INTEGRATING ARTIFICIAL INTELLIGENCE INTO THE MORTGAGE CREDIT MARKET

ROSAMARÍA GARCÍA TERUEL

ASSISTANT PROFESSOR OF CIVIL LAW
UNIVERSITY OF BARCELONA

WORKING PAPER 1/2023
**Abstract:** In order to avoid borrowers’ over-indebtedness, Directive 2014/17/EU introduced the obligation to assess the creditworthiness prior to concluding a mortgage loan. To carry out this task, novel technological applications that use artificial intelligence and machine learning are emerging, as well as the use of alternative data, which ensures more predictive power. The purpose of this article is to analyze whether it is possible to use this technology in accordance with Directive 2014/17/EU and the General Data Protection Regulation. It also discusses the possible advantages and challenges that this technology may have.

**Title:** Integrating artificial intelligence into the mortgage credit market

**Keywords:** mortgage credit directive, data protection, alternative data, creditworthiness assessment, housing

**Resumen:** Para evitar el sobreendeudamiento de los prestatarios, la Directiva 2014/17/UE introdujo la obligación de efectuar una evaluación de solvencia previa a la conclusión del contrato de crédito inmobiliario. Para llevar a cabo esta tarea, están surgiendo nuevas aplicaciones tecnológicas basadas en inteligencia artificial y machine learning, así como el uso de datos alternativos, que aseguran tener un mayor poder predictivo. La finalidad de este artículo es analizar si, de acuerdo con los preceptos de la Directiva 2014717/UE y del Reglamento General de Protección de Datos, es posible usar esta tecnología, así como las posibles ventajas y retos que conllevaría su aplicación.

**Título:** La integración de la inteligencia artificial en el mercado del crédito hipotecario: interacciones entre la protección de datos y la directiva de crédito inmobiliario en las evaluaciones de solvencia.

**Palabras clave:** directiva de crédito inmobiliario, protección de datos, datos alternativos, evaluación de solvencia, vivienda.

**Resum:** Per a evitar el sobreendeutament dels prestataris, la Directiva 2014/17/UE va introduir l’obligació d’efectuar una avaluació de solvència prèvia a la conclusió del contracte de crèdit inmobiliari. Per a dur a terme aquesta tasca, estan sorgint noves aplicacions tecnològiques basades en intel·ligència artificial i machine learning, així com l’ús de dades alternatives, que asseguren tenir un major poder predictiu. La finalitat d’aquest article és analitzar si, d’acord amb els preceptes de la Directiva 2014717/UE i del Reglament General de Protecció de Dades, és possible usar aquesta tecnologia, així com les possibles aventatges i retes que comportaria la seva aplicació.

**Títol:** La integració de la intel·ligència artificial en el mercat del crèdit hipotecari: interaccions entre la protecció de dades i la directiva de crèdit inmobliari en les avaluacions de solvència.

**Paraules clau:** directiva de crèdit inmobliari, protecció de dades, dades alternatives, avaluació de solvència, vivienda.
SUMMARY

1. INTRODUCTION ......................................................................................................... 3
   1.1. THE EMERGENCE OF FINANCIAL TECHNOLOGIES ........................................................... 3
   1.2. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TO ASSESS THE CREDITWORTHINESS …… 4
2. CREDITWORTHINESS ASSESSMENT WITHIN THE MORTGAGE CREDIT DIRECTIVE: DIFFERENCES WITH
   TRADITIONAL CREDIT RISK ASSESSMENTS .......................................................................... 5
3. LEGAL POSSIBILITIES OF USING AI FOR CREDITWORTHINESS ASSESSMENTS ................................. 6
4. CHALLENGES TO THE PROTECTION OF PERSONAL DATA WHEN USING NON-TRADITIONAL DATA .......... 8
5. POSITION OF THE EUROPEAN COMMISSION AND FURTHER DEVELOPMENTS IN THE AI ACT ........... 10
6. REFERENCES ........................................................................................................... 11

1. INTRODUCTION

1.1. THE EMERGENCE OF FINANCIAL TECHNOLOGIES

New technologies applied to the financial sector, also known as Fintech, have been
developed over the last decades. Financial institutions have traditionally relied on
technologies to better perform their services, especially when an important amount of
data shall be assessed. However, the Great Financial Crisis (GFC) 2008 was one of the
most important factors that make financial technologies emerge (Arner et al, 2015), with
the commitment to make traditional processes more effective and to achieve a better
inclusion of the most vulnerable in the financial sector (International Monetary Fund
Annual Report, 2020). Financial technologies include electronic payments, digital
advisory systems (the so-called ‘roboadvisors’, algorithms that manage clients’
portfolios; Abraham et al. 2019), peer-to-peer lending through crowdfunding service
providers¹, big data techniques², machine learning (ML), and also artificial intelligence
(AI) to assess the creditworthiness of potential borrowers, i.e., the process to determine
the ability of borrowers to repay a loan.

¹ This phenomenon has been recently regulated by Regulation EU 2020/1503, on European crowdfunding
Regulation 2020/1503 intends to harmonise the legal framework of crowdfunding service providers so
that they shall follow the same procedure to operate. Among others, some obligations are imposed on
these providers: the obligation to undertake due diligence on each project, to inform national supervisors
of the total amount of crowdfunding projects granted, prudential requirements, etc.

² Big data, including the so-called ‘data science’, refers to the automated processing of complex and large-
scale data through informatic algorithms and advanced techniques of data processing, according to the
European Parliament Resolution of 14.3.2017, on fundamental rights implications of big data: privacy,
data protection, non-discrimination, security and law-enforcement (2016/2225(INI)).
This paper aims to assess the possibilities of using AI and ML for creditworthiness assessments in the mortgage credit market, evaluating their opportunities and challenges and whether it is possible to use them according to Directive 2014/17/EU. The further effects that these technologies may have in the field of protection of personal data are also taken into account.

1.2. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TO ASSESS THE CREDITWORTHINESS

Artificial intelligence (AI) is a software system that works as a human mind does to undertake certain tasks (Raynor, 1999). The ISO/IEC 22989:2022 also defines ‘artificial intelligence systems’ as ‘engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives’. These AI systems are usually autonomous, meaning that they may work and modify their goals without external intervention, and use machine learning (ML) and deep learning (DL) techniques, that allow AI to improve their performance through the experience and the use of a vast amount of data and computer power (McKenzie, 2018). Therefore, AI may decide the course of action without human intervention, learning from the experience, which may make them unpredictable.

AI is used for several applications: self-driving cars, software to detect diseases, home assistants, natural-language processing, etc. Apart from that, AI, jointly with ML or DL, can also be used to assess whether a certain borrower will repay or not a given amount. And, to better assess the creditworthiness of the borrower, AI may make use of ‘alternative or non-traditional data’, obtained mainly from public websites, such as data from social networks. According to Pascual (2021), following the reports of the Center for Financial Inclusion at Action and of the Asian Banking School, non-traditional data may include shopping patterns, data from social networks (marital status, political affiliations...), geolocation, payment of utilities, as well as public information found through public websites, such as crimes databases, information in newspapers, connections in professional social networks, etc.

Hurley and Adebayo (2016) detected some credit scores using non-traditional data in the United States, such as the FICO – Expansion Score, which incorporates information from utility bills, membership club records, judgements, etc., or LendUp, using social network data, such as how quickly a user scrolls through its site. However, is it still under discussion whether these technologies may contribute to a better performance in creditworthiness assessments or the challenges that they pose to consumers. Some studies suggest that AI based on alternative data found in social media may have a high classification accuracy compared to other methods, predicting more than 80 per cent of the defaults (Zhang, et al. 2016) while making the process more agile. Óskardóttir et al. (2019) confirm that using telco data (e.g., calling behavior or geolocation) may have the

---

3 See also Banco de España (2022) Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost. Documentos de trabajo n. 2032.
same predictive power as traditional data\textsuperscript{4}, so that ‘thin credit files’ could especially benefit from using alternative data. Niu et al. (2019) also consider that social network data may positively contribute to the financial inclusion of thin credit files. Data from social networks indicated that several parameters have an impact on defaults: people with more contact information are more likely to default, and the longer a mobile number is used, the lower the probability of a default. All these results suggest that incorporating alternative data through AI and ML techniques could improve financial inclusion in the mortgage sector. This could facilitate access to housing of people traditionally excluded from the mortgage credit market. Moreover, being creditworthiness assessments a difficult task that increases bureaucracy (Livada, 2019), the use of these technologies could speed up this process.

Nevertheless, as Remolina (2022) points out, AI and ML may also pose several challenges, such as the risks of biases and discrimination and personal data breaches. Regarding discrimination, as stated by some authors (Remolina, 2022, Pascual, 2021, p. 570), while using a vast amount of data, AI may discriminate people on grounds of race, nationality, gender, etc. (e.g., the more certain races default, the more probability the AI considers a person of that race as defaulter). This risk is also highlighted by the European Union within the AI Act\textsuperscript{5} (recital 37), confirming that ‘AI systems used to evaluate the credit score or creditworthiness (…) may lead to discrimination of persons or groups and perpetuate historical patterns of discrimination, for example, based on racial or ethnic origins, disabilities, age, sexual orientation, or create new forms of discriminatory impacts’. Moreover, the applicant may not be aware of the reasons why a certain loan has been rejected: on the one hand, algorithms are often treated as a trade secret, so its owner does not have the intention to reveal the logic behind it or the data they use. On the other hand, complex algorithms using ML may start to take decisions in a non-explainable way, losing human control (Pascual, 2021).

2. CREDITWORTHINESS ASSESSMENT WITHIN THE MORTGAGE CREDIT DIRECTIVE: DIFFERENCES WITH TRADITIONAL CREDIT RISK ASSESSMENTS

Directive 2014/17/EU\textsuperscript{6} (hereinafter, MCD) took an innovative approach to both prevent over-indebtedness and to increase the transparency of the European Mortgage Market through requiring a creditworthiness assessment of consumers who apply for a mortgage loan (Collado, 2019). Traditionally, consumer protection in the credit market was ensured by obliging lenders to properly inform borrowers. However, this type of

\textsuperscript{4} This study concludes that, inter alia, not having a potential defaulter neighbour reinforces the idea that applicants will not breach their obligations.


protection was deemed ineffective after the GFC and the EU decided to tackle some reckless banking practices through the implementation of the ‘responsible lending’ principle, which includes the obligation to undertake a creditworthiness assessment.

According to Article 4.17 MCD, ‘creditworthiness assessment’ means the evaluation of the prospect for the debt obligation, resulting from the credit agreement, to be met. The lender shall undertake a throughout assessment of the creditworthiness, according to Article 18.1 MCD. If the assessment indicates that the consumer will probably default, lenders shall deny the credit (Article 18.5a MCD), while not being forced to concede the mortgage loan if the result is positive. The lender, to undertake the assessment, may consult internal or external sources (Article 20.1 MCD).

Before the MCD obliged lenders to assess the creditworthiness, it was a common practice to analyse prospective borrowers through ‘credit risk assessments’ (Collado, 2019), using computing technologies and credit scores. With this system, lenders could evaluate the risk associated with crediting a certain borrower and decide whether to grant credit or not: the aim of credit risk assessment is to determine if a particular operation is within the lender’s risks, while the creditworthiness assessment aims at preventing borrowers from getting into over-indebtedness. As Cherednychenko and Meindertsma (2019) confirm, ‘the lender’s duty to assess the consumer’s creditworthiness (…) should by no means be limited to the assessment of credit risk and should also include the borrower-focused creditworthiness check’. Another important difference between traditional credit risk assessments and creditworthiness assessments is the consequence of a certain result: in the credit risk analysis, if the borrower is deemed to default, the lender may decide to grant the credit or not, considering their risk policies. However, in creditworthiness assessments, if the result indicates that lenders will probably breach their obligations, lenders are obliged to deny the mortgage loan (Article 18.5a MCD).

Lenders traditionally used statistical methods for their credit risk assessments, including complex software systems that are also being used nowadays for creditworthiness assessments. However, AI and ML differ from these traditional statistical techniques, as AI based on ML may become more sophisticated and complex, analyzing a vast amount of data (including non-traditional data), while learning to perform this task through experience (Pascual, 2022). Moreover, while statistical methods assess fixed criteria (e.g. income, debt, age, working experience…), the criteria applied by an IA may not be clear or explainable.

3. LEGAL POSSIBILITIES OF USING AI FOR CREDITWORTHINESS ASSESSMENTS

When analyzing whether a lender, in a mortgage loan, may use AI, ML, and alternative data to undertake the creditworthiness assessment, it is relevant to note that the MCD does not specify which method to follow for that purpose. Also, the MCD only indicates that the assessment shall be carried out based on information on the consumer’s income and expenses and ‘other financial and economic circumstances’. This fact gives room for Member States to regulate specific procedures to undertake the assessment or data required.

Nevertheless, some articles of the MCD could limit the use of AI systems as well as of non-traditional data. First, art. 18 MCD foresees that:

‘Member States shall ensure that the procedures and information on which the assessment is based are established, documented and maintained’.

This obligation imposed on Member States involves that the steps required to undertake the assessment shall be clear enough among actors. However, when using complex AI systems (especially those based on ML or DL), the outcome of this software may not be explainable at all, as commented before. Some AI take decisions following what they learned previously, making it a difficult task, even for the software developer, to understand the variables taken into account. Therefore, it is doubtful whether, to fulfill what is established in Article 18.1 MCD, the lender shall also maintain, document, and establish the logic of the AI and the parameters taken into account to consider a borrower as a potential defaulter or non-defaulter.

Although the MCD does not directly address this problem, the General Data Protection Regulation\(^8\) (GDPR), in its turn, protects natural persons when processing personal data by automated decision-making means, including profiling. According to Article 4.4 GDPR, profiling means ‘any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular, to analyze or predict aspects concerning that natural person’s performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements’. As confirmed by recital 71 GDPR, profiling may include creditworthiness assessments, as they involve the prediction of a person’s performance through automatic means.

GDPR may have a twofold impact on creditworthiness assessments for mortgage credits.

- First, the borrower will be entitled to obtain human intervention from data controllers (i.e., lenders). Also, borrowers may ask for an explanation of the decision reached after the automated assessment (Article 22.3 GDPR).
- Second, when the credit is denied solely based on the automated processing of data, the borrower (data subject) will be entitled both to contest the decision

and not to be subject to it (Article 22.1 GDPR), with some exceptions included in Article 22.2 GDPR to be analyzed on a case-by-case basis.

As a result, although the MCD does not specify whether the AI could be used for creditworthiness assessments, in principle, it is feasible to include these software systems in this process, provided that the GDPR rights are met: the decisions, in principle, could not be solely met on what the AI decides, and the decision-making process shall be explainable to the borrower (to check if the decision was taken properly, without discrimination, etc.). Complex AI based on ML or DL may not be appropriate for this task. This interpretation also goes in line with the Charter of Fundamental Rights of the European Union9, which recognizes the protection of personal data (Article 8) and non-discrimination of individuals (Article 21), and with the Ethics Guidelines for Trustworthy AI of the EU Commission, promoting the ‘principle of explicability’10.

4. CHALLENGES TO THE PROTECTION OF PERSONAL DATA WHEN USING NON-TRADITIONAL DATA

The second controversial aspect of using AI in mortgage creditworthiness assessments is the possibility of using non-traditional data. The MCD does not specify which data may be collected or used for that purpose, since the intention of the EU legislator was to give room to Member States in this field. According to Article 20 MCD:

‘The assessment of creditworthiness referred to in Article 18 shall be carried out on the basis of information on the consumer’s income and expenses and other financial and economic circumstances which is necessary, sufficient and proportionate’.

While transposing the MCD, some countries referred to the specific information that had to be collected. For example, Spanish Act 5/201911 foresees that the evaluation shall consider ‘relevant factors’, including employment status, current income, foreseeable income, assets owned, savings, fixed expenses, and commitments already assumed. Also, the Catalan Consumers’ Code12 in Article 263-2 mentions that the lender shall consider, among other criteria, the consumer’s current and future income, savings, debts, and financial commitments, according to the information provided by the

---


10 Ethics Guidelines for Trustworthy AI. Independent high-expert group on artificial intelligence set up by the European Commission. 2019.


consumer. The Portuguese Aviso do Banco de Portugal núm. 4/2017\textsuperscript{13} foresees in Article 5 that the lender shall take into account, among others, the following elements: nature, amount and characteristics of the credit agreement; age and employment status of the consumer; income; regular expenses; and compliance with other obligations assumed by the consumer.

As stated, both MCD and national transposition establish non-exhaustive data to be collected, so that, from this point of view, non-traditional data could theoretically be collected. However, as stated in the MCD, the data shall be ‘necessary, sufficient and proportionate’. According to Collado (2021), ‘necessary’ and ‘sufficient’ mean that the required information shall allow the lender to undertake a proper assessment; the sufficiency will depend not only on the personal situation of the consumer but also on the length and amount of the credit (ECJ 18.12.2014\textsuperscript{14}). As it pertains to alternative data, the challenge lies in determining if it is ‘proportionate’: Article 20.3 MCD refers to data that is 'limited to what is necessary to conduct an adequate creditworthiness'. This definition suggests that it cannot be used more information than the one necessary to undertake the assessment, which would be in line with the principle of data minimization of Article 5.1.c GDPR. According to this principle, the processing of personal data shall be adequate, relevant, and limited to what is necessary considering the purposes for which they are processed. Also, if lenders want to use alternative data, this shall relate to ‘financial and economic circumstances’, as stated in Article 20 MCD.

Considering that some AI are making use of alternative data for creditworthiness assessment, the Proposal of a Directive on Consumer Credit\textsuperscript{15} specifically regulates this phenomenon. According to the explanatory memorandum of the proposal, 'The principles of data minimisation, accuracy, storage limitation as laid down in Article 5 of the GDPR govern the use of data to conduct creditworthiness assessments'. Without diverging from the GDPR, this proposal aims to address the concerns identified in the processing of personal data that are specific to practices observed in the consumer credit market, i.e., the use of alternative sources of data for creditworthiness assessments or the transparency of assessments carried out using machine learning techniques. Article 18 of the proposal also foresees that Member States shall ensure that the consumer is entitled to obtain human intervention and a clear explanation of the assessment when using automated processing of data, while data obtained in social networks or other special categories of data, such as health data, cannot be processed.

In light of this, when interpreting the concept of 'limited' of the MCD with other sources, such as the GDPR, non-traditional data can't always be used: if alternative data is considered 'personal data' (i.e., when a natural person can be identified because of an

\textsuperscript{13} Published in 2 S, 184 Supl. Part E, of 22.9.2017.

\textsuperscript{14} Case CA Consumer Finance SA v. Ingrid Bakkaus, Charline Bonato and Florian Bonato. ECJ Judgement of 18.12.2014 (C-449-13).

identifier, such as location data, name, social identity, etc. Article 4.1 GDPR), and if there is enough traditional data to undertake the assessment, using AI and alternative data shouldn’t be proportional to meet the lender’s obligation. On the contrary, when the alternative data is not considered ‘personal data’ (i.e. no natural person can be identified through the data processed), or even when the lender cannot undertake the assessment with traditional data (e.g. in specific cases of ‘thin credit files’), using alternative data through AI would be possible, provided that the assessment is not solely undertaken by the AI, the process is explainable to the borrower and the data refers to economic and financial circumstances.

5. POSITION OF THE EUROPEAN COMMISSION AND FURTHER DEVELOPMENTS IN THE AI ACT

The EU Commission called for advice from the European Banking Authority (EBA) in 2021 regarding several issues related to the MCD\textsuperscript{16}. In the report, the EU Commission recognizes new challenges of digitalization for the mortgage market: in particular, the use of AI and non-traditional data to assess the creditworthiness of mortgage borrowers. The EU Commission asks the EBA if there are any possible risks for consumer protection arising from the use of AI systems for mortgage borrowers’ creditworthiness assessments. The EBA replied in 2022\textsuperscript{17} confirming that, considering the detriment that consumers may suffer because of decisions based on automated systems, these risks (e.g., discrimination) may be addressed in the Artificial Intelligence Act (AI Act).

Indeed, the AI Act proposal specifically foresees these types of systems. AI Act regulates different types of obligations depending on the kind of AI which wants to be deployed. Some AI systems are explicitly forbidden, such as those that implement subliminal techniques beyond a person’s consciousness, those AI used by public authorities for the evaluation of the trustworthiness of natural persons, or real-time remote biometric identification systems, with some exceptions (Article 5 AI Act). According to Article 52 of the AI Act, ‘certain AI systems’, such as AIs that recognize emotion, or AIs that manipulate image, audio or video content, must inform natural persons of their interaction with the software. Finally, AI Act foresees ‘high-risk systems’, imposing more obligation on its deployment. Annex III includes a list of ‘high-risk systems’, including AI used by a judicial authority to interpret facts and the law, polygraphs and similar tools to detect the emotional state of a natural person, and ‘AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems put into service by small scale providers for their own use’ (Annex III para. 5).

Therefore, AI for creditworthiness assessments in the mortgage market, once the AI Act enters into force (expected in 2023), will be considered ‘high-risk’ and subjected


\textsuperscript{17} EBA. Opinion of the European Banking Authority on the European Commission request for technical advice on issues related to the Mortgage Credit Directive (EBA/Op/2022/07). 23.6.2022.
to obligations included in Chapter 2 Title III. Among others, providers of AI for creditworthiness assessments shall implement risk management systems\textsuperscript{18}, training, validating and testing data sets, drawing up technical documentation before the AI is placed on the market, ensuring a certain level of traceability of the AI (enabling automatic recording of events, checking the input data and monitoring operations), ensuring transparency to users, so that they may interpret the output, or ensuring human oversight by natural persons in order to prevent risks to fundamental rights (e.g. against discrimination practices). Consequently, following AI Act entry into force, the service provider of a creditworthiness assessment through AI shall meet several obligations to prevent the possible risks of discrimination, data protection, etc.

It is still unknown the effects that these obligations will have on developers of those AI systems, and whether the use of new technologies will be hindered by ensuring consumers’ protection. For the moment, the Association of Consumer Credit Information Suppliers (ACCIS) gave feedback on the AI Act\textsuperscript{19}. ACCIS considers the AI Act to be too broad and includes creditworthiness assessments using software that has been tested for several years without any risks (i.e., those utilizing statistical methods rather than machine learning). Further, ACCIS considers that, in some cases, even when using ML that poses higher risks to consumers, operating with this software has no significant impact on the assessment of creditworthiness, so they can be excluded from Annex III’s ‘high-risk’ list. What is true is that the risks of these technologies depend on the intended use and on the logic behind it (e.g., ML, DL, or just statistical methods), as well as the possible explicability of each decision. A strong regime imposed on those low-risk systems may hinder the use of novel technologies, which are useful for protecting consumers and integrating certain groups (thin credit files) into the mortgage credit markets.

6. REFERENCES


\textsc{Banco de España} (2022) Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost. Documentos de trabajo n. 2032.

\textsuperscript{18} The risk management system requires a continuous process of detecting foreseeable risks, risk management measures, proposing measures to eliminate or reduce risks as much as possible, provision of adequate information, etc.

\textsuperscript{19} Association of Consumer Credit Information Suppliers. Feedback on the European Commission’s proposal for a Regulation laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.


Este obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.