

## **WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata 1 December 2020**

### **Aku Turppo: The blockchain technology of EUIPO – and AI projects (transcript translated to English)**

*"...So it's precisely through the API platform that we could start thinking about combining other databases as well as copyrights and industrial property rights..."*

So my name is Aku Turppo. From last year's September until the end of this year's September, I am a Pan-European Seal Trainee. I graduated from the Faculty of Law last year, and the first step after that was that I went to EUIPO. Another thing in the background was that I wrote my master's thesis on the blockchain technology from the perspective of music copyright. Which has led me to be in contact with Anna. In the EUIPO's new strategy plan 2020-2025, one of the main programmes of which is digital evolution, which practically covers all these new technological projects. In my presentation, I will discuss the blockchain technology projects, of which there are two. Then, I will – relatively briefly – discuss AI-related (solutions). I would also like to bring up data quality and things related to integrations precisely from the perspective of metadata, with regard to the Commission's announcements last week.

So if we begin with blockchain register, where this began and what the schedule is. So at the end of last year, within EUIPO, they succeeded in making a proof-of-concept of this blockchain register. Hence its further development got greenlighted. In July, the proof-of-concept was successful in loading and sharing thousands of IP right registries in almost real time, which is a significant step considering there are millions of trademarks and design rights these days on the EUIPO level, too. And with the blockchain technology register, other offices besides EUIPO would be connected with it as well. So the next step is that we begin a cooperation process with the first pilot offices. As far as I understand, there will at first be two national EU offices that will join up. And indeed, the target is that it's ready in 2024 according to the strategy plan. Of course, there are various other steps and analysis points, too, to track how things proceed. The goal is to get 26 IP offices from the EU – EUIPO is one of those 26 – and in addition to those, five international IP offices outside the EU. So I made this little chart which demonstrates in a simple fashion how the blockchain register would work. So here in the middle, we have the "IP register nodes" that are practically the computer servers that run the blockchain database. So, each office involved practically has a node of their own, and each office has their own system linked with those others. And back-office systems are linked with those. For example, one can imagine that if one of these IP offices would be PRH(Finnish IP Office), so when there's an application sent at PRH, it would go through their systems to the PRH node, after which it would be in an agreed process connected with the blockchain, from which – in almost real time – copies would be transmitted to all offices. For example, at EUIPO's end, a user could see from the EUIPO's search system – practically in real time – that there's an application like this written at PRH. So that's practically the big idea and sought advantage here, that all offices' databases would be continuously synchronized with each other. Of course, the goal

## WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata 1 December 2020

here is to relocate EUIPO's main databases – TMView and DesignView – to become a direct part of the blockchain. So that it would give out the real-time data transfer to those operating systems. And one big advantage in this is the strong information security of the blockchain technology – they say it's a reliable system both thanks to its cryptographic protection and its general trustworthiness and transparency. What's more, the transparency and openness are helpful in quality assurance as well, if you consider data – different operators can have easier access to the data. After all, it is possible that each office can personalise their own user interfaces according to what local users at their end would rather use, so it is not a prerequisite that everyone would use the same mould when it comes to operating systems. As to the openness of blockchain, the purpose of the register is to replace EUIPO's current "Open Platform", from which you can now publicly, freely download copyright and design right data.

To go a bit deeper into what kind of blockchain technology we're talking about, at EUIPO, they decided to use a blockchain system called "multichain". They say it has the best sides of both Bitcoin and Ethereum. However, it's a closed system in which they try to weed out monopoly from a single operator with the help of diversity. They say such a diversity consensus fits institutional operations very well. Anyway, Bitcoin is compatible with multichain, so in case one would wish to use Bitcoin in trademark payments, or, there is also the possibility of connecting several blockchain systems with each other. Another blockchain technology may involve more handling of real money, but it's linked with the blockchain register, which could – thanks to the smart agreements – help the processing in the future, as well as payment control. If we think about the possibilities, but of course, it's still too early to go deeper into that. However, what's important in multichain with regard to the register is that it enables one to modify use permissions and consensus as one desires – which, in turn, enables changing functionalities as well as metadata, so if you imagine a right holder who sells all their trademarks, say, as a part of a company acquisition, this data, the data of the new owner has to be modified in the meta-data, so the modifiability of the closed system is vital from the register's point of view. And indeed, when it comes to closed systems, the involved offices would practically make up a governance body, so those are the ones who take part in the functionality and the modifying process, rather than the right holders (of IP).

These days, in EUIPO, they process transfer-of-rights applications and things like that, so this system enables that. What's more, the multichain enables data confidentiality, which is important, because applicants may send documentation related to trade secrets either with their applications or in other processes, so if you send something like that through the register, it would be good to be able to classify such documentation data as confidential – from the otherwise open blockchain database. Multichain also enables the governance of external keys – storing of so-called private keys outside the nodes, which were shown before; the IP register nodes are the core of the functionality. So in a way, these private keys can be governed outside of that. The way I see it, it means that, for example, PRH retains

## **WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata**

### **1 December 2020**

full right to modify their own data. Because of the national trademark, PRH only has the right to modify its metadata – the others in the system cannot access that, precisely because of the governance of private keys. But from this perspective, the system has to be flexible, for if a national trademark, say, is changed into an EU trademark, at that point, the “seniority” would be used, so however, when it comes to EU trademark, the governance right belongs to the EUIPO.

Another blockchain technology that is being advanced is about counterfeit products. So already in 2018, EUIPO arranged a Blockathon forum and competition, based on which they created a Blockchain Use Case documentation on product verification system. So both of those are linked here, in case you want to see in more detail how they have described a potential product verification system to be. In practice, there are four main points of view. Authenticity here means that there is a need to prove and ensure that the products received are authentic. So the vision is that, in this blockchain system, right holders could – so to speak – create individual identifiers for every product, so in blockchain terms, they could token their products. At the stage an individual product moves from one party to another, they change the token between the digital wallets in the blockchain system. Out of the supply chain and the change of token, a delivery database is formed, a log of the entire supply chain, which could – for example – help customs to analyse risks as to whether a product is fake or authentic. In the use case, they have given an example in which a big goods container would also be individualised with such a token. And then the individual products in that container, also tokened, would be combined with the aid of a mathematical algorithm. The idea is that there would be no need to, at least not every time it passes from one party to another, to open the container to check whether the authentic products are still there – rather it would create trust in the fact that the system helps to keep the correct products in and that it has not been changed in any ways. From the execution point of view, another vision is that the system would automatically create transaction warnings, say, if it noticed anomalies in the supply chain, so it would give you a heads-up in case there’s a certain goods container that has to be checked, having been shipped via an anomalous route or something like that. From the origin’s point of view, the idea is that consumers could benefit from a logistical database log, in that they could check themselves where the product has been manufactured, what raw materials have been used, their origin, and they could also check the supply chain.

There is a strong environmental aspect to this: if a consumer wants to be more knowledgeable about the origin country of the products, what the total environmental impact is. Next, the purpose is to begin the pilot implementation at the beginning of next year, which – however – will be more focused on a high-level plan regarding infrastructure. The purpose is also to map out its further development. I tried asking my colleague what kind of technical execution this implementation includes, but they couldn’t say anything about that as of yet; I gather that in the next few days, they will give more information about the pilot implementation. However, the final target and idea is to connect EUIPO IP

## WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata 1 December 2020

Enforcement Portal (that is used for cooperation and data transfers between customs and right holders) also to the possible future blockchain register. So then we would have a three-system interplay to help in this counterfeit-product game. Related to counterfeit products, there are two different pilots in development, one of which would make use of augmented reality; the idea is kind of that the customs authority could make hotspot check-ups in the field of the products to be checked with respect to the original product. The way I see it myself is that the authority would have the products to be checked on the spot, and with the aid of augmented reality, they would get the original product next to it, visualizing the original product, and then perhaps they could do approximate comparison, whether – say – an Adidas or Nike shoe looks similar, the product they have at hand and what the original shoe has been, say, from a design point of view. So the other is a Big Data pilot, the idea of which is to get real-time data for the right holders so they could analyse the state and development of applications for action, which is what customs do when, for example, in the IP Enforcement Portal, they can fill out such an application for action, so to be able to give big data to right holders, so they could assess and analyse effectiveness themselves.

Next, AI solutions. So the first one is **image search**, which is already a feature in use in EUIPO's eSearch Plus and TMView databases. So here you can, say, upload a logo to the image search, and then you should get search results for similar trademarks that are already registered. So this was launched already in 2018, and it is in continuous development. There may be different opinions as to how well it works as of yet – however, it is constantly improved on. Another AI development is chatbot. Was it as early as next year that they're going to launch it? In any case, it is a means of support meant for the application process itself, so it would help with basic questions – how to fill out the application, something related to the code of conduct. In any case, there is a strict limit that it cannot be legal counselling. For example, were a client to ask about the classes they should register their trademark in, it would be legal counselling so that EUIPO customer support couldn't answer them either. Of course, if you contact them by phone, you can get guidance as to where to look and how to think about the classification, but they will not give you an answer how you should do it. In any case, as for the classification, (they're developing) AI-based services, also for registering, which would help the user to classify products and services, to compare between different trademarks, and also to compare between products and services, so the idea is to create added value for the registering process, so it would be easier to predict how well your own registering will succeed – still leaving the responsibility to the user themselves.

I also want to bring up a couple of **language-related AI solutions**. Already these days, in EUIPO's Case Law search in eSearch, one is able to get automatic translations of IP rights, so currently, it supports 30 language combinations, Finnish to English being a relatively recent addition. There is no English-to-Finnish at least yet. The idea is to give IP information in legal cases, say, if someone is unable to understand the processual language of the legal case. For example, someone abroad

## WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata 1 December 2020

doesn't understand Finnish but yet has found an interesting legal case from the databases, so in that case, the automatic translation service would provide them with a relatively good-quality translation based on which they can get at least the gist of the legal case. There are already 150,000 legal cases in there, also from EUIPO Board of appeals and national courts of law. Another tool in development that's perhaps interesting from a lawyer's point of view is a decision-making tool. When making applications, there are many repeated phrase types and expressions, so when the tool starts writing a sentence, it gives you a few options of ready phrases. So the idea is to speed up the decision-making. Then, regarding metadata, a data-quality tool that was implemented in September, first in the Lithuanian office, and later, it is supposed to be taken into use by all national authorities. It's called Data Quality Validator and it's about visualisations – a set of dashboards – which would give you a realistic overview on the data quality of trademark and design right in the said IP office. So it's not about the volume of applications, rather it's purely about – to my understanding – quality of the metadata, whether there are errors or something like that. And the idea is that IP offices would get a proactive role in improving the quality of their own register data, which would naturally reflect on the EUIPO and their database as well. But at this point, I would like to point out a minor criticism that I also saw in my own work when I was doing data analytics, because currently, EUIPO's data warehouse, the update cycle is one day, or, for example, today we get yesterday's data. So it's not possible at this time to get data out from the data warehouse in real time – for now, but they are developing something to change that in taking up cloud services. And I guess another one is that the current data warehouse is not a historical structure. So if you think about it, how it shows in practice, say, in the analysis and in the meta-data is that, if the right holder's representative is changed, it's practically the new representative who replaces the right holder's representative's ID in the data, in the data warehouse. Practically that means that one is forced to make a supposition that the new representative has filed all the right holder's trademarks or design rights. So in that sense, the data warehouse has kind of dragged behind the other AI solutions and stuff, but to my understanding, there will be improvements next year – luckily – from a data economy point of view.

**I would also like to bring up API's** *i.e* programme interfaces mentioned in previous presentations. Currently, EUIPO provides private programmes interfaces precisely to national IP offices, in addition to a few third parties, through which they have a more direct access to these IP data sets, whereas the public channel so far has been the bulk downloads off the Open Platform. However, public programme interfaces are currently in development in order to distribute the data in a more open and extensive way. So the estimated time of the availability for these public programme interfaces is early next year. The API platform, too, will be supported by a developer portal that includes technical data, and the idea is to make use of it to cherish a lively community of developers. So it's precisely through the API platform that we could start thinking about combining other databases as well as copyrights and industrial property rights – for example, at this point. And kind of tangential to

## **WORKSHOP 1: Copyright Infrastructure - Digiskills and metadata**

### **1 December 2020**

this, a new feature has been introduced in the programming interfaces in connection with the EU trademark application, so at the end of the electronic application, one can get a notification as to whether the trademark is a free EU domain in the database governed by EURid. Of course, when it comes to EU domain, it is a more limited domain field than that of coms and orgs. In any case, the idea here is to bring databases together and to combine programme interfaces, so that the IP system as a whole would provide users with more. One also has the opportunity to get notifications in case someone registers an EU domain name that's compatible with their own trademark, which make their own decisions on whether they need to react to it somehow. And EURid and EUIPO here are naturally developing and investigating into whether this could be executed the other way around, so that when you would register an EU domain, you would get some sort of information as to whether it's also free as an EU trademark. Another very recent one from last week is the name registration database of farm products and food, which was published by EUIPO, this GIView, just last week. So of all those and Commission's publications last week, I thought about the combining of trademark and copyright databases. After all, EUIPO is currently responsible for the orphan works database, and also for the development of the Out-of-Commerce portal, which is expected to be deployed next summer, in June. So if you consider the EUIPO Observatory is responsible for IP research, with its focus much on malpractices and execution. For example, I was involved in an investigation on dynamic blocking orders, regarding the internet, which has very much to do with copyright enforcement – precisely in the piracy domain. So all the data that EUIPO has on copyright and other research that's been done. So including those industrial property rights databases. But in addition to that, it would be an interesting idea that there would be a creative author searching a copyright database for copyrights of a record company or a publisher or a band, they could be informed on the side on the trademark possibility – whether it would be a good idea to licence a trademark with the same party, say, for advertising purposes or something like that. So also through the programme interfaces, if one searched for a band, for example, it would be possible to open, for example, the EUIPO database to see whether that band already has a registered trademark. So such integration and combining of programme interfaces is what just crossed my mind, from the perspective of last week and the combining of all the meta-data. That's basically it. Thank you!