'd, computer scientists, psychologists, marketing scholars, because I think their lawyers need to talk with others. With other people. We usually tend to be a, in our room, we need to be with other people, and we need to talk different languages.

So, yes. So in my last year I had the fortune to. To have this discussion about explanation of AI vulnerability. My most, my mind, because to topic of research into the last years was vulnerable data subjects in Europe and beyond. And uh, so mostly, fairness, fairness of AI. General concepts.

And from, I mean, it's few months that I've been appointed co-director of the bras of privacy hub, where we try to elucidate these concepts, new policy proposals in Europe and beyond we try to be a reference for researchers, policy makers and, uh, lawyers. So yes, that's all.

Eric: And, and before we started kick into it do you have a personal experience that, uh, maybe when you know, really focused your attention on what you're focusing on now,

meaning like some personal experience where it all clicked something early on which you can think back to what, what would that.

Gianclaudio: Yes, I have some of them some are more personal and other are more professional experiences that imposed me to start at this topics. Let's say, a couple of, general experiences that I can talk about is uh, first of all, the selection of students I am a professor in France where there is an algorithm that uh, decides which kind of students can be in some high level universities.

So there are. And there was an automated process some years ago called potato soup that, uh, true AI could decide what students could be accepted in a high level school. And for me, this was the first motivation to study justification and explanation of AI. Smaller examples are some colleagues at the university that were trying to have smaller loans for buying a new computer and so on.

And they just received a no. So also the explanation of decisions was something that really motivated me and my research about vulnerable people in AI. I think it was driven by a couple of personal experience. First of all, being member of LGBT community was a big let's say. Stimulus for my research and vulnerability.

Yes, I think my research was very much influenced by my personal experience as for everyone,

Eric: guess. Great. I read your paper on algorithmic impact assessments. And we are going to drop it into the show notes. It's something that I hope that we can talk more about today in addition to some other uh, eventually to GDPR.

But do, do you want to give a little bit of a backdrop for the paper?

Gianclaudio: Yeah, sure. I think in the GDPR, one of the biggest innovations and revolutions was the, two of them one that's had a lot of attention was uh, how to address automated decision-making so how we should regulate automated decisions.

So algorithms that take decisions and how we can protect people that treat. Strong impacts that have strong impacts on the basis of disability decisions. The second great novelty of the GDPR was the impact assessment data protection impact assessment in the U S there's the environmental impact assessment and a European one was a bit built on the example of the environmental impact assessment with my quarter, professor Margo Kaminski she's a professor in Colorado law school.

We had the idea that these two tools, so regulation of automated decisions and data protection impact assessment had to talk, had to be connected. And this was essential for two reasons. First, it is good for customers, consumers, individuals, because it helps to reach better. Decisions and better explanations of decisions.

Secondly, it helps data controllers companies because they really have this two duties, protect people when there are automated decisions and do an impact assessment. If we manage to propose a model that can combine these two tools, it would be win-win for companies and for individuals. So that's why we had the idea to write this paper that we had quite fortunate so far.

Yes.

Eric: Excellent. And maybe as a bit of a precursor to that maybe you could review the evolution of GDPR regulation around the issue of algorithms processing personal data.

Gianclaudio: Sure. Yeah. Yeah, sure. So basically in the GDPR we have um, article 22 about right. Not to be subject to automated decisions.

So basically when there is an algorithm that takes a decision on individuals and when does decision. Has a legal or similarly significant impact effect. So for example uh, a denial of a bank of a request to a bank or um, accessing to a particular school or university whatever, there are many examples that we can make.

In this cases, we have a right, not to be subject to this decision unless we give explicit consent or there is a member state law that authorizes this. So for example I dunno in Germany, it is authorized that, uh, some algorithms can take decisions about insurance. So it depends on international law and a third one.

Is, uh, so basically you have, sorry. I said consent member, state law, or necessary for the contract. So the third one is contract when it's necessary for a contract in which I am a party. Okay. Uh, in this cases, there are additional safeguards that they should I have as a customer, first is to receive information in general about the algorithm. So we should understand how the AI works. Secondly we should, be able to contest the algorithm to contest the decision. Thirdly, we have a right to have a human in the loop, so we have a right to have a human that can change or. Approve the machine decision.

And the last thing is the right to say my own opinion in the in the automated decisions, in addition to these four higher risks. So for data processing, all the higher risk, for example, a window is high, a large amount of sensitive data, hospitals dating apps, and so on. I have also right to receive specific explanation of the decisions about me.

So you see it's really a risk based approach with some safeguards, right? So this is basically how did you get your protects individuals that receive automated decisions.

Eric: And there are other jurisdictions that have similar protections. Like for example, the Canadian government has I think they have an online algorithmic impact tool which is actually, I looked up quite interesting.

State run agencies and departments are required to perform. But, uh, you what other jurisdictions have something that, that even closely resembles GDP.

Gianclaudio: Um, the last part that maybe I had to clarify to last, and so was that an additional point is also auditing algorithms. This is also part of the GDPR auditing algorithm, Troy risk assessment.

And this is the link between the impact assessment and automated decision making rights and safeguards. If we can apply this model. So make having, analyzing the impact of algorithms and use this analysis as an explanation to individuals. So what is happening? This was the real core of our article. We can have the win-win situation it was explaining before.

So we have. Countries and jurisdictions where some forms of auditing of algorithms and algorithmic impact assessment is already every reality. You mentioned Canada, of course, they have a directive about algorithmic impact assessment and how they do that. Just, yeah, we can say public authorities or forest, I mean are asked to do this impact assessment of algorithms.

But they, they are, there are some easy ways to do that. There is a software in which they can um, self-assess their algorithms. This is what is missing in Europe. We have wonderful principle-based law, but we miss sometimes. Automated form of compliance or let's say facilitated form of compliance, right?

So of course we have Canada that has this example. It's a symbol to, to do a graphic impact assessment in Canada, but it's a bit a superficial approach because you cannot really go deeper and deeper because it's just a general software use for all. Public offices, then we have other examples around the ward, right?

We have the Brazilian data protection law that also has the right to receive explanation of algorithms. Um, and then we have uh, even in China, it was just approved a few weeks ago was the beginning of November that the Chinese, Dre, I mean, a good rhythmic bill was approved in which basically individuals have a right also.

To understand the logic and the functioning of the algorithms. And also, uh, there should be some form of risk assessment all around these jurisdictions. So yes, we have other existing example in the U this was just a proposal. We have the Biden proposal. Sorry. The why then proposal was a proposal about algorithmic impact assessment.

So automated decision-making impact assessment. So some assessing the impact of automated decisions. It was the first proposal of a 2018. Denny was we proposed. Congress, and it's still pending. I'm not very optimistic, but I'm enough to, a specialist of U S law. And then a, yeah, basically these are the most, interesting, existing, laws about, about explanations of AI and impact assessment.

Yes.

Eric: And the white in account, algorithmic accountability act that applies to entities that use store or share personal information to conduct automated decision system impact assessments and data protection, impact assessments. I agree. I don't think it's going to be a 20, 22, probably not a 20, 23.

But it, you I think there's someone with a shop in the U S before we get to that. But and in China is the China one is law, but does it only apply to private entities, or does it also apply to governmental entity?

Gianclaudio: Yeah, it applies mostly too. It applies to private end of this, actually. So yeah, it's the other side of the moon.

If we consider Canada that applies to also only two public entities, right? It's the fear of big brother. And in China you have just the problem of small brothers. So yeah, basically you have this two different if I can add something in the U S I forgot to mention that some states have legislations about impact assessment algorithms and so on.

So basically California, Virginia, so we have some existing proposals of legislation. So yes, I think something is coming soon, right?

Eric: Um, yeah, it's going to be difficult for her to play hopscotch with all the states in the U S which ones do and which ones don't after a point. And is, is the problem with the current construct is that it's maybe a little too simplistic in terms of the, just offering individual explanation.

The

Gianclaudio: problem of automated decisions and their explanations is that if a companies have a duty to explain the logic of complex algorithms of a black box algorithms well, um, there might be a very limited approach in complying with this duties, right? So basically what can happen is that companies might easily go with naive form of explanation.

So I can make an example if you go on Google ads, so you receive banners everywhere, right? Banner, some Facebook banners on the behavioral advertising everywhere. And then we can click. Why does ad it happens? Google has as an automated button for that. And usually the reply will receive.

Because in this time of today and in your country, in your region, in your area, this is an interesting content to share. Of course this is not a satisfactory explanation, or we can make also other examples banks, our bank takes decisions based on algorithm. And if we go on the general privacy policy of our bank, what we receive is just to some naive

explanation saying, sometimes we receive, we do some form of automated decisions to help us and to, to, to, to make our process quicker.

Of course this is not enough, right? So what we really need is to find the balance between effectiveness. So comprehensibility of explanations but also, detailed. The level of granularity. Uh, of course there should be different explanations for different audiences, right? So we should have a more specific explanations and not so easy and somber explanation for experts, our doctors using algorithms when they do the triage or when they treat COVID symptoms and whatever or when they do surgery and so on.

But there should be at the same time, some are simpler explanations where the balancing between simplicity effectiveness. And let's say completeness is more towards simplicity, of course. So there are different audiences and different kinds of simplicity we can say. But maybe what we really need is not just a ma explanation.

We need the justification. We need to understand why the algorithm is. Accurate right. Limited to what was necessary and not beyond that and so on. So yeah, explanation is okay with the different audiences, but justification is even more important.

Eric: Great. And with regards to encryption how does encryption fit into this?

Does it actually complicate the ability to make these explanations?

Gianclaudio: This is a, this is a very nice point because last year we conducted a series in Brussels. It was called the tech talks. And basically this was a common problem because encryption is a privacy announcing technology.

But at the same time, sometimes the, the sessions behind this private sentencing technology are difficult to understand for users, right? So you have a paradox. That's using more privacy friendly technologists means um, black boxes for individuals. So more difficult systems to understand.

So again, we should find the good balancing between protection. And so the effectiveness of encryption, for example, and at the same time uh, level of explanation, Of this of the systems, that there are many examples we can do. For example, secure multi-party computing or computation is a, it's a real good example in which algorithms can be used through a pet privacy enhancing technology.

But at the same time uh, average individuals could not understand how this is protective or not. And this is a concern, not big tax because big attack, if they invest the money in encryption, they also want it individuals feel safer and protected. So they also want to use privacy non-sick technology as a marketing tool.

And if it's difficult to explain, it's difficult to use it to leverage right this, this technologist. So yes, this is a very important point. And I think it's. In the future will, this discussion will be more and more important. Now it's just coming. I think in the future, we will discuss more and more.

I am sure I have no answers now, but you have you know, I agree with the question.

Eric: And I imagine that something like homomorphic encryption is probably even more challenging than secure multi-party computation, just because of the complexity and the fact that everything's encrypted and we're at least in secure multi-party computation, it, it, you it, it preserves your data set, but it doesn't necessarily limit you the same way that homomorphic really, it, it could, since everything is encrypted very hard to get behind the, um, the algorithms.

Um, can you, can you explain the concept of collaborative governance and how it fits into GDPR, and you know, your proposal around algorithm, algorithmic impact assessments?

Gianclaudio: Yeah, that was one of the concepts that in the paper where the most, let's say legal ones. So basically collaborative governance It's an ambiguous term, right?

Because in, do you, has it as a meaning in Europe and in the European union, it has another meaning. Let's say that in general, we can say that collaborative governance is when it's not just a regulator to control and check the respect of the law, the compliance with the law. And it's not even just a single company that needs to bear costs and burden of all compliance.

There's a cooperation between the regulator, the watchdog, the survey, let's say oversight authorities and the companies. So in the GDPR collaborative governance, for example, is quite evident in the data protection impact assessment. Why? Because. The concept of accountability in the GDPR is basically based on the notion of companies that are doing their self-assessment, but their self-assessments should be prudent and wise enough because the regulator could always check if they're respected the minimum threshold, the minimum level.

So it's a form of mutual trust between companies and state cooperation. There are even more advanced forms of cooperation that we can mention, even beyond the paper, for example uh, regulatory sandboxes, where basically it's the state that makes an agreement with the companies of certain sectors, for example, big tax, social media, or banks, whatever and decides some uh, their obligations from the written law.

In order to achieve a better level of a governance compliance and enforcement, the state sometimes cannot check all data processing and respect of all data protection principles, right? So they have to delegate to the private partner, some form of self-assessment. And

of course it's easier for them to check a self-assessment then starting from zero, and check everything. So this is a bit the idea of collaborative governance.

Eric: Great. And in terms of what would need to occur in order for an algorithmic impact assessment to be, required under GDPR, is there any existing guidance or making which you know, it could be implied, or would it need some more formal guidance or would it need something.

Gianclaudio: According to us. So according to my co-author and me we, uh, we already have an algorithmic impact assessment into law because basically data protection impact assessment should be performed in cases of high risk data processing activities and the amount of three cases of a higher risk data processing activities there.

So ways profiling an automated decision making. It's one of the tree. So in this case is you have to do an impact assessment. What we suggest is that since in cases of automated decisions, companies also need to give explanations contestation possibility to have human in the loop, et cetera. It would be wise to combine these two duties.

So to combine an impact assessment. There are requirements about automated decisions. So in our view, it's not a proposal for a reform. It's just a way to better understand and interpret what's already there. And the benefits are for all parties, as I already said, right? Because the benefits are for companies and the benefits are forum, individuals because individuals would bet we'd have a better picture and to companies would have an easier life.

Of course, there are some problems in this. I cannot say that it's all good. The problems are for example, trade secrets and you know, companies would not like to share impact assessment reports to all customers because they are not obliged to, we have some guidelines, the guidelines interpreting the data protection impact assessment duals require that, just some summary.

Our share it. It's not even mandatory. It's a best practice. So this is the thing may be in which we should work more in the future. So make the impact assessment reports more available to users, and this is something that's, it's still missing into law. So we w what we say it's, it would be wise to do that.

It would be better because it would be a better form of compliance and also better rights for individuals, but to form some safeguards for individuals. But at the same time companies are still free to have a very limited approach to the existing tools. And the limited approach is not against the law.

It's just a limited approach.

Eric: And, and it occurred to me that, the algorithmic impact assessment is probably less about privacy rights, per se, although it does impact personal information and probably more about outcomes, is that a way of looking at it. Personal outcome, something that impacts your personal rates or impacts you personally, that you're not otherwise privy to versus, my, my information was disclosed.

Yeah. So in, in, sorry.

Gianclaudio: No, I was just saying, yes, it's exactly that the, the ancestor of the data protection impact assessment was privacy impact assessment. The end was just assessing them the impact on privacy of individuals, the GDPR. So did you repeat in general that the production regulation focuses not just on privacy, focuses on all fundamental rights and freedoms that are at stake when you process personal data?

So exactly what you were saying, any forms of outcomes from data processing discrimination. Manipulation stigmatization of course, loss of control of data and invasion of privacy, invasion of home safety of body. So health or safety of you know, environment and so on and so forth. So yeah, we should have a big approach and that's why it's difficult to automate the impact assessment because it's so big.

And you know, diversified that it's difficult to really automate and impact assessment.

Eric: Before you were in, you were talking about like a doctor, how there's one, possibly more complicated disclosure you know, in terms of algorithmic decisions and there may also be a simpler disclosure um, is that, what is that sort of what you meant by, in your paper, by a multi-layered explanation?

And do you want to develop that a little bit?

Gianclaudio: Yes, exactly. I think multilayer explanation was, I mean is again not Explicit in this way in the GDPR and in other legislations. But we think that interpreting the GDPR means to have different layers of explanations. Why? Because we can have um, so explanations can be different considering the moment in which we give the explanation before that the decisions are that the decision is taken or after wards and the moment the timing is relevant because the explanation can be very different.

If the decision has been already taken, or if you're just explaining the general logic of algorithm, then you can give, you can consider the audience of as I was saying before as also different, different, parties, you can have But two kinds of experts computer science experts.

So just other computer scientists that try to adapt and uncle rhythm that has already been built. Then you have experts, sectoral experts, doctors, or other kinds of experts of certain

sector, bank, financial experts. So people that decides to use algorithms or debt. The first one that uses algorithms. So what in Europe are called users of algorithm.

And then you have customers, even in customers, we could make a difference because there is different level of literacy of data literacy, digital literacy, and so on of course, explaining something to you know, privacy activist is different from explaining something to a, an elderly person or a or a kid.

Okay. But yes, audience is relevant. And in general, just looking at the explanation, we can have three layers of explanation, at least three layers. The most general one is the general explanation. I explained how the algorithm works and this is not grinding blur. It's very general, right? Computer scientists called his explanation global explanation.

This is relevant because it. It's useful to understand from the outset, how the algorithm works, et cetera. But then we have the individual explanation, which is why a certain decision was taken on me. Why Gianclaudio received that explanation, that decision, and it's different from the general one, because it's just addressing my own situation.

And then it's interesting to notice that there is an intermediate layer. The intermediate layer is the group based explanation means that sometimes similar decisions are taken for similar clusters of people or similar groups. And this might be very useful more than the individual one.

Why? Because I could understand what my cluster is. What does my group. I am in a vulnerable group. If I am in a minority group, if the company is taking advantage on some vulnerabilities that they might have, for example, the group might be very general black people or gay people, or also um, fun of some football team or, it can be very specific, and it would be very helpful to enable contestation and the right to say my own opinion in the decision.

So yes, this should be the tree layer set list.

Eric: And the individual one that would be probably, unless it was I, I would imagine that the individual one would probably trend toward the group over time since I, I can't imagine that, unless you just insert variables in, you know, a template fashion, it's certainly not going to be individualized.

Shifting gears a little bit. What is the pro-business use case for algorithmic impact assessments? How do you foresee it actually having a positive business impact or for enterprises deploying these outfits?

Gianclaudio: I think there might be several positive impacts for businesses in, uh, performing, an algorithmic impact assessment.

First of all, an answering trust of individuals and this is, it seems obvious, but you know, there are now some uh, empirical studies on customers. And for example, I just, last month, I was asked to do peer review of one study about this, and usually customers tend to assess to, to, to, to value first the price of some algorithmic systems, secondly, understandability, and let's say trustworthiness.

For example impact on the environment, but just at the church level. Uh, uh, the, the top two are price and transparency. And of course, this is the benefit of the seller. And why don't we say to seller or in generally in general companies, of course we are not talking about big tech because big tech is so powerful that they don't even need to find you know, trust and trustworthiness.

Of course they need for general thing, but you users, uh, have a Lockean effect, right? You wouldn't change from Facebook just because Facebook is not transparent. Yeah. It's more a group of people, but right. The same was for Grindr. For example, there was a case in Europe against grinder in which basically.

The, the, the court, the Norwegian data protection authority, they said, they, they are quite monopolist in the field of dating apps for LGBT people, that it would be quite difficult to give freedom to individuals. So apart from this big monopolistic examples in general financing, crosswalk announcing trust of individuals is the first benefit.

Then there are some other benefits. And the other benefits are that usually data controllers. So companies don't have a clear view of what might be the problems and inaccuracies and problems and gaps and whatever of algorithms, so doing an impact assessment or ditching the algorithm might be really beneficial for them to discover someone one to.

Gaps and problems in joker Reno, and also our model of an algorithmic impact assessment implied a big participation of individuals in the process. How, because we mentioned that exercising individual rights could be some necessary steps in the impact assessment mechanism. So basically if I contest the algorithm, I'm also giving a feedback in the impact assessment loop.

So I am a giving my personal contribution to the improvement of geography. And of course it might be considered a naive approach in which we, we all trust each other and so on, but a collaborative model in which the participation of individuals is really guaranteed, can be really beneficial for the algorithm itself because many biases of the machines are, undetected and then wanted.

It's very rare that company says, yeah, I want to discriminate black people, gay people, disabled people. It might happen. It happened in the past. But usually they are you know, problem of framing, problem of uh, statistical biases and so on. So if you allow people to

interact more, to understand, to participate, to sit to the table of the impact assessment, through their automated decision, making rights, contestation explanation, and so on, you will really create an environment of continuous improvement of your algorithm.

Uh, inter eh, you know, a proof of that is that there is a re. Computer scientist of big tech and small tech that are doing that. For example, the IBM lab in New York is already proposing the multilayer explanation system in which individuals can have you know, a day or say so, yes, I think there are benefits for businesses.

Eric: And um, I'm gonna shift gears a little bit to connect this to concepts of web three O which I have to define since when three O means a lot of different things to a lot of different people, but basically the notion of people having their own data, like retaining their own data in a wallet or something that's protected and engaging on decentralized servers.

That information isn't necessarily being collected by, like the Googles or just large, even central ISP, but, basically every you know, every destination in the web three O it's, it's decentralized, and it's more, how should I say it? There's a different level of interaction, a different level of trust.

And certainly if you're in that. You know, that could potentially be you know, a powerful grounds for building an audience to a particular platform in web three O where you're saying I'm not taking your data, but not only that, when I do take your data, here are the conditions upon which I'm going to do it.

And this is my we've done. W we're responsible and making that case, to its participants as to how it's processing that data. What are your thoughts are on that?

Gianclaudio: Uh, yeah, this is one, another complex issue because of course, new forms of let's say data governance of course, uh, challenging for the traditional structures that we use to protect data.

Sometimes we might discourage some Technologies or good advancements in particular, uh, um, summarizing a bit to your question. Rephrasing a bit, your question with the question I received in, uh, in, a conference organized by Facebook a few weeks ago, the question was, is still is there a problem of fairness, even if we decentralize the web, even if we don't own the data.

So the problem of can do we have a problem of fairness and lawfulness and so on, even if data are anonymous and we decentralized systems and so on. And I mean, of course being a privacy activist, my answer would be yes, because it's not just the tweet. Can, the tension today is too.

Really try to solve with some technologies, some big governance and management problems, but actually technologist should not be a way to, it should not be a passport to, for, like privacy enhancing technologies should not be a passport to, for not complying with the law. And so on in particular, just to make some examples, fairness is still, is there still a problem of fairness if we anonymize data, if we create decentralized systems?

Yes. If we have an impact based approach, as we were saying before, if we focus just on the impact on individuals, we can have a diverse impact even in, in a web of tree points or a scenario because I can manipulate people even without knowing who they are. I can distort. Consumer's behaviors, even if we really don't have ideas on who is a male, female gay, lesbian, and so on.

So basically the, the point is that it's wonderful to have an excellent approach and to have a justification of algorithm ex-ante. So from the beginning, but it's also important to have a periodical ex post you know, analysis of what's happening on individuals and trying to protect them.

And yeah, of course there might be an accusation of being paternalistic and European usually have this you know, kinds of acquisition and in particular from us, et cetera. But I think it's not paternalism is looking at individuals as powerless. Players into the market and trying to look, not that this technologist as a way to be free from other Bardens this technologists.

So the web three.zero in so on is important, but we should still take care of the effects manipulations. And I mean, we have many examples from Cambridge Analytica to the Facebook papers and so on. You know,

Eric: right. And it sounds like what you're saying is that you know, whether it's web two or web three, the same issue arises, which is, there there's still a collection of activity or information or generalization. You don't necessarily even need to know the identity of people. It's very easy to de-anonymize or even.

Or even make a very particular algorithmic assessments of the actions of people who meet certain characteristics without even identifying who they are. And it sounds to me like, like what you're saying is, and I agree with the mill I'll first, let you say it before. I agree with it. But uh, you know, that the notion of whether of a decentralized server or a centralized, Facebook or Google taking that data and running it through algorithms and, you know, making, engaging in activities that may impact the user is, is similar or even potentially the risks could be even.

Gianclaudio: Yeah. Yeah, exactly. I'm not identical. But yes, for, for the impacts on individuals of course it might be dangerous to think that the decentralized platform might be the solution decentralized platforms are of course you know, interesting because they

help the data minimization. So beta necessity, I just collect very limited amount of data, but at the same time, we are in a in an age where big tech has so powerful big data analytics and machine learning models.

They really don't need that information. Um, you know, they, they really do need to know. My name is Han, right? So basically I think that the point today is to protect, um, you know, in a wide way, individuals, not just their data in a nineties, where in an odd way approach. So, yes, I, I think I think it's important also, it would be dangerous to, to have a limited view in which decentralizing systems means that this data is not considered personal data because you cannot identify individuals. And so you escape from the GDPR and escaping from the GDPR doesn't mean just escaping from individual rights.

It also means escaping from the fairness principle. Impact assessments, and it's, so it would be really a way to say we do this, so we didn't have to respect other laws. And of course it's not acceptable, and this is why I think it's good to have AI based. The legislation's not just a personal database.

Legislation and that's why many countries in the wards, even European union are going towards artificial intelligence act, not looking at personal data, but looking at any form of interaction between humans and AI, but maybe this is another topic.

Eric: Yeah. We'll, we'll touch on that. With regards to, and maybe this also equally applies to AI, but you know, do you see a potential risk of the outcomes of these assessments becoming You know, becoming politicized or effectively becoming a policy decision.

Let's just say that a business was forced to prove a negative to the regulators, prove to, in your impact assessment, you need to come to the conclusion that you are not discriminating. You need to prove the, these negative needs to prove that you're not. And it may be difficult. Sometimes I know when I used to work with developers and it was always a barrier as a lawyer with developers, you'd say, you need to prove to me that you don't do this.

And they're like by definition, I don't do this because it's not programmed into the system. I'm like no, no, no, no, you still were regulated. You have to prove the negative, but in D but do you see proving the negative? Potentially as, as forcing policy decisions that even weren't necessarily avert.

Gianclaudio: Well, this is a really interesting challenge because basically my definition of uh, good with me justification was exactly improving the negative. So proving why and how, the algorithm is not discriminatory. It's not manipulative, it's not stigmatizing unfair, inaccurate, and so on. And of course in legal terms, negative proofs are quite impossible, uh, in Latin, I'm Italian.

I like to use Latin when I do law and enlightened, we say provato the abolish the devil, proof. We should be pragmatical and we sh and I think it's important that regulators like the, in Europe we have data protection authorities in the us, you have mandatory this, including federal trade commission and so on, should give some clear guidelines on how to deal with this negative proofs to avoid some political, uh, or some distortion of the existing or the future proposals.

So basically if you of course, it's difficult to prove that you are not discriminatory, but you could prove what are the reasonable steps that you took in order to avoid, to be discriminatory. And these reasonable steps should be persuasive enough to convince the data protection authority or the FTC, or the regulators in general.

That you were right. And that you are taking this problem seriously in serious account. Okay. So I really think that, of course it would be impossible to guarantee that you are 100% free from discrimination, but that's why we need to risk-based approach discrimination is a risk. The risk based approach is not that you delete the risk is that you mitigate the risk.

You know, the two components of risks are likelihood and impact severity, and you can play on one of the two or on both of these levels. You can say we made discrimination less likely. Tu to occur or you, we made the severity of discrimination, less likely. Why? Because for example, we didn't, we decided not to process sensitive data so that discrimination could be not on, I don't know, gender sexual orientation and so on.

You know, I, I think this is important and we have some, in, uh, I uh, I had the real like positivist approach just a few seconds ago saying that we should have regulators to give guidelines. Also bottom up approaches, codes of conduct certification mechanism, all of these forms of bottom up compliance forms could be helpful because if, for example, all uh, cloud service providers or all banks decide that reasonable steps in order to improve.

To be not to discriminatory art 1, 2, 3, and four. It would be easier to avoid you know, this distortion and civil convention of existing or future requirements. Yes. But of course this is a very big issue. And so yeah, we will see in the future.

Eric: Yeah. And, and so what about after you've created the algorithmic impact assessment you know, are you envisioning like full publication?

I know we talked about trade secrets may not make that feasible but how do you view you know, good practices or ideal practices around the publication of those assessments versus a summary or something? More tailored to disclosure?

Gianclaudio: Yeah. Um, I don't think that a full publication of impact assessment might be.

Beneficial because in that case does the risk that companies could have a really limited approach to impact assessment. If they know that they have to disclose everything, probably they would not take this task seriously, or they will have a very limited activity, a

Eric: litigation risk as well. Yeah.

Gianclaudio: Yeah, exactly.

But what competence should understand is that the report could say we, we saw this problem and we mitigated it, and this is a form of justification. This is a form of negative proof that as we were saying, just one minute to go. So basically it would be wonderful to have a clear view of the risks that they saw coming and the mitigations that they proposed.

Why? Because if they don't even see risks then the regulator I noticed individuals could see. The guys, you, you are social media, you didn't see that discrimination is a problem. Of course you might say that you have mitigation for it, but of course it's a problem, right? I can make some personal, you know, experience from that.

I, as, as you. So, as you said at the beginning, I am also expert assessor of a European union proposals, research proposal, ethics experts. And basically what we do is to do a risk assessment. We see whether I know how some proposals, some research proposals for European union funded projects can have high risks for fundamental rights and freedoms, and, uh, and we see the self-assessment of individuals. And when we see of companies, sorry, companies, so universities, whoever wants to apply for a grant and it's much better to say. The risk written on the paper, say, yes, we see that there are big risks here with some explanation on why you treat the risks compared to other people they say no, we see no risk because we are perfect.

So I think it's much more you know, persuasive to see a big risks on the table, but addressed rather than a perfect world in which everything is. Okay. Yeah. So I think it's quite obvious, but companies sometimes tend to have the second approach.

Eric: Because there's more protective and it's a stamp, but I get ya.

And so, so let's say that I'm a company and I want to do an algorithmic impact assessment today. And, but I don't have any experience in doing it where w what resources are available.

Gianclaudio: We have, um, not many existing resources available. It depends on what is your company. And so what are two level of risks that to have for your data?

Of course, I think that as I was saying before, the Canadian software for algorithmic impact assessment is a good proxy. I know it's just for public candidates in Canada, but it can be

it's open source. You can go on the website and do it. And so why not apply this also to private companies all around the world?

This is this is better than nothing. Then we have another tool that is it's called PR is the privacy impact assessment tool of , the French data protection authority. They have something similar to the Canadian algorithm, but it's, it's good for small and medium enterprises, or let's say for small enterprises, but this is not in my view appropriate for high risk big, uh, companies.

Why? Because it, it has a very superficial approach, the approach in that case, it's just based on. Cyber security risks, which is not the risk that we were addressing now, of course, cyber security is as important, but now we are talking about new kinds of security. We call w we are talking about cognitive security.

So the blocking the company from entering in my mind and manipulating my behaviors or discrimination and so on. So yes, there are some not many tools actually. We were trying to develop a project about automating, impact assessment, but it was not easy because we were waiting for new legislative proposals to have a very, like, um, why approach, comprehensive approach.

So we will see. But yes, I think, it would, it would not be easy for small companies to start. Yes, this is a problem, of course.

Eric: And, and to your knowledge, Are there big companies that are social media companies perhaps that are trying to tackle this

Gianclaudio: well? Yes. I mean, there are forms of innovative impact assessments that are being performed.

I was mentioning IBM, um, of course not the whole system, but some particular, parts of the algorithmic impact assessment proposal where applied, for example, I was mentioning participatory approach. So participation of individuals in the loop, and we know that Twitter, for example, has taken some steps in this direction.

In the direction of making participation of the users of representatives, of vulnerable people in the loop. So this was already, something. So yeah, th there are some um, privacy associations that are trying to propose some models. For example, algorithmic justice league was also proposing some, uh, forms of participation in impact assessment and so on.

I think one company that for sure we'll need, and we'll take some steps in the direction of algorithmic impact assessment. And the better impact assessment is exactly Facebook or Metta because basically Facebook papers showed quite clearly that there they have a problem in addressing risks. They know risk, they can recognize risks coming.

But they were not able to prove, to give the negative proof to saying we care about teenagers, having problems on Instagram. We care about misinformation. We care about distortion of behaviors of kids, right? So I think that the impact assessment justification model algorithmic impact assessment could be really useful for Facebook in the next a months.

And I have some. I've been talking with them and probably did, was take some steps in this direction.

Eric: Great. They, they could probably use it from a publicity perspective. Before we, you know, shifting gears out of the uh, away from the algorithmic assessment there've been some other developments in GDPR, which I know you're quite knowledgeable on.

One of which is the platform workers proposal which also goes to the use of algorithms but what platform workers you know, need to know about it protection of, and also protections related to the presumption that their employees and not just independent contractors. Do you want to develop a little bit on?

Gianclaudio: Yeah, sure. I think so platform work directive was proposed just few weeks ago. It was proposed at the end of November. No, actually it was the beginning of December and yes, it's, it's a great innovation that's compliment of big data production drills, basically. Yes. Uh, I, I know that in this post, in this podcast, you have already addressed the notion of co-operatives and the gig economy, and so on the European union approach to gig economy and platform workers and so on for what concerns our readers.

So I will just address this now because we wouldn't have time to address the whole proposal, but just for what concerns algorithm is that workers should be aware. Of the logic of the algorithm to requirements that the algorithm uses. So the data are the kind of categories of data that your group uses, and this is the transparency part, but it was a bit red in the GDPR.

What I really care about is another provision, which is consultation and participation of workers in decisions about the adoption of automated decision making systems. So it's the participatory approach I was mentioning before finally, clearly in the letter of the law, we have, we will have, if it will be approved when it will be approved trade unions and workers will be asked to give their opinion.

And so in the assessment of the automated decision before the dis adopted, and I think it might be very interesting. I think workers wouldn't change their mind if they understand that the algorithm delivery algorithm or I don't know just each algorithm take some data riders. Of course it would be better because it could, you could control if there are some discriminations.

And we know that for example, there was a scandal with lava and other. So we know that there are lots of views you know about algorithms used by platforms. But what is interesting is that workers are asked for their opinion. And this could, because we didn't need to take, I think, algorithms for a yes or no choice.

There might be a yes choice, which some safeguards. Yes, we don't need to open the black box, but we can put close to the black box, some protection. In the business model, not in the algorithm, in the business model. And I think this is the next step going beyond the challenge to open the black. And mitigate the effects of the black box.

Maybe that's what the platform directive is trying to do, but from a work,

Eric: sorry. Great. Something, we'll have to explore in, in future episodes. And also I wanted to touch on before I let you go on the AI act as well, which you've also written about. I think you wrote an op-ed in New York times and maybe some other places.

Gianclaudio: Yeah. Yeah, my Coulter was a professor. Frank Pasquale was the one who invented the black boxer concept applied to algorithms and law. And yes, we the AI act, the artificial intelligence act proposed last April, by European commission. Now it's under discussion with the browser privacy hub. We are trying to inform about the discussion of different policymakers, the parliament, the council, and so on will be approved probably in the spring 2023.

So not now but the discussion is at a good level. Basically it's a risk-based approach again in which, uh, algorithm users and developers and users are not customers. Users are companies that use algorithms will be asked to take some design duty. Steps in order to mitigate the risks of algorithms.

And there is a clear list of mitigations from transparency to human oversight, to data management plan, to checking accuracy of input of algorithms in a curiosity of outputs, contextualization. And so on. Then there is a list of forbidden practices, the blacklist, and this is something that I think it should have an international approach.

Social scoring, like the Chinese examples of some years ago social scoring was an example of a. A blacklist practice and go in the list of the EU AI act B uh, subliminal, uh, techniques, leading to mental manipulation, leading to physical or psychological damages. Manipulation exploiting a vulnerability is based on age disability and economic and social conditions or, um, uh, in discriminate facial recognition by police and law enforcement authorities.

So you see this is a blacklist list, and then we have some, um, um, Uh, other forms of AI. So non-risky AI that are free to, to be commercialized, et cetera, but they can adopt some codes of conduct. And of course it's not perfect. There are some problems, for example, in

my view, emotional recognition should be considered high risk at the moment is considered limited risk.

But of course, emotion recognition is something that really touches our privacy, um, and, uh, similar things might be safe for manipulation at the moment. The only form of prohibited manipulation is manipulation leading to physical or psychological damages. What about economic damages? Yes, it's of course there are room for improvements, but I think it's one of the most advanced legislation in the world and in the op-ed in the New York times of work, suggesting that the us should take.

Inspirations from European union proposal in us now there is the bill, the proposal for this deal of AI, but the bill should be a soft law. And so we really hope that there will be some hard, slow approach on AI or even in the U S

Eric: yeah um, certainly that's a developing story as well too much to touch on it at this point in the podcast, but thank you for the summary.

It was great to have you on and to learn, about, um, about the algorithmic impact assessment and start to figure out how it how it actually applies or could apply. So thank you.

Gianclaudio: Thank you very much, Eric. It was my pleasure. Thank you. .