Working Paper
Fondazione Giacomo Brodolini

Prospects on Risks, Liabilities and Artificial Intelligence, empowering Robots at Workplace Level

The EU Regulation 2024/1689, with the related EU and Domestic Legal Frames, compared to the U.S. Legal System

di Michele Faioli

Fondazione Giacomo Brodolini

24

Working Papers della Fondazione Giacomo Brodolini

Direttore scientifico della collana
Anna Maria Simonazzi, Presidente della Fondazione Giacomo Brodolini

Direttore scientifico della serie "Scuola Europea di Relazioni Industriali - SERI"
Michele Faioli, Università Cattolica del Sacro Cuore

Fondazione Giacomo Brodolini
00185 Roma - Via Goito, 39
tel. 0644249625
fax 0644249565
info@fondazionebrodolini.it
www.fondazionebrodolini.it

ISBN: 9788895380643

About the author

Michele Faioli is an Associate Professor at Università Cattolica del Sacro Cuore (abilitato/tenured as full professor). He teaches courses in the areas of labor relations and comparative/EU labor law, including industrial relations, social security, tech law. Faioli's comparative research mainly looks at collective bargaining impact on labor relations, AI and robots operating at workplace level. As Visiting Fellow at the ILR School of Cornell University and at the Fordham Law School he carried out investigations on international labor rights, global trade and unions strategies. Michele has written widely on tech and labor law (see his book, "Mansioni e macchina intelligente", Giappichelli, 2018), undeclared work, social security, pension funds and other topics for a variety of law reviews and journals. After his appointment to the CNEL (Consiglio Nazionale dell'Economia e del Lavoro) in 2018, he has conducted investigations also on the possible social applications of blockchain to the EU labor market and social security systems. Prof. Faioli chairs the SERI - Scuola Europea di Relazioni Industriali. He directs the EUROFOUND Italian Observatory.

https://docenti.unicatt.it/ppd2/en/docenti/27259/michele-faioli/profilo

Working Paper
Fondazione Giacomo Brodolini

Prospects on Risks, Liabilities and Artificial Intelligence, empowering Robots at Workplace Level

The EU Regulation 2024/1689, with the related EU and Domestic Legal Frames, compared to the U.S. Legal System

di Michele Faioli

24

Sommario

Introduction

Abstract: The tech transformation is forcing us to rethink new solutions for mapping risks and opportunities arising from interactions between robots, empowered by AI, and workers ("IWRs"). While doing so, we must keep in mind that the debate is not about artificial intelligence and robots, but about us, who will have to live and work with them. Following this line of thought, we are aware that there are challenging changes ahead. Such changes should be investigated in relation to the occurrence of new social and psychophysical risks connected to IWRs. My essay is built on the conviction that robots empowered by AI ("AI/R"), also in the form of the so-called "Frontier AI", should not merely augment humans in their capabilities, but that AI/R and workers can become better together, thereby boosting worker wellbeing and productivity. The absence of a unified scientific framework makes us blind to various forms of reciprocity that exist when AI/R share the same physical and social space as workers. My thesis is aimed at buttressing the forefront of understanding and shaping the future of work regulation around AI/R, the related liabilities regime, also with reference to possible damages to compensate and risks to mitigate, moving from a legal comparison between the EU and U.S. labor systems in the field of accidents at work and occupational diseases. The questions from which I move are as follows: what do we want artificial intelligence, including advanced robotics, to do for us, or rather "with" us in the workplace? Who regulates this interaction between the worker and the AI/R? And how should that interaction be regulated? Furthermore, given that interaction, what might happen if the responsibility for any injury to the worker depends directly/exclusively on the AI/R? Do insurance regimes for accidents at work/occupational diseases, already included in our social security systems, even cover work related injuries connected directly to AI/R? And how efficient, in terms of damage mitigation, are the procedures that the safety at work regulation imposes in the context of advanced technology production and where the worker co-operates with the AI/R? Upon the results of an ongoing transdisciplinary research project, I also intend to develop a basis for a new academic labor law field concerning the IWRs regulation (here labelled as "Robot Labor Law").

Keywords: Labor Law and Industrial Relations, Robot, Artificial Intelligence, Frontier AI, Risks, Liabilities Damages Compensation, Harms, Injuries, Accidents at Work, Occupational Diseases, Mitigation Measures and Insurance Plans, Social Security Schemes.

Summary[1] **: 1.** The Technological Framework in which the Research is situated (Foundation Models, LLMs, Frontier AI, Generative AI). **2.** Legal Comparison between the U.S. law and the European Law governing Labor and Artificial Intelligence. Legal Genotypes and Phenotypes. **3.** AI, Workplace related Risks and Personal Injuries. The EU Legal Frame. **4.** AI, Workplace related Risks and Personal Injuries. The U.S. Legal Frame. **5.** Conclusions. For a New Branch of Labor Law (RLL - Robot Labor Law)

1 This investigation stems from my previous essays and studies that were already published in law journals and presented in several conferences/seminars. See my essay, M. FAIOLI, Robot Labor Law. Linee di ricerca per una nuova branca del diritto del lavoro e in vista della sessione sull'intelligenza artificiale del G7 del 2024, in Federalismi.it, 2024, vol. 8, p. 182. Such an investigation was developed in the context of the first activities of the constituting research center WSP Co-Lab, Workplace and Social Policies Co-Lab (Università Cattolica del Sacro Cuore), within the paths of discussion with social partners organized by the SERI-FGB (Scuola Europea di Relazioni Industriali - Fondazione G. Brodolini) as well as in relation to the PRIN 2022, Project "SafetyChain: Social Blockchain for Implementation of a Digital Wallet for Occupational Health and Safety Training," project code "20225TZB2P", MUR – https://www.mur.gov.it/it - funded by the European Union - Next Generation EU.

1. The Technological Framework in which the Research is situated (Foundation Models, LLMs, Frontier AI, Generative AI). Legal Comparison between the U.S. law and the European Law governing Labor and Artificial intelligence.

The purpose of this study is to initiate a dialogue with colleagues and technology experts, as well as with unions and employers' organizations, on the adequacy of current legal schemes referable to artificial intelligence and robotics operating at a workplace level, for the prevention and mitigation of new work hazards, as well as implementing protection against them. This is especially in view of the policies that may be made on this issue following the G7 in 2024. The questions from which we move are as follows: what do we want artificial intelligence, including advanced robotics, to do for us, or rather "with" us in the workplace? How can we avoid that the artificial intelligence/robot (hereinafter also "AI/R") should never act to our detriment? Who regulates this interaction between the worker and the AI/R? And how should that interaction be regulated? Furthermore, given that interaction, what might happen if the responsibility for any injury to the worker depends directly/exclusively on the AI/R? Do insurance regimes for accidents at work/occupational diseases, already included in our social security systems, even cover work related injuries connected directly to AI/R? And how efficient, in terms of damage mitigation, are the procedures that the safety at work regulation imposes in the context of advanced technology production and where the worker co-operates with the AI/R?

Such questions do not pertain to the lack or insufficiency of technology, but to problems that are ours because they are totally human, referable to the rules we give ourselves, both on an individual and a collective level. To positively affect the problems that originate from such questions, referring to the regulation of advanced technology in the workplace, one must preliminarily look at the field of the human, and ask what we intend to do to improve that field. Finally, we must ask ourselves what goal is really intended to be achieved by regulating and verifying compliance with that regulation.

The following remarks are, of course, beyond the scope of ethical analysis, which is almost geared, even internationally, to discern between what is the AI/R super-power and what is not, and, consequently, between what is good and what is bad for humanity in the creation of this AI/R super-power[2]. Here we would like, instead, to examine the perspective of the regulation of what is not yet fully known and, in particular, of the regulation of the AI/R in active cooperation with the worker in highly technological environments, composed of new models of artificial intelligence, taking into consideration that the general socio-economic and industrial framework, in which this scientific contribution arises, can perhaps be compared with that of the start of nuclear energy experiments and, more recently, the great financial crisis of 2008[3]. In both cases, macro-regional, national, domestic regulation served little or no purpose. Perhaps it even came too late and proved inefficient. Instead, the ability to create overall governance of the whole affair (during and after World War II for nuclear power, and from 2008 onward for the financial crisis) had to be transferred to an international regulator. The results, in those cases, have been partial not conclusive, perhaps not satisfactory, but certainly pointing in the right direction: regulating the phenomenon at the national/domestic level serves little purpose, while it may be more useful to regulate the conduct of those who manage/govern that phenomenon at the national level by imposing international standards and creating related supervisory authorities.

This, to some extent, is already happening around the definition of the AI that the OECD has recently identified, also to initiate forms of cooperation between different jurisdictions[4]. Western legal systems, at the transatlantic level[5], are also moving towards an AI common regulation. European and U.S. legal regimes are looking for answers on how to regulate the complexity of the phenomenon, based on a legal assumption that is evolving very quickly and must apply transnationally.

---

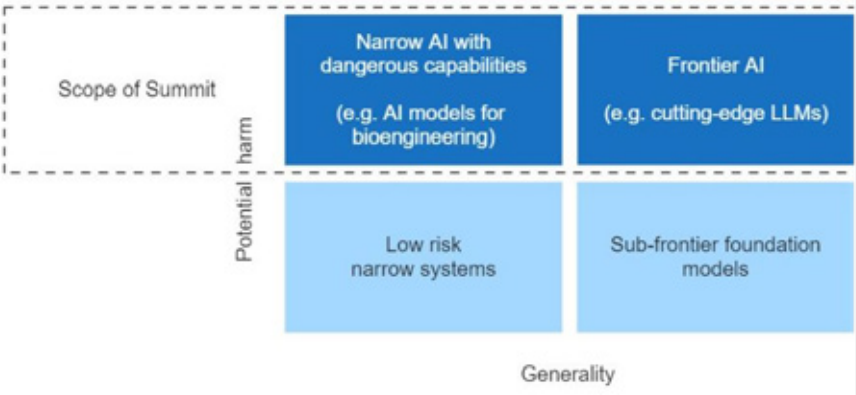
2 See the work of the UN Commission on the Ethical Aspects of Artificial Intelligence, AI Advisory Board - https://www.un.org/en/ai-advisory-body. For the most important recent academic investigations, see the book by D. J. GUNKEL, Person, Thing, Robot. A Moral and Legal Ontology For the 21st Century And Beyond, MIT Press, Cambridge, 2023, who also recalls the foundational studies of P. RICOEUR, Il giusto, Effatà, Turin, 1-2, 2005 and R. ESPOSITO, Persone e cose, Einaudi, Turin, 2014.

3 The insight is from L. ZINGALES, B. MCLEAN, Who Controls AI? With Sendhil Mullainathan, in Capitalisn't, Dec. 21, 2023.

4 See the documentation here - https://oecd.ai/en/wonk/ai-system-definition-update

5 The outcomes of the dialogue between the European Union and the United States of America can be analyzed at https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council_en#objectives-of-the-partnership -

An attempt is being made to pose a legal notion of AI/R that is shown to be as suitable as possible for the future context, not for what we already know today or has happened in the past. This notion coincides, at least in Europe and the United States of America, with that of "Frontier AI Models", here also next-generation or frontier artificial intelligence[6]. From a technological point of view, we can define Frontier AI as that which falls under the structure of so-called foundation models, which also normally possess a certain ability to determine damage. LLMs, BERT, DALL-E, GPT-3 also fall under the definition of a foundation model. It is a model that is built on a large database and is adaptable to an almost infinite set of tasks (so-called downstream tasks). The diagram below is taken from the papers of the AI Safety Summit in London, 2023.



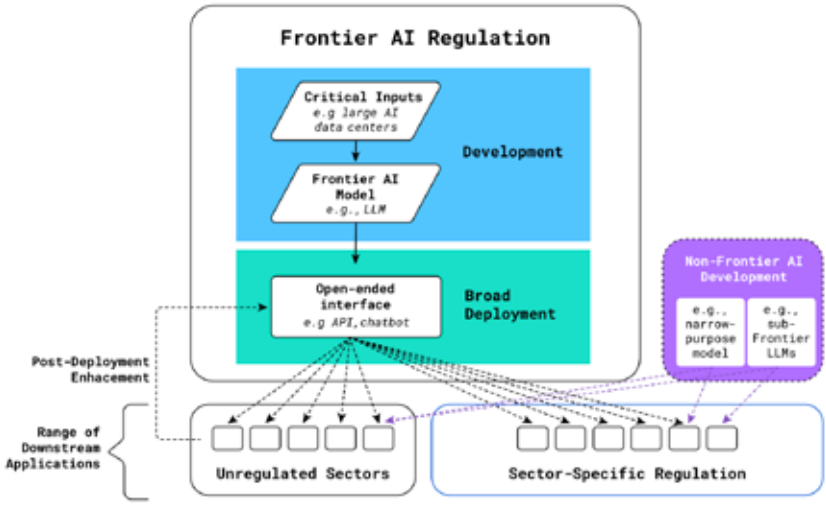
6 *Hereinafter I do avoid using the term "general-purpose AI" to avoid confusion with the use of that term in the EU legal frame and other legal orders. Frontier AI systems are a related but narrower class of AI systems with general-purpose functionality, but whose capabilities are significantly advanced. For the definition of Frontier AI, see the results of the intergovernmental conference AI Safety Summit, London, 2023, and, in particular, the paper Frontier AI: Capabilities and Risks - Discussion Paper, October 25, 2023 at* https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper *- "For the purposes of the Summit we define frontier AI as highly capable general-purpose AI models that can perform a wide variety of tasks and match or exceed the capabilities present in today's most advanced models. Today, this primarily includes large language models (LLMs) such as those underlying ChatGPT, Claude, and Bard. However, it is important to note that, both today and in the future, frontier AI systems may not be underpinned by LLMs, and could be underpinned by another technology." See also the wide investigations concerning the legal definition of Frontier AI and General-Purpose AI carried out by F. G'SELL, Regulating under Uncertainty: Governance Options for Generative AI, available at SSRN:* https://ssrn.com/abstract=4918704*, August 06, 2024.*

In relation to such a definition, we are interested in placing at the center of labor law reflections on the AI/R (also in the form of the Frontier AI) acting in cooperation with the worker. This means that we are going to explore the way by which (i) on one hand, it may be protected, and (ii) on the other hand, it may be mitigated those damages/injuries/harm that result from downstream tasks carried out at workplace level, in the context of interactions between person and machine (IWRs), managed by the AI/R[7].

I intend to investigate Frontier AI models, also empowering robots that operate at workplace level, that are "highly capable foundation models that could exhibit sufficiently dangerous capabilities. [...] Foundation models, such as large language models (LLMs), are trained on large, broad corpora of natural language and other text (e.g., computer code), usually starting with the simple objective of predicting the next token. This relatively simple approach produces models with surprisingly broad capabilities. These models thus possess more general-purpose functionality than many other classes of AI models, [...]. Developers often make their models available through broad deployment via sector-agnostic platforms such as APIs, chatbots, or via open-sourcing. This means that they can be integrated in a large number of diverse downstream applications, possibly including safety-critical sectors"[8]. See the diagram below for a recap of this idea.

7 *In the civil and contract law field, concerning AI/R and liabilities, it is pivotal the analysis carried out by A. BERTOLINI, Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules, in Law Innovation and Technology, No. 5(2), 2013, p. 214, F. BATTAGLIA, N. MUKERRJI, J. NIDA-RUMELIN, Rethinking Responsibility in Science and Technology, Pisa University Press, Pisa, 2014 and E. PALMERINI, E. STRADELLA, Law and Technology. The Challenge of Regulating Technological Development, Pisa University Press, Pisa, 2013.*

8 *See M. ANDERLJUNG, J. BARNHART, A. KORINEK, J. LEUNG, C. O'KEEFE, J. WHITTLESTONE, S. AVIN, M. BRUNDAGE, J. BULLOCK, D. CASS-BEGGS, B. CHANG, T. COLLINS, T. FIST, G. HADFIELD, A. HAYES, L. HO, S. HOOKER, E. HORVITZ, N. KOLT, J. SCHUETT, Y. SHAVIT, D. SIDDARTH, R. TRAGER, K. WOLF, Frontier AI Regulation: Managing Emerging Risks To Public Safety, 2023, in* https://arxiv.org/abs/2307.03718



The essay moves, therefore, from an observation of reality: the technological transformation, due to next-generation artificial intelligence (i.e. Frontier AI), is forcing us to lay out a new mapping of the actual risks and possibilities for innovation arising from the interaction between workers and intelligent machine. This confronts us with more complex questions than we have been asking to date. To this end, we choose to focus our view on new social and psycho-physical risks arising from such a person/intelligent machine interaction. While we are convinced that greater productivity and greater well-being can certainly result from that interaction, we cannot fail to investigate what is about to happen in workplaces re-planned by advanced technology. The classical notion of risk probably cannot longer be adequate because of the operational presence of a "third element" between employer and worker[9] . This third

9 *On the notion of intelligent machine as a "third element" of the employment relationship I refer to my previous studies. In particular, see M. FAIOLI, Mansioni e macchina intelligente, Giappichelli, Turin, 2018 and the theoretical line defined by some of my writings, including in particular M. FAIOLI, Data analytics, robot intelligenti e regolazione del lavoro, in Federalismi.it, No. 9, 2022, p. 149; M. FAIOLI M, Artificial Intelligence: The Third Element of the Labor Relations, in A. PERULLI, T. TREU, (eds.), The Future of Work. Labour Law and Labour Market Regulation in the Digital Era, Alphen aan den Rijn, Wolters-Kluwer, 2021; M. FAIOLI, Unità produttiva digitale. Perché riformare lo Statuto dei lavoratori, in Economia & lavoro, Rivista di politica sindacale, sociologia e relazioni industriali, No.*

element exercises powers, confronts obligations as well as may result in damages, also creating forms of contractual and extra-contractual liability. The essay, as part of a broader, transdisciplinary research project, intends to initiate a confrontation, including academic discussion, to be able to create a theoretical substrate of a new branch of the labor law that pertains to the study of the regulation of artificial intelligence/robots in active interaction with workers in technological production contexts (also referred to here as Robot Labor Law - "RLL"). The essay, defining the comparative method chosen (section 2), sets up the analysis of the framework of European rules and the related labor law context, whose first task will be to define the kind of risk arising from AI/R, including in its operation in the workplace (section 3). It continues with an analysis of the current U.S. system of AI/R regulation, which also seems to have as its purpose the geo-political drag on the global scale of AI/R normalization (paragraph 4). In paragraph 5, an attempt will be made to outline some of the contents of the new, transnationally relevant branch of the "Robot Labor Law" - RLL, pending the completion of the research project underlying this essay.

1, 2021, p. 41.

2. Legal Comparison between the U.S. law, the European Law governing Labor and Artificial intelligence. Legal genotypes and phenotypes.

The method for comparing the EU/U.S. legal regimes can move from the construction of a genotype, from which, at a later stage, specific legal phenotypes can be derived, valid for observing the two legal systems[10] . The genotype can be constructed in relation to three issues: (i) What risk and harm does that legal system intend to select for protection to be arranged in relation to AI/R's ability to interact with the worker? (ii) What mitigation and prevention measures can that legal system introduces to address the need for protection related to those risks and harms? (iii) What institutions will such a legal system put in place to implement that protection, including in terms of individual and collective enforceability?

The three problems allow us to understand the extent to which the concerning law is more oriented towards: a risk-prevention norm or, on the contrary, towards a punitive norm that is more oriented toward punishing the harm and fixing the corresponding reparation. We know that the notion of risk, individual and collective, is never neutral, at least when observed through the eyes of a law scholar. It is the result of a very specific legal policy choice, in this case made at the European and U.S. level. To have decided to place in the notion of risk this (potential) damages/harm to the person resulting from AI/R conduct, even in the species of Frontier AI, is to have predetermined a path of regulation, which no longer coincides (or does not exclusively coincide) with the claim of compensation for a harm by the person who suffers it, but above all with a norm that uses a certain way of regulating risk management, in continuity or analogy

10 Here we follow the comparative law method developed, at the Cornell Law School, by R. B. SCHLESINGER, Comparative Law. Cases, Text, Materials, Foundation Press, New York, 1988 edition, G. GORLA, Diritto comparato e diritto comune europeo, Giuffré, Milan, 1981 and by R. SACCO, Introduzione al diritto comparato, Giappichelli, Turin, 1990. We also refer to L. MENGONI, Diritto vivente, in Jus, No. 1, 1988, p. 14, G. BENEDETTI, L'elogio dell'interpretazione traducente nell'orizzonte del diritto europeo, in Europa e diritto privato, No. 2, 2010, p. 413. See also P. SANDULLI, M. FAIOLI (a cura di), Attività transnazionali. Sapere giuridico e scienza della traduzione, Edizioni Nuova Cultura, Rome, 2011.

with other sectors where there are already established practices of risk management (energy, environment, etc.). Even in the AI system, one can decide to regulate such damages *ex ante*, making use of the notion of risk, with anticipatory, mitigation or precautionary models, introducing a kind of transplantation of legal institutions already used elsewhere and probably useful here as well.

Hence, it can be considered that the law shows active participation, almost in a logic of conceptual construction, in the definition of AI/R, precisely because of the type of risk (individual and collective) that is intended to be selected and then, because of the norm, mitigated, prevented, ensured, etc.

The risk marks a direction, and, in some ways, a datum to be observed in the future[11]. Harm has already occurred[12]. Law reacts with respect to risk with a series of preventive schemes. In relation to harm, the law poses restorative protections because of an event that has adversely affected the person, setting the rules on "if," "who," and "how much/how" (whether the harm that has occurred should be compensated, who is obligated to compensate, how much/how the harm should be compensated)[13] . Where

---

11 For a very useful map of the concept of risks related to AI/R see the MIT AI Risk Repository, available at https://airisk.mit.edu *– that is periodically updated. The scientific field concerning risks and legal theory is quite wide. I am not going to set up an exhaustive of risk legal theory. Rather, I introduce in this essay those aspects of the legal doctrine concerning risk which may be more relevant for my purposes. For the theoretical approaches mainly related to the common law systems, see the overview realized by J. STEELE, Risks and legal theory, Oxford, Hart, 2004. The legal notion of risk has been the subject of important studies and theories, also elaborated by the Italian labor and social security law scholars. See F. SANTORO PASSARELLI, Rischio e bisogno nella previdenza sociale, Giuffrè, Milan, 1948; more recently, P. LOI, Il principio di ragionevolezza e proporzionalità nel diritto del lavoro, Giappichelli, Turin, 2017 and Il rischio proporzionato nella proposta di regolamento sull'IA e i suoi effetti nel rapporto di lavoro, in Federalismi.it, No. 4, 2023, p. 239. See also T. TREU, Il diritto del lavoro: realtà e possibilità, in ADL Argomenti di diritto del lavoro, No. 3, 2000, p. 467.*

12 G. ALPA, Danno "in re ipsa" e tutela dei diritti fondamentali (diritti della personalità e diritto di proprietà), in Responsabilità civile e previdenza, No. 1, 2023, p. 6; P. SIRENA, Danno-evento, danno-conseguenza e relativi nessi causali. Una storia di superfetazioni interpretative e ipocrisie giurisprudenziali, in Responsabilità civile e previdenza, No. 1, 2023, p. 68; V. ROPPO, Pensieri sparsi sulla responsabilità civile (in margine al libro di Pietro Trimarchi), in Questione Giustizia, No. 1, 2018, p. 108; P. GALLO, Quale futuro per il contatto sociale in Italia?, in La Nuova Giurisprudenza Civile Commentata, No. 12, 2017, p. 1759; V. ROPPO, Responsabilità contrattuale: funzioni di deterrenza?, in Lavoro e diritto, No. 3-4, 2017, p. 407. See also A. BOLLANI, Il danno alla persona nel diritto del lavoro, tra influssi della civilistica e necessari adattamenti, in Giornale di diritto del lavoro e di relazioni industriali, No. 176, 2022, p. 593.

13 For an overview of the current theories concerning contracts and civil law, see the studies elaborated by A. D'ADDA, Danni "da robot" (specie in ambito sanitario) e pluralità di responsabili tra

one regulates risk, in a sense subsuming the damage within the notion of risk, one is merely performing a descriptive operation, based on statistics, probabilities and samples. By reason of this, one indicates the ways by which the possibility of the occurrence of that risk can be reduced, by imposing audits, due diligence, permit/permit requirements, etc.

In general, risk has in it some mechanism for selecting the event to be prevented, insured, mitigated, etc. because the rule on risk itself is based on a calculation to prevent, insure, mitigate the risk itself. In the case of artificial intelligence, this holds even more true because it is the law itself that fixes or updates the notion of artificial intelligence, moving from an indefinite variability of technological definitions, identifying one or a few for the desired legal effects at that historical moment and in that socio-political context. And it is in relation to that notion of artificial intelligence (to exemplify, today it is Frontier AI, tomorrow it will be something else) that a series of protective mechanisms and behavioral obligations are triggered.

The potential risk from AI/R, even managed by Frontier AI, is almost always related to a context. It depends on the context in which the AI/R is placed. Risk mapping cannot be done only once, and for all, because risks change due to several interacting factors, with the consequence that what can be defined as risk at that time is risk. The framework of factors can change and, as a result, so does the risk, which, perhaps, in some cases, can undergo a kind of downgrade into something else. Moreover, there is a risk arising from AI/R, managed by Frontier AI, only if with its use comes increased exposure to harm: which, at least in the legal logic of this study, means looking at risk with reference to people who also interact in workplaces with forms of AI/R, managed by Frontier AI. From that viewpoint, the stronger the human context is, given that it is based on education and training to handle the eventual problems of that AI/R, the lower the impact of risk.

The three problems posed above (what risks, what mitigation measures, what institutions) become a kind of general outline in order to be able to more effectively compare EU/U.S. legal regimes, keeping simultaneously in mind, tech law, as it has been developing in recent times, and labor law, which reacts (as has always been known) to the technology introduced at the firm level in relation to more factors, in addition to those already under investigation in some way (management, coordination,

sistema della responsabilità civile ed iniziative di diritto europeo, in Rivista di diritto civile, No. 5, 2022, p. 805. See also G. CALABRESI, E. AL MUREDEN, Driverless cars. Intelligenza artificiale e futuro della mobilità, Il Mulino, Bologna, 2021.

control, variability of tasks, opinion surveys, discrimination, etc.). There is, in fact, an elective field to carry out this examination: that of safety in the workplace, personal injury and related insurance (for the Italian system, see the INAIL regime[14]). It must be understood that advanced technology (AI/R - Frontier AI) can be viewed as a tool which may cause personal injury to the worker due to a violation of safety regulations, and as a tool that helps prevent personal injury. In other words, the intention is to observe, through the research project that has been initiated, how AI/R, also in the form of the Frontier AI, in the context of its functions as a third element, can coordinate, control, and direct human labor, as well as varying its tasks and theoretically applying disciplinary sanctions. On the one hand, it can cause harm and, on the other hand, at least hopefully, also capable of preventing it.

Hence the centrality of the object of the present study: knowing that law is called upon to adapt to technology[15], not the other way around, it is no longer sufficient (only) to understand what limits to place, through law and/or collective bargaining, with respect to the powers that AI/R - Frontier AI can exercise[16] . Instead, we must also grasp what harms and risks may result from AI/R - Frontier AI's conduct/choices, when used in interaction with the human person in the workplace, and, therefore, what criteria for imputation of liability we are required to update and why, given the production environment that is evolving day by day. Furthermore, we must understand how to prevent and mitigate such risks and damages, including through innovative firm-level organizational procedures and advanced technological systems that can be introduced with the specific function of preventing/mitigating any risks/damages from Frontier AI.

14 See the general description of the social security regime managed related to the accidents at work and occupational diseases covered by the INAIL regime at https://www.inail.it/portale/it/multilingua/english.html

15 To evoke one of the strongest ideas of G. GIUGNI, Il processo tecnologico e la contrattazione collettiva, in F. MOMIGLIANO, Lavoratori e sindacati di fronte alle trasformazioni del processo produttivo, Feltrinelli, Milan, 1962.

16 In this regard see the structure of the theory in M. FAIOLI, cit., 2018, p. 93. and p. 211.

3. AI, Workplace related Risks and Personal Injuries. The EU Legal Frame.

Normally, when faced with such complex phenomena as those related to injury to the worker attributable to an intelligent machine, the law scholar suggests intervening with a norm, the economist believes it is sufficient to introduce a price for inefficiency such that those who must comply with the norm will do so[17] . There are law scholars who have combined the two perspectives. These certainly include, with differing views, G. Calabresi[18] and R. Posner[19] , who have applied the logic of costs and benefits to law: one sets the starting point (in our case, the personal injury to the worker attributable to Frontier AI - "S"); then, one establishes the objective of the analysis (minimizing social costs - "C"); finally, by various steps, one selects the most economically efficient way ("V") to achieve that result ("R"). Now, using an exposition scheme in line with that theory, we might have the following: to realize R (outcome), moving from S (situation), one must calculate C (difference between costs) and identify V (efficient way)[20]. Let us adapt the argument to the case we are concerned with here: C, the cost of damages attributable to AI/R can neither be passed on to workers in any way, nor can it be fully internalized by the employer, since these are damages arising not from the employer but from AI/R, which we assume here to be the third element of the employment relationship. The result R certainly does not coincide with trying to avoid such damages altogether because that would cost too much. It is, instead, much more sensible to assume that we should reduce such damages as much as possible, to the maximum extent according to the most appropriate applicable technology (for the Italian system, see the assumptions of Art. 2087 of the Italian Civil Code, also in relation to Act No. 81 of April 9, 2008). The efficient way V will be able to be achieved, in some cases, by general prevention and, in many others by specific prevention, defined by measures that are aimed at identifying the safest way for intelligent machines to operate in workplaces. General prevention is based on the idea that the employer

17 *R. COASE, The Problem of Social Coast, in Journal of Law&Economics, No. 3, 1969, p. 1.*

18 *G. CALABRESI, Transaction Costs, Resource Allocation and Liability Rules. A Comment, in Journal of Law&Economics, No. 11, 1968, p. 67; G. CALABRESI, Melamed, Property Rules, Liability Rules and Inalienability. A view of the Cathedral, in Harvard Law Review, No. 85, 1972, p. 1089.*

19 *R. POSNER, Economic Analysis of Law, Aspen, Boston, 1992.*

20 *See E. AL MUREDEN, Costo degli incidenti e responsabilità civile quarant'anni dopo. Attualità e nuove prospettive nell'analisi economico-giuridica di Guido Calabresi, in Rivista di diritto civile, No. 4, 2015, p. 1026.*

can assume the costs of introducing AI/R at workplace level and related compliance with certain parameters. General prevention has the identification of risk and the models for managing that risk as a prerequisite, in relation to the level of tolerability (individual/collective) that is intended to be delineated, the verification that each person can carry out on the risk arising from a certain conduct (in this case of AI/R) and its evaluation (high, medium, low risk), as well as the supervisory actions carried out by public authorities. General prevention poses numerous critical issues when, as in the case of AI/R, there is a counterbalance to be made with the rights of the human person, those most important, fundamental rights pertaining to the protection of dignity. Special prevention is structured according to the model of (tendentially strict) liability of the person who determines the damage, with insurance to cover the costs, insurance that can be entrusted to the market, or as is the case in many European countries, to a social security scheme (for Italy, INAIL).

This general theory is suitable for interpreting the European standards, which are predominantly set up with a view focused on general prevention, risks to be protected against and referring to a certain variable notion of AI/R (see the Regulation EU 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence - Artificial Intelligence Act hereafter the "Regulation")[21]. The approach to risk has been regulated according to a pattern already known in Europe, coinciding largely with risks arising from the commercial circulation of products or with risks pertaining to the environment, energy production, etc. The special prevention, that relates to damages to be compensated, is assigned to a directive (see Proposal for a Directive of the European Parliament and of the Council on the adaptation of the rules on non-contractual civil liability to artificial intelligence - Directive on liability by artificial intelligence - COM/2022/496 final - hereafter "Directive"). The Regulation and the Directive belong to a much broader European regulatory strategy on artificial intelligence, which consists of at least four areas of regulation (artificial intelligence, data governance, digital markets and services, and digital platforms[22]). The Regulation also selected the legal concepts that may be

21 See the recent studies of S. S., NUNO, The Artificial Intelligence Act: critical overview, available at SSRN: https://ssrn.com/abstract=*, June 2024; S. WACHTER, Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond, in Yale Journal of Law & Technology, Volume 26, Issue 3, 2024; M. D. MURRAY, Generative Artificial Intelligence -Where did it come from? How does it work? available at SSRN:* https://ssrn.com/abstract= *June 2024. See also the legal problem investigated by M. SHAAKE, The Quest for Global AI Governance: the UN AI Advisory Body, April 2024 in* https://cyber.fsi.stanford.edu/events/april-2-quest-global-ai-governance-un-ai-advisory-body

22 See https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_en - See the studies by M. ALMADA, N. PETIT, The EU AI Act: Between Product Safety

currently used to state what Frontier AI may be considered pursuant to the EU legal regime, indicating the "General-Purpose AI Model" as an AI model trained with large datasets, capable of performing a wide range of tasks and integrated into various downstream systems or applications (see Article 3 – Definitions, Article 51 – GPAI Models with systemic risk, Article 25 – Providers' responsibilities Articles 53 and 55 – GPAI Model providers' obligations, Article 50 – Transparency obligations, Article 88 – Enforcement)[23].

With reference to general prevention, the Regulation lays out a framework that is designed according to a logic that sees, on the one hand, macro-risks referable to the market and its actors (enterprises/consumers) and, on the other hand, micro-risks that those operating in the AI/R market are called upon to detect both through preventive verification and through analysis following the introduction of AI/R into that market. The risks arising from AI/R are predominantly related to people's health, safety, and fundamental rights. There are unacceptable risks (first type – see prohibited AI practice enlisted in Article 5 and Article 99, Para. 3)[24], risks that are high, but under certain conditions acceptable (second type – see Article 6, Annex I – Conditions to be a high-risk AI system, Article 6, Annex III – List of high-risk AI systems, Articles 16, 22, 23, 24, 26, 31, 33, 34, 50 – Obligations on parties), and finally risks that are completely acceptable because they are minimal (third type). Anything that can lead to the unacceptable risks is subjected to an absolute ban. That which leads to high risks is subjected to a system of preventive and *ex post* management procedures. To mitigate

---

and Fundamental Rights, 2023, in SSRN: https://ssrn.com/abstract=4308072 *and N. MORENO BELLOSO, N. PETIT, The EU Digital Markets Act (DMA): A Competition Hand in a Regulatory Glove, 2023, in SSRN:* https://ssrn.com/abstract=4411743 *; A. BARTOLINI, Artificial Intelligence and Civil Liability, 2020 in* https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.*pdf*

23 A general-purpose AI model may be also classified as having "systemic risk" if it has high-impact capabilities, such as significant computational power (exceeding 1025 FLOPs) or other criteria set by the European Commission (Article 51). FLOPs (Floating Point Operations per Second) measure the computational power of a system by counting the number of floating-point calculations it can perform per second. This high computational threshold indicates the model's ability to handle extensive and complex tasks, necessitating robust regulatory oversight. Systemic risks may include major accidents, disruption of critical sectors, serious consequences to health and safety, and actual or reasonably foreseeable negative effects on democratic processes, and public and economic security. Systemic risk increases with model capabilities and model reach.

24 AI/R that uses subliminal techniques or exploits the vulnerabilities of a specific group of people, creating discrimination, or used by or on behalf of public authorities for the purpose of assessing or classifying the trustworthiness of individuals, or capable of "real-time" remote biometric identification in publicly accessible spaces for the purpose of law enforcement activities, given certain conditions.

anything which involves minimal levels of risks, various forms of self-regulation are promoted. For high risks, those of the second type[25] , the AI/R is subjected *ex ante* to a verification of compliance, after which it can be placed on the market. This verification is carried out because the AI/R may be a safety component of a product (or itself a product) or alternatively, has an impact on fundamental rights[26]. In the first case (security component/product) there is a role for third parties who have powers of notification, inspection, etc. In the second case (fundamental rights) there is no role of third parties facilitating the adaptation process because there is an absolute ban on the use of that type of artificial intelligence. In general, *ex ante* verification moves from the identification of risks and the probability assessment of their occurrence, under conditions of normal predictability. It is also placed in relation to the assessment of other possible risks arising from the analysis of data collected by the *post*-market monitoring system and the introduction of risk management measures. Compliance is based on an authorizing act that can be obtained, following *ex ante* verifications, even with the help of experienced third parties.

There is also an obligation to carry out certain *in itinere* and *ex post* verifications to enable the possible adoption of mitigation measures and the ongoing management over time of the risk arising from AI/R. The European standards impose several protections related to the collection and processing of data by which the AI/R system is fed into the marketplace. For *ex post* verifications, a system of public records, monitoring, reporting and remedies of the possible cause that creates that risk is defined. AI/R systems that may result in specific manipulation risks are subject to specific transparency requirements if they interact with humans, are used to detect emotions, are used to define social categories based on biometric data or may generate false content.

---

25 High-risk AI systems are those intended for use as safety components of products subject to ex ante third-party conformity assessment and those that have implications primarily in relation to fundamental rights explicitly listed in Annex III of the Regulation. This list of high-risk AI/Rs in Annex III of the Regulation contains several AI/Rs whose risks have already materialized or may materialize. However, the Commission may extend the list of high-risk AI/Rs used within certain predefined areas by applying a set of criteria and a risk assessment methodology.

26 These include for certain, given the reference in the preamble of the Regulation, the right to respect for private life and protection of personal data, non-discrimination and equality between women and men, freedom of expression and freedom of assembly, to ensure the protection of the right to an effective remedy and to a fair trial, the presumption of innocence and the rights of the defense, the rights of workers to just and fair working conditions, a high level of consumer protection, the rights of the child, and the right to a high level of environmental protection and the improvement of the quality of the environment, including in relation to the health and safety of persons.

In relation to institutions, it should be noted that a European Artificial Intelligence Committee will be established, with representatives of the Member States and the Commission. At the national level, Member States will have to designate competent authorities to monitor the application and implementation of the Regulation. The European Data Protection Supervisor will act as the competent authority to supervise the European Union's institutions, agencies and bodies. There will also be the creation of a Europe-wide database for so-called high-risk systems that have primarily fundamental rights implications. The database will be managed by the Commission and fed with data made available by AI/R providers, who will be required to register their AI/Rs before placing them on the market.

On the other hand, with reference to specific prevention (i.e. damages compensation), we know that domestic systems are not fully "suitable" to handle liability actions for damages caused by AI/R-based products and services because the characteristics of AI/R and its relative complexity, autonomy and opacity, can make it complicated to discharge the burden of proof. The Directive introduced a "dual track" of protection. Awareness of the inadequacy of national rules also has impacts on the investment side. Ambiguous jurisprudence without *ad hoc* AI/R legislation may create many compensatory expectations, block the insurance system and, in fact, technological development. The dual track is achieved by adapting the already existing discipline on the strict non-contractual liability of the producer for damage caused by defective products and by introducing a rule on the non-contractual liability for fault from facts arising from AI/R. Specifically, the Directive deals with non-contractual liability litigations. The Directive places two powerful tools in the hands of the plaintiff/injured party. The first tool concerns the possibility of applying to the court for an order to disclose evidence. This is a preliminary litigation step by which an order can be obtained from the competent judge to have those pieces of evidence disclosed that are deemed relevant in relation to the high-risk AI/R that is suspected to have caused harm. The judge should order such disclosure limited to what is needed to support a claim for compensation. The second instrument relates to the introduction of a rebuttable presumption that is designed to define the causal link between the AI/R's non-compliance with European standards and the conduct/outcome of the AI/R (or, possibly, the failure of the AI/R system to produce an outcome) that caused the harm.

At this point, having identified the EU regulations that may be of most interest to us for purposes of this study, let us try to indicate below where they may intersect with work and risk management[27]. An initial insight can be gleaned from Recital 57 of

27 *To preliminary recap such mix between labor law regime/AI legal frame, there are at least*

the Regulation. That Recital makes a very negative, perhaps even disproportionate judgment on the use of AI/R in the management of workers. It cautions that the recruitment and selection of personnel for making decisions on promotion and termination of employment contracts, as well as for the assignment of tasks, monitoring or evaluation of workers should be classified as high-risk systems, as such systems may have a significant impact on the future of such persons in terms of their future career prospects and livelihood. It justifies this by insisting that throughout the hiring process, as well as for the purposes of evaluating and promoting individuals or continuing employment-related contractual relationships, such systems may perpetuate historical patterns of discrimination, for example against women, certain age groups, persons with disabilities, or persons of certain racial or ethnic origins or sexual orientation. These are all issues that have already been probed, including in recent scholarly studies[28], on which an articulate and plural doctrine is gradually being formed. Here, however, the fallout points of the reflection that follows is deliberately different, i.e., more in keeping with the dynamics of the factory (in the sense of a production unit) that evolves digitally because it has decided to change the production-organizational structure by orienting it towards robotics and artificial

these Recitals to be considered: Recital 9 (no prejudice to existing Union law, in particular on [...] fundamental rights, employment, and protection of workers [...]), Recital 19 (place of work/unit and the notion of "publicly accessible space should be understood as referring to any physical space that is accessible to an undetermined number of natural persons, and irrespective of whether the space in question is privately or publicly owned [...]"), Recital 44 (work activities and emotions), Recital 48 (AI and the fundamental rights protections), Recital 57 (AI high risks and employment, workers management and access to self-employment, in particular for the recruitment and selection of persons, for making decisions affecting terms of the work-related relationship, promotion and termination of work-related contractual relationships, for allocating tasks on the basis of individual behavior, personal traits or characteristics and for monitoring or evaluation of persons in work-related contractual relationships), Recital 58 (AI and social security benefits), Recital 92 (AI and no prejudice to obligations for employers to inform or to inform and consult workers or their representatives).

28 For the purposes of this study see the ideas developed by some Italian labor law scholars: M. CORTI, L'intelligenza artificiale nel decreto trasparenza e nella legge tedesca sull'ordinamento aziendale, in Federalismi.it, No. 29, 2023, p. 163; U. GARGIULO, Intelligenza Artificiale e poteri datoriali: limiti normativi e ruolo dell'autonomia collettiva, in Federalismi.it, No. 29, 2023, p. 171; L. IMBERTI, Intelligenza artificiale e sindacato. Chi controlla i controllori artificiali?, in Federalismi.it, No. 29, 2023, p. 192; A. ALAIMO, Il regolamento sull'intelligenza artificiale, in Federalismi.it, No. 25, 2023, p. 133; L. TEBANO, Poteri datoriali e dati biometrici nel contesto dell'AI Act, in Federalismi.it, No. 25, 2023, p. 198. Si v. anche S. CIUCCIOVINO, La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro, in Lavoro, Diritti, Europa, Jan. 2024; M. MAGNANI, L'intelligenza artificiale e il diritto del lavoro, in Bollettino ADAPT, Jan. 2024; M. PERUZZI, Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela, Giappichelli, Turin, 2023.

intelligence[29]. Consequently, it will be more in line with the important procedural system of anticipation and mitigation of AI/R risk in the workplace designed by the June 2020 European Frame Agreement, signed by unions and employer organizations[30], and taken up by Art. 26, para. 7, of the Regulation, which stipulates the obligation for the deployer of high-risk AI to inform unions and workers. We pointed out that there is a field to carry out this investigation that has been, at least to date, scarcely investigated and that is considered central to being able to read in a unified way all the aspects that are already the subject of academic analysis (AI/R and the exercise of employer powers). This is a very interesting scholarly field and considered useful to identify the substratum of a new branch of labor law that I call here "RLL," (Robot Labor Law) because it stems from the observation of the reality where AI/R operates in full interaction with people. It is the field of analysis referring to injury to the worker and, indirectly, to that of measures to mitigate such injury, including workplace safety procedures and the insurance regime for accidents at work/occupational diseases (for the Italian system, the INAIL regime). As anticipated above, advanced technology (AI/R, also in the form of the Frontier AI) is understood for the purposes of our research in two ways: as a tool on which personal harm to the worker may also depend, by reason of a violation of safety regulations, and as a tool that helps prevent harm to workers, by reason of elements that enable new machine-person interactions, new skills, and new forms of intelligent intervention to protect the person (e.g. exoskeleton, remote forms of intelligent control and monitoring, innovative forms of technological intervention, implemented by means of drones or other anthropomorphic robots, in the event of probable risk, including replacing workers in the context of textile, chemical, pharmaceutical, energy production, etc.)[31] .

---

29 See, by way of example, the recent Saipem case and the related firm level collective agreement of January 15, 2024. Worker protection measures were defined in the context of next-generation artificial intelligence systems. This involves constant monitoring implemented through an "artificial intelligence technological solution" that makes it possible to prevent behavior, even unconscious behavior, of non-compliance with safety obligations at construction sites. Images from cameras placed at the construction site are processed, with all the precautions that the law provides, and, artificial intelligence, selecting cases of riskiness, sends smart notifications to workers in danger.

30 See ETUC Protocol - Business Europe in https://www.etuc.org/system/files/document/file2020-06/Final%2022%2006%2020_Agreement%20on%20Digitalisation%202020.pdf *- See also Recital 92 of the Regulations, which provides for a direct link to the right of workers to be informed and consulted, including through union representatives, on anything related to high-risk artificial intelligence applied in the workplace.*

31 The corporation we have invited to participate in the research project are Eni, Enel, Ferrovie dello Stato, Leonardo, Hera, O-I International, Intesa San Paolo, and Saipem. Among the first employer associations that have given their willingness to be part of the research work are Confcommercio, Federdistribuzione, ABI, Confartigianato, CNA, Federfarma, Confindustria Digitale Anitec-Assinform. CGIL, CISL, UIL have been informed about the project.

Although under the first profile (damage caused by AI/R) there are at least two elements to be studied to verify the adequacy of the current legal framework on occupational risk and liability, one cannot fail to emphasize that, upstream of these elements, there is a methodological premise to be reiterated. We know that there is a hierarchy of sources that is created because of the distinction between matters of European competence, on which the Regulation and the Directive intervene, and labor and social security matters, that are of domestic competence. On the one hand, there is the European norm governing what is prohibited, what can be done and what should not be done in the absence of certain requirements (see above), and, on the other hand, the domestic labor/social security norm ensuring protection from harm resulting from occupational risks. The consequence of this may be summarized as follows: if a certain technological activity were prohibited by the European norm, it could hardly be the subject of compulsory insurance against accidents at work and occupational diseases.

The two elements we study to verify the adequacy of the rules on the INAIL insurance from damage resulting from AI/R in the workplace are referenced by the Italian social security system, keeping in mind that this essay represents a first phase of a broader investigation, including comparative ones. The first element concerns protection from occupational risks, which is granted by the legislature exclusively to workers who are generally exposed to a certain type of risk. According to the Italian Constitutional Court of March 2, 1991, No. 100, there is a dissociation between the legal basis of the insurance for accidents at work/occupational diseases and the related statistical-insurance method of risk. This dissociation derives from Article 38, Para. 2, of the Italian Constitution according to which the system of insurance protection is not aimed at guaranteeing in itself the risk of injury or disease, "but rather these events insofar as they affect the ability to work and are connected by a causal link to an activity typically assessed by the law as deserving protection." Thus, the protection is related to certain typical activities, beyond the concrete extent of their relative danger. This determines a complex picture to analyze because in all productive sectors the law over time has selected workers whose activity is deemed, for political and social evaluations, linked to that historical moment, to be more exposed than others to the risk of injury/disease. The professional activities, at least in the Italian system, are defined as dangerous by means of a double referral: on the one hand, it is dangerous if it is carried out on the basis of the use of some listed machines, the characteristics of which are identifiable from time to time by the norm and, on the other hand, regardless of these machines, it is dangerous if it is included in an list, not susceptible to analogical application. The second element relates to the income guarantee, which is not equivalent to the damage suffered. There is a gap between the

extent of the damage to the worker and the economic benefit payable because it is not compensation for the damage, but mere indemnity, since INAIL insurance cannot cover all the damage suffered. This indirectly determines the justification for the so-called employer's exemption from civil liability under Article 10 Act No. 1124 of June 30, 1965. This exemption scheme is partial. Indeed, it can be considered that it is now much reduced in its relative guaranteed function with respect to the employer. As it is the result of a consolidated case law[32], it can be said that the protection of the worker who has suffered an occupational injury or disease is now assigned, in the Italian system, to the competition between INAIL rules (compulsory insurance) and civil code rules on employer liability. Where the former (INAIL) ceases to operate, the latter (civil code) intervenes: the employer, theoretically, would not be called to answer with reference to the items of damage absorbed by INAIL insurance, but this almost never happens because if there is an offense and a judgment establishes that the accident occurred due to an act referable to the same employer or to the person in charge of management/supervision, the employer is called to answer for all the damages caused to the worker[33]. In addition, for damages not covered by INAIL insurance, so-called supplementary damages, there is no exemption whatsoever and the compensation system under the Civil Code applies in full (Italian Constitutional Court of July 18, 1991, No. 356).

Considering the description of the two elements, the lines of research to be developed to nurture a debate on the theoretical assumptions of the new branch of labor law (Robot Labor Law - RLL) stem from some questions that need to be asked. There are two preliminary questions about the adequacy of such a regime with respect to AI/R injury in the workplace: (i) Can the (arguably outdated) system of classification of the level of danger of professional activities also be considered sufficient to guarantee the worker's safety with respect to AI/R operating in advanced production units? (ii) Can the employer's system of liability causation, as we know it, also be applied *sic et simpliciter* in the case of AI/R, even in the form of Frontier AI? And again, within these questions, there is further room for inquiry: would it be sufficient to extend

32 For the Italian system, see Constitutional Court July 11, 2003, No. 233; Constitutional Court April 24, 1986, No. 118; Constitutional Court June 19, 1981, No. 102; Constitutional Court March 9, 1967, No. 22. See G. CORSALINI, L'azione di regresso dell'INAIL e il significato della sua autonomia, in Rivista del Diritto della Sicurezza Sociale, No. 3, 2023, p. 617.

33 Pivoting on the preventive function of Article 2087 of the Italian Civil Code and Act of April 9, 2008 No. 81, the Italian case law has held that any failure to comply with the safety obligation integrates the case under Article 590 of the Italian Criminal Code (see for the Italian case law - Supreme Court Aug. 25, 1995, No. 9000; Supreme Court Feb. 12, 2000, No. 1579), with the consequence that the employer is generally called to answer for all damages caused to the worker if the event integrates the extremes of a crime and there is a violation of occupational safety obligations.

the INAIL list of professional activities, absorbing AI/R as well? And then which AI/R, also the Frontier AI o new forms of AI? What about a scenario in which the AI/R would indirectly control machines already considered hazardous? Or, conversely, what if the AI/R controls, even in relation to downstream tasks, machines that are not considered useful in defining risky or unlisted work, but in any case, by reason of management by the AI/R, capable of causing harm? How should the notion of prevention be re-interpreted in the face of Frontier AI and, therefore, responsibility for choices that cannot be imputed to the employer? Why then should the employer be held accountable for the actions of an AI/R capable of self-determining (almost or always more) integrally, as in the case of Frontier AI? What does an AI/R risk/damage prevention system mean in the context of a business organization shaped, controlled, monitored, and coordinated by intelligent systems that replace, in terms of the third element of the employment relationship, the employer in the exercise of typical powers of control, coordination, monitoring, etc.? What processes need to be put in place? What are the most efficient ones to guarantee the person of the employee and, at the same time, not block technological and financial innovation that the employer might carry out? What safeguards and joint examination procedures can be negotiated at firm level, also considering European best practices? This would in any case not be sufficient because one must also observe the *in fieri* phase and the *ex post* phase at the introduction of an AI/R at firm level, prompting further questions: are the procedures that the current regulation on safety at work imposes[34] still efficient for the purpose of damage mitigation, even in the context of advanced technology production or where the worker co-operates with the AI/R? Can compliance, even certified by third parties, with such a procedural safety at work rule set-up result in more efficient forms of exemption for the employer in the presence of AI/R? Can AI/R risk prevention measures include special functions of joint institutions dealing with private/collective supplementary health care funds (Act of December 30, 1992, No. 502) or safety at work (Act of April 9, 2008, No. 81)? Can the special functions of supplementary health care funds include, by indication of the establishing collective bargaining agreements, forms of psycho-physical health monitoring referable to the advanced and AI/R technology most used at the firm and/or sector level? Could the psycho-physical health data reworked by such supplementary health funds be used

34 See the studies of P. PASCUCCI, Salute e sicurezza sul lavoro, responsabilità degli enti, modelli organizzativi e gestionali, in Rivista giuridica del lavoro e della previdenza sociale, No. 4, 2021, p. 537 nonché Sicurezza sul lavoro e cooperazione del lavoratore, in Giornale di diritto del lavoro e di relazioni industriali, 2021, No. 171, p. 421; M. GIOVANNONE, Responsabilità datoriale e prospettive regolative della sicurezza sul lavoro. Una proposta di ricomposizione, Giappichelli, Turin, 2024; F. MALZANI, Tassonomia UE e vincoli per l'impresa sostenibile nella prospettiva prevenzionistica, in Giornale di diritto del lavoro e di relazioni industriali, No. 177-178, 2023, p. 75.

to define by collective bargaining the health policies for prevention of occupational risk, with specific referability to the type of risk, sector, and AI/R used in a more suitable way? For the reprocessing of such data, could one imagine the use of a digital worker/citizen booklet (digital wallet), accessed by supplementary health funds and other paritarian institutions, also to bring together information on safety training, psycho-social risk situations, etc.? Would it be desirable to set up a digital wallet for recording safety training, into which information on prevention and health of the AI/R cooperating worker would also flow? Could the application of contractual regulations on supplementary health care funds that provide for this type of risk prevention initiative result in a reduction of the INAIL premium due to certified compliance with certain AI/R anticipatory health protocols? And, again, would it be possible to provide a reduction in INAIL social security contributions for employers who invest more and better in certified advanced technology aimed at mitigating risk and/or introducing targeted forms of risk prevention through the intervention of private/collective supplementary health care funds (Act of December 30, 1992, No. 502)?

4. AI, Workplace related Risks and Personal Injuries. The U.S. Legal Frame.

In late 2023, a measure specifically regulating Frontier AI was issued by the President of the United States of America. This is the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, hereafter also referred to as "EO," which is part of a broader strategy consisting of the Blueprint for an AI Bill of Rights, the AI Risk Management Framework, and the establishment of the National AI Research Resource[35] . The EO requires artificial intelligence corporations to conduct so-called AI Red Teaming, which is defined in internal, firm-level procedural terms as structured testing on the possibilities of AI/R vulnerabilities from which various kinds of risks and harm to human persons may result ("a structured testing effort to find flaws and vulnerabilities in an AI system, often in a controlled environment and in collaboration with developers of AI. Artificial Intelligence red teaming is most often performed by dedicated "red teams" that adopt adversarial methods to identify flaws and vulnerabilities, such as harmful or discriminatory outputs from an AI system, unforeseen or undesirable system behaviors, limitations, or potential risks associated with the misuse of the system."). Such an obligation is related to those corporations that develop next-generation artificial intelligence systems, here labelled as Frontier AI, and, in the U.S. legal system, also referred to as "Dual-Use Foundational Model." This implies an indirect selection of those bound to this obligation in relation to the capacity expressible by the intelligent machine. For the U.S. standard, a Dual-Use Foundational Model is the system educated in relation to a wide range of data, capable of self-supervision, capable of handling billions of parameters, cross-applicable in many contexts, capable of performing tasks from which serious risks may arise for security in general, for national and economic security, for security related to the health of citizens, for combinations of these, such as the impact on protective barriers concerning the use of chemical, biological, nuclear elements, cyber-attacks, etc., or deviation from human control. There will be a consolidation of the normative definition

35 See the reconstruction by R. CALO, Artificial Intelligence Policy: A Primer and Roadmap, in 51 U.C. Davis L. Rev., No. 51, 2017, p. 399 and, more recently, by M. E. KAMINSKI, Regulating the Risks of AI, in Boston University Law Review, No. 103, 2023, p. 1347. See for case law issues intertwining the regulation of artificial intelligence in the United States of America the analyses of B. ROGERS, Data and Democracy at Work. Advanced Information Technologies, Labor Law, and the New Working Class, MIT, Cambridge, 2023; O. LOBEL, The Equality Machine: Harnessing Digital Technology for a Brighter, More Inclusive Future, Public Affairs, New York, 2022.

of Dual-Use Foundational Model. The EO assigns secretaries to the competent State to further update various technical details to select what is already Dual-Use Foundational Model today. Specifically, in Article 4, Para. 2, a number of elements are identified that pertain, on the one hand, to the power of the AI/R model (any model that was trained using a quantity of computing power greater than 1026 integer or floating-point operations, or using primarily biological sequence data and using a quantity of computing power greater than 1023 integer or floating-point operations), and, on the other hand, to the capacity of the digital infrastructure (any computing cluster that has a set of machines physically co-located in a single datacenter, transitively connected by data center networking of over 100 Gbit/s, and having a theoretical maximum computing capacity of 1020 integer or floating-point operations per second for training AI). In addition, the EO imposes a specific obligation on service infrastructure providers, such as Amazon, Google, Microsoft, etc., to monitor and report to public authorities the conduct of foreign nationals who instruct artificial intelligence systems with the above-mentioned power characteristics also for possible cyber-attacks on the country's security (so-called "malicious cyber-enabled activity"). This means having made a different choice from the European one: in the U.S. legal system, monitoring of all possible AI/R applications, including those of the new generation, is carried out through a system of large clouds. The positive aspects of this choice are offset by many critical issues, which relate to possible abuse by cloud infrastructure providers or the ability of subjects to carry out real investigations of possible malicious cyber-enabled activities.

We have indicated above that the genotype that can be constructed to initiate the comparison between the European norm and the U.S. norm arises in relation to three issues: (i) What risk and harm that legal system intends to select for the protection to be provided in relation to the Frontier AI's ability to interact with the worker? (ii) What mitigation and prevention measures does that legal system intend to introduce to address the need for protection related to those risks and harms? (iii) What institutions that legal system intends to put in place to implement that protection, including in terms of individual and collective enforceability?

It is understood that the U.S. norm has a broader geo-political scope than the European one, having set the definition of Frontier AI based on criteria pertaining to the power of the information system and the infrastructural system that supports it[36] . This may

36 See the reconstruction done by some Stanford University scholars, especially the director of the Center for Artificial Intelligence Research (HAI), FEI-FEI LI, The World I see. Curiosity, Exploration, and Discovery at the Dawn of AI, Flatiron, USA, 2023.

also have an impact in the field of the case law inquiry (AI/R determining or preventing harm to the worker in advanced production units). In particular, at least three elements are intertwined in the North American system of labor condition protection, which are mitigation, insurance, prevention. The first element relates to individual litigation that each aggrieved worker can not initiate to be compensated for the harm suffered (Torts Suits)[37]. The second element concerns the provision of a financial benefit and/or medical services to the injured worker (Workers' Compensation). This is an insurance scheme which, in the event of an occupational injury or illness, is activated in favor of the worker[38] . In order to access this system, it must be investigated on a case-by-case basis whether or not the occupational accident or illness may result in the payment of the benefit[39]. This involves conducting a preliminary test that is based on an analysis of the facts (is it really a work accident? Did the accident occur "in the course of the employment" or "arising out of the employment"? What does the injury to the worker consist of?). The notion of an accident/illness is generally open-ended, and on a case-by-case basis there is a jurisprudential test of inquiry[40]. Application of the Workers' Compensation regime theoretically creates a form of exemption of the employer from other liabilities and claims. There are not situations that may determine the right of the worker to add to the social-insurance benefits also compensation for further damage[41]. The third element relates to the structure of prevention and safety at work, which is structured according to the model of risk anticipation and management of the effects of risks through certain institutions that oversee the proper fulfillment of the employer's obligations to protect[42]. From my point of view, the questions already posed above, which were useful in initiating the lines of research on RLL, in relation to the Italian and to some extent the European system, can also be repeated, with some

37 Exclusivity is one of the founding principles of workers' compensation. In exchange for a no-fault system, workers's compensation may be considered the exclusive remedy against the employer for a worker injured. There are some exceptions to the exclusivity remedy principle that may allow an injured employee to bring a tort suit against the employer. See Farewell v. Boston&Worcester Rail Road Corp. Supreme Judicial Court of Massachusetts 45 Mass. (4 Met) 49 (1842).

38 See R. EPSTEIN, The Historical Origins and Economic Structure of Workers' Compensation Law, in Georgia Law Review, No. 16, 1982, p. 775.

39 The administration of Workers' Compensation varies from state to state. Some states have welfare-like administration, with centralization of activities in one agency. Other states have mechanisms similar to private insurance.

40 See Matthews v. R.T. Allen&Sons, Suprem Judicial Court of Maine, 266 A.2d 240 (1970).

41 See Millison v. E.I. du Pont de Nemours&Co., Supreme Court of New Jersey, 501 A.2d 505 (1985).

42 Reference is made to the OSHA agency discipline described and reported here - https://www.osha.gov For more theoretical reflection on the function of occupational safety in the U.S. system, see this study A. P. BARTEL, L. GLENN THOMAS, Predation Trough Regulation: The Wage and Profit Effects of the Occupational Safety and Health Administration and the Environmental Protection Agency, in Journal of Law&Economics, No. 30, 1987, p. 239.

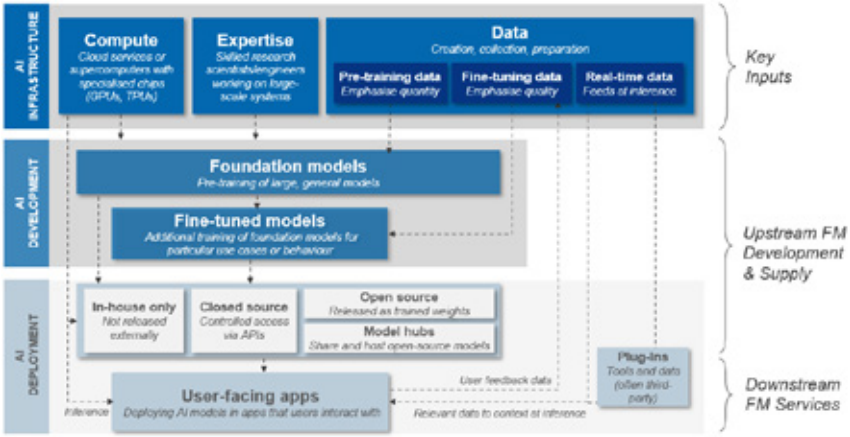
differences, for the North American system. These questions can be enriched by the following: are the measures to prevent psycho-social risk arising from AI/R operating at firm level adequate or not? Does the North American system provide sufficient economic resources for injury compensation and prevention? Is such a system fair with reference to the various types of individual employment contracts, unionized and non-unionized workers, migrants, etc.? Is an open system of defining occupational injury/illness, not pre-defined through lists, efficient with reference to AI/R operating at firm level?

5. Conclusions. For a New Branch of Labor Law d (RLL – Robot Labor Law).

Any reflection cannot fail to start from a snapshot of the whole system that is evolving, both technically and normatively, at the level of Western legal systems. It is what in the jargon of any risky activity can also be called the Vantasner Danger Meridian, a kind of line that fixes the timeline of the real riskiness. This may determine an alarmist technophobia or bring closer the more desirable pragmatic attitude of the regulator. We understood that AI/R, managed by Frontier AI, imposes on us a (transnational/national) regulation that cannot fail to follow the entire life cycle of the intelligent machine. From the initial development phase, up to testing and then introduction into the market, with periodic checks on the outcomes, whether positive or negative, including in terms of harm to the workers. The more advanced the artificial intelligence, as in the case of Frontier AI, the more risks there are. And this becomes even more true for forms of AI/R that cooperate, interact, coexist with workers. Frontier AI risks may arise from three factors that regulation, including labor law regulation, must somehow intercept[43] . The first factor relates to unexpected problems, generally arising from the fact that Frontier AI presents the actual ability to self-determine. It may do this precisely at the stage of distribution in the marketplace, or in performing tasks for and with end users. The second factor concerns the issue of security in the deployment phase. End-users could demand further development from Frontier AI that is not aimed at legitimate ends, in the sense of legally unjustifiable, which are not entirely foreseeable (see above) and could multiply its harmful scope. The third factor relates to the problem of proliferation, since Frontier AI is generally open-sourced, allowing anyone to develop additional models, even those with purposes that are not legally justifiable. This often involves model theft or cyber-attacks that allow access to patterns that have some strategic value. The diagram that may clarify these factors is taken from the documentation of the AI Safety Summit in London, 2023:

---

43 Here we follow the approach that was developed by the group of scholars and experts, whose work became the subject of reflection and a point of reference for the policies of law by the intergovernmental conference dealing with Frontier AI in London in 2023, by the European Union, and by the United States of America. See M. ANDERLJUNG, J. BARNHART, A. KORINEK, J. LEUNG, C. O'KEEFE, J. WHITTLESTONE, S. AVIN, M. BRUNDAGE, J. BULLOCK, D. CASS-BEGGS, B. CHANG, T. COLLINS, T. FIST, G. HADFIELD, A. HAYES, L. HO, S. HOOKER, E. HORVITZ, N. KOLT, J. SCHUETT, Y. SHAVIT, D. SIDDARTH, R. TRAGER, K. WOLF, cit., 2023.



To meet these challenges regulation, including labor law regulation, would be called upon to set (i) international safety standards[44], (ii) enforceable models of compliance with these standards, (iii) transparency and information indices and schemes, and (iii) a macro-regional, no longer just domestic/national, insurance model against Frontier AI-derived labor risk. Let's start with international standards, which should be the object of a form of conventional or treaty regulation, signed at least among the most relevant Western legal systems, including certainly the United States of America, the EU institutions, and Great Britain. Standards should be the subject of study, experimentation, risk auditing, data collection and academic, social, etc. comparison. Auditing schemes aimed at verifying standards should be introduced. This would make it possible to construct standards that are elaborated through the technique of risk assessment, with the weighting of what is actual capacity for danger arising from AI/R (also in the form of the Frontier AI), even at firm level, and its controllability. Assessment by specialized third parties, auditors or the like, would in fact create greater credibility based on what is made known externally, even with the help of such specialized third parties. Hence it would move the construction of standardized protocols on how AI/R should be distributed and what safeguards should be introduced as well as the processes for periodic review of risks and measures to be taken.

44 *In the United States of America, the National Institute for Standards and Technology has created the AI Risk Management Framework. The National Telecommunication and Information Agency has started a path to introduce AI-related insurance policies. In Great Britain, the AI Standards Hub has been established. The European Union has asked the agencies CEN and CENELEC to develop standardized safety models on artificial intelligence.*

But all this would probably not be sufficient in any case. Compliance models should be made enforceable, balancing their impact on potential technological-economic development. This is the real challenge for a lawmaker. An avenue of voluntary adherence to these standards should be promoted, also at the transnational level, with the creation of codes of conduct (self-regulation), certifications, issuance of special prior qualification/license in order to be able to create and then deploy, also for person/machine collaborations, AI/R and Frontier AI systems, supervisory authority actually active in inspections, etc. This would make it possible to activate cross-checking systems in high-tech production units, new occupational safety procedures linked to certified vocational training for each worker collaborating with AI/R, special patents for such workers, etc. Hence it could be possible to build up a compulsory, at least macro-regional, European insurance system, referable to what in Italy is INAIL, for AI/R risks at workplace level. We know that the greater the scope, ordinal and geographic, of an insurance/pension/social security scheme, the greater the benefits for all. The imponderability of this approach specifically relates to the impact on the rule that prevents the harm and risk to the worker as well as allows compensation for the harm under advanced AI/R systems such as Frontier AI. It is an imponderability that stems from the fact that the most appropriate combination of prevention at workplace level and compensation for further harm, in the event of the exercise of employer powers by Frontier AI or any form of artificial intelligence even more advanced, is not yet well known. This determines an impact, also based on interdisciplinary and transnational scientific studies that reference reality, not just the mere idea, such as the one we intend to carry out. With reference to the fact that one cannot fail to verify the adequacy of social-insurance regimes for accidents at work and occupational diseases with respect to the new generation artificial intelligence, adequacy to be measured with respect to the obligation to contribute benefits and the employer's liability regime, while still verifying the efficiency of safety at work systems. These safety at work systems are all or almost all set in the past, not in the future, that is, for a machine subjected (with enormous variability) to the indications of the worker, not for an intelligent machine, Frontier AI, that coordinates, imposes, controls, directs even the worker's performance.

All this represents the most important challenge for the scientific community, the policymaker and, ultimately, for those who work in the industrial relations systems, unions and employer organizations. It is a direction to sense and then decide to follow, knowing that direction is always more important than speed because it represents the whole over the part.

I Working Papers nascono dall'attività di progetto e di studio del gruppo di ricerca della Fondazione Giacomo Brodolini.
Sono uno strumento agile di informazione che permette la sistematizzazione e la diffusione dei lavori realizzati sulle principali tematiche d'interesse della Fondazione.

Working Paper
Fondazione Giacomo Brodolini

Fondazione Giacomo Brodolini