# Identity verification systems: the Distributed Know Your Customer platform

Diego Ponte, Department of Economics and Management, University of Trento, diego.ponte@unitn.it

Carlo Rizzi, Cherrychain, carlo@cherrychain.it

## 1 Introduction

Over the last decades, the need for a secure and interoperable Know Your Customer (KYC) process has become growingly important, especially in the financial sector. Such importance is also highlighted by different international and European regulations such as EIDAS, PSD2, and AML. Nonetheless, empirical findings show that the majority of current identity management systems are based on inefficient technical models. These show many criticalities in managing the KYC process and generate costs of up to 500 million USD per year per bank (Parra Moyano & Ross, 2017). Without an international standard for a standardized KYC procedure, it is difficult for companies to remain compliant both locally and globally; companies have to follow different regulations in the countries in which they operate, with non-standard KYC programs (Garber et al, 2021; NTT, 2020). Such systems are mainly based on centralized or federated models that have many drawbacks. Self-sovereignty (SSI) models based on distributed systems might theoretically outperform the traditional models. Indeed, the usage of solutions based on Distributed Ledger Technologies (DLT) could make identity management much more efficient and could improve the KYC process, leading to time and cost savings as well as the better overall security of data. On the other hand, distributed systems for KYC are still in an exploratory phase (Benchaya et al, 2022; Qadir et al, 2022). While many scholars have explored such model mainly based on reviews, theoretical or conceptual approaches (Eduardo Demarco, 2020; Singhal, at al. 2020; Stockburger, et al. 2021; Ostern, & Riedel, 2021; Čučko, & Turkanović, 2021), this paper shows the benefits of a peculiar blockchain-based KYC model, called Distributed Know Your Customer (DKYC) which was tested using a sandbox. Based on the results of the sandbox populated with an ecosystem of institutional bodies and other regulated companies of the financial and non-financial sectors, this paper discusses the contribution of self-sovereignty models for KYC by showing the benefits and implications of the DKYC model. The model showed to be able to facilitate and make the process of exchanging data among the actors of the ecosystem faster and less onerous. The authors elaborate on the results to show the main benefits of such a solution: more efficiency in terms of time and costs and new opportunities in terms of new business models.

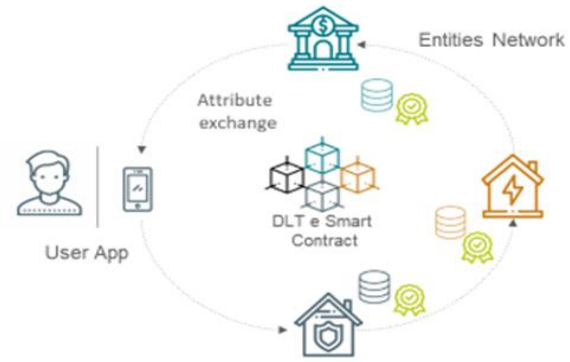
## 2 The Distributed Know Your Customer model

The DKYC model proposed in this paper provides that the data and documents are made available, always with the explicit consent of the customer, by one entity to other thanks to the fact that the system allows the sharing of data directly between entities (figure 1). In this sense, an entity can assume the role of custodian of sensible data; the data can be shared only with the permission of the customer (EU 2018). Therefore, the same user has the opportunity to check and arrange the documents already kept by different organizations with which it has an active relationship. For example, a user can open a new account at a Utility, thanks to the secure, transparent, and unalterable exchange between reliable data sources held by the Entities in the blockchain network. DKYC is also a distributed system as the data and documents reside where the business relationship was born and kept. The balance between confidentiality and source reliability is achieved thanks to the traceability of the compliance level used in the opening of the report by each entity that classifies each available attribute (data or document). In the DKYC the same data or document is accompanied by the level or degree of compliance for which it was subjected to validation in the KYC process. For example, the same residential address is validated by two anonymous entities of different sectors (Bank, Utility, Telco), but each attribute is weighted by a KYC class applied during the data acquisition (e.g., Class 1 Bank, Class 2 Utility, Class 3 Telco). In conclusion, the main characterizing elements are:

- the control of the consent flow by the user. The consent to share data and documents is controlled by the user as it has to grant the exchange of data between entities;

- the distribution of data and documents. This can be made by the Entities that generate or store data based on an existing and active relationship;
- the traceability of the origin of data and documents. Traceability guarantees the confidentiality and reliability of the source.

**Figure 1: DKYC process**




## 3 A real application of the DKYC: the OKYC PLATFORM

The Onboarding Know Your Customer (OKYC) is a platform based on the DKYC model that simplifies the processes of opening a new relationship to save time and costs thanks to the safe, transparent and unalterable exchange of data between different sources, leaving the user with full control of the flow of data. Indeed, the OKYC platform supports an innovative digital process for the transfer and updating of the available information in line with the needs of users and participating entities. The technological architecture is described in the following report (Cetif, 2021). The platform was initially validated through a Sandbox organized by a committee held by the IT company that developed the model and platform, a consultancy firm, and a multinational IT firm. The Sandbox is a useful environment for analyzing, developing, and experimenting with use cases of various types which, with the appropriate technological adaptations, will be able to see a rapid entry into a subsequent production phase (Brown & Piroska, 2022; Fahy, 2022; Granell et al, 2022). Since their role as regulators of the financial and insurance ecosystem, the first experimentation saw the involvement, as observers, of the national central bank and the national insurance supervisor authority. For 4 months, between 2020 and 2021, different entities participated as users of the platform; the entities were: 5 international banks, 1 multinational Utility in the power sector, 2 IT services companies, 2 international companies providing financing and payment services, 1 regional public administration and a philanthropic organization of the banking sector. The OKYC Sandbox allowed the exchange, within the ecosystem, personal and/or additional data of the customer to contribute to an enrichment of the information base available to the customer. During the Sandbox the actors of the ecosystem have done a total of 1063 operations exchanging 18,579 attributes (e.g., street of residence, identity card number, ...) meaning an average of 10 operations per user. Specifically, these operations covered different aspects, 654 (61%) transfer operations among entities, 187 (18%) updates of data, 222 (21%) automatic propagation of data, that is the automatic diffusion of updated data between the entities of the ecosystem. To evaluate the experimentation phase and collect useful data and information on the onboarding and KYC processes, two questionnaires were prepared, one for those who played the role of customers and one for the experts who represented the companies. The first questionnaire assessed the strengths of the platform, the customer experience, the amount of time saved in onboarding, the intuitiveness of the processes, and the value attributed to the OKYC service. In particular, attention was paid to the evaluation of the user's time savings for onboarding operations and the possibility of using the OKYC service with companies belonging to different industrial sectors. With the OKYC Sandbox, users experienced a significant decrease in time spent on KYC. This decrease in the time spent for the onboarding was quantified as a reduction of 46% in the physical channel and by 64% in the digital channel of time spent for KYC. The second questionnaire was submitted only to the group of experts to collect data on the economic savings generated by the use of OKYC for the company, both in the physical and digital channels, on the performance and effectiveness of the platform. After asking the experts to quantify the cost of current KYC procedures, the experts estimated an overall cost saving of around 48% compared to traditional methods.

## 4 Advantages of the OKYC solution for the onboarding process

The project is one of the first Sandbox based on the DLT technology. The experimentation showed numerous benefits for all the participants of the ecosystem, especially for simplifying the onboarding process, saving time and cost thanks to a secure and unalterable exchange of data, and leaving the customer in full control of it. It also represents a starting point for future developments in the distributed management of digital identity (table 1). From the empirical case, it appears that in the DKYC model, the validation and verification of data and documents are strengthened by the presence of more operators. Furthermore, for a new entity willing to join the ecosystem, it could receive such data and documentation more easily thus allowing saving of time. The type of data as well as of the documents, not being static, may vary according to the types of business relationships, therefore the set of data and documents available will be enriched over time, as the user uses the system to access new products and services also from different sectors. In this sense, another distinctive element also emerges, namely that the existence of different active relationships generates a responsibility that is distributed between the different operators who are custodians of the data and documents of the same subject. Finally, the traceability and classification of individual attributes generated through the mapping of metadata make the origin of data and documents known during data acquisition, respecting the levels of consent required according to the legislation. Traditional KYC systems can reach the first three characteristics of the KYC process (table 1) as their theoretical models are not designed for sharing between the entities' network. Rather these are focused on the digitization of the data acquisition processes. The solutions based on SSI can be seen as complementary to the previous ones as they allow the creation of a network of entities for the exchange of data and documents, thus extending the ability to trace the responsibilities and the origin of the data transparently. The systems based on the DKYC model, in addition to making aspects such as Validation and Traceability more efficient, add the qualifying aspect of the level of compliance of data and documents.

**Table 1. Main peculiarities of the DKYC model**

| Characteristics of KYC | Traditional KYC models | Distributed KYC |
|---|---|---|
| **Validation of data** | The ratio is 1 to 1 between the customer and the entity. | The ratio is N to M as the customer presents data and documents validated by at least one previous KYC. |
| **Verification of data** | Requires an active search by the Entity from third parties to confirm the customer's statement. | The entity might receive data and documents from other entities that have carried out a KYC on the customer. |
| **Dataset** | Fixed and static based on sharable documents (e.g., diploma, driver's license, bank statement, ...). | Variable and dynamic based on past relationships (e.g., dataset necessary for the opening of the different relationships for each product and sector). |
| **Responsibility** | It is mainly based on the customer. | The customer's statement is supported by functional relationships (reinforced onboarding). |
| **Traceability** | For the entity, the origin of data and documents is known after data acquisition (onboarding). | The source of data and documents is made known to the entity during data acquisition (reinforced onboarding). |
| **Compliance** | It is not contemplated a priori, but only *a posteriori.* | It is meta information that accompanies the data and documents acquired during the acquisition phase. |

## 5 Implications and Conclusions

Based on the sandbox results, in this section the authors conclude by summarizing the possible benefits of such a model. These benefits affect the KYC process from an economic, organizational, and regulatory perspective.

5.1.1 **Economic efficiency.** From the experimentation, the results allow indicating a substantially greater efficiency in the KYC processes with the only limitation of the different levels of automation that can affect the individual operator in the processes of verification and validation of data and documents. The introduction of a DKYC model promises a reduction in the use of human, technical, and organizational resources between 40 and 55%.

5.1.2 **Lean verification mechanism and security**. The effect of distributed responsibility on the validation and verification of data and documents also has an impact on the organization of resources. The model allows greater efficiency in the control capacity given by a better focus on the attributes that require a higher level of verification. The attributes that are already validated by other entities require a lean verification process. At the system level, consequently, it is expected to be able to reduce the risks and increase the security by allowing a faster interception of fraudulent behaviors.

**5.1.3 Scalability of the platform.** The data and documents are stored in the databases of the various custodians. Custodians keep data and documents. In this sense, the model is easily scalable because the technological structure is distributed over different actors.

**5.1.4 Privacy.** The platform allows two levels of privacy. The first level is linked to the consent for data processing. The user's consent to the "DKYC" service. It allows the application to map the metadata referring to the identity attributes of the user at the level of the single entities. Furthermore, the consent is specific to each relationship that is opened. The transfer of data and documents from custodian A to another entity is safeguarded from the consent given to custodian A. In this way, the data transfer takes place safely with control by the user because the user always knows which data is being transferred (it has ownership of the transfer). The second level of privacy is linked to the physical storage of data. Data and documents are not stored on the DLT platform. These are off-chain. Thus, they cannot be copied from the platform for malicious purposes.

In conclusion, the authors state that the model here presented might allow for several direct benefits. These are greater transparency, simplified management of customer data, and a simple and fast sharing of data necessary for the opening of a new relationship. Indirect benefits include a general speeding up of the identity verification process and a significant reduction in related costs.

## References


Benchaya Gans, R., Ubacht, J., & Janssen, M. (2022). Governance and societal impact of blockchain-based self-sovereign identities. Policy and Society, 41(3), 402-413.

Brown, E., & Piroska, D. (2022). Governing fintech and fintech as governance: The regulatory sandbox, riskwashing, and disruptive social classification. New Political Economy, 27(1), 19-32.

Cetif (2021). Onboarding Know Your Customer, Outcome Report

Čučko, Š., & Turkanović, M. (2021). Decentralized and Self-Sovereign Identity: Systematic Mapping Study. IEEE Access, 9, 139009-139027.

Eduardo Demarco, A. (2020). Analysing blockchain/distributed ledger technology in capital markets and know your customer process. Journal of Securities Operations & Custody, 12(1), 58-71.

EU (2018) Anti-money laundering (AMLD V) - Directive (EU) 2018/843

Fahy, L. A. (2022). Fostering regulator–innovator collaboration at the frontline: A case study of the UK's regulatory sandbox for fintech. Law & Policy.

Garber, E., Haine, M., Knobloch, V., Liebbrandt, G., Lodderstedt, T., Lycklama, D., & Sakimura N. (2021) Gain digital trust, how financial institutions are taking a leadership role in the digital economy by establishing a global assured identity network, September 2021.

Granell C., Mooney P., Jirka S., Rieke M., Ostermann F., van den Broecke J., Sarretta A., Verhulst S., Dencik L., Oost H., Micheli M., Minghini M., Kotsev A., Schade S. (2022) Emerging approaches for data-driven innovation in Europe: sandbox experiments on the governance of data and technology, 2022, European Commission, https://data.europa.eu/doi/10.2760/511775

NTT (2020) Know Your Customer Strategies for a successful due diligence program during the COVID-19 pandemic

Ostern, N. K., & Riedel, J. (2021). Know-your-customer (KYC) requirements for initial coin offerings. Business & Information Systems Engineering, 63(5), 551-567.

Parra Moyano, J., & Ross, O. (2017). KYC optimization using distributed ledger technology. Business & Information Systems Engineering, 59(6), 411-423.

Qadir, F., Gani, G., & Jeelani, Z. (2022). Blockchain-Based Banking: Theory and Applications. In Applications, Challenges, and Opportunities of Blockchain Technology in Banking and Insurance (pp. 1-18). IGI Global.

Singhal, N., Sharma, M.K., Samant, S.S., Goswami, P., Reddy, Y.A. (2020). Smart KYC Using Blockchain and IPFS. In: Gunjan, V., Senatore, S., Kumar, A., Gao, XZ., Merugu, S. (eds) Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies. Lecture Notes in Electrical Engineering, vol 643. Springer, Singapore.

Stockburger, L., Kokosioulis, G., Mukkamala, A., Mukkamala, R. R., & Avital, M. (2021). Blockchain-enabled decentralized identity management: The case of self-sovereign identity in public transportation. Blockchain: Research and Applications, 2(2), 100014.