

Sosiaalisen median seuranta ja analysointi teollis- ja tekijänoikeuksien loukkaamisen yhteydessä



Tiivistelmä

Verkkokaupan kasvusta on paljon dokumentaatiota, mutta ei ole selvää, miten eri teknologioiden ja kulutustottumusten nousu on vaikuttanut teollis- ja tekijänoikeuksien (IPR) loukkaamiseen internetissä, erityisesti sosiaalisen median alustoilla. Tämän vuoksi Euroopan unionin teollisoikeuksien virasto (EUIPO) päätti toteuttaa tutkimuksen, jolla saataisiin parempi kuva teollis- ja tekijänoikeuksien loukkausten määrästä ja esiintymistäajuudesta sosiaalisessa mediassa (¹). Tutkimuksessa tarkasteltiin kolmea kiinnostuksen kohteena olevaa tekijää paremman ymmärtämyksen saamiseksi nykyisistä toimista ja suuntauksista, jotka liittyvät väärennettyihin fyysisiin tuotteisiin ja digitaalisen sisällön piratismiin.

Tutkimuksen ensimmäisen osan tarkoituksena on antaa kattava kuva sosiaalisen median käytötavoista, jotka liittyvät mahdollisiin teollis- ja tekijänoikeuksien loukkaamistoimiin tai niiden edistämiseen. Toisessa osassa arvioidaan fyysisiä tuotteita ja digitaalista sisältöä koskevien teollis- ja tekijänoikeuksien loukkausten suhteellista osuutta sosiaalisessa mediassa verrattuna aitoihin tuotteisiin tai lailliseen tekijänoikeussuojattuun digitaaliseen sisältöön. Kolmannessa osassa on tarkoitus määrittää keskeiset indikaattorit, jotta teollis- ja tekijänoikeuksien loukkaamiseen liittyvät liiketoimintamallit voidaan tunnistaa paremmin sosiaalisessa mediassa.

Analyysi toteutettiin lounaamalla tietoa neljältä sosiaalisen median alustalta (Facebook, Twitter, Instagram ja Reddit) kuudessa Euroopan valtiossa eli Saksassa, Espanjassa, Ranskassa, Italiassa, Puolassa ja Ruotsissa, vaikka keskustelujen jäljitettävyyttä tiettyyn maahan tutkimukseen valituilla sosiaalisen median kanavilla ei ollut aina mahdollista. Tämän rajoitteen lieventämiseksi käytettiin kielellisiä kriteerejä keskustelujen kohdentamiseksi tutkimuksen kohteena oleviin maihin. Lisäksi englantia lisättiin tutkimuksessa mukana olevien kuuden Euroopan unionin (EU) kielen (ranska, saksa, italia, puola, espanja ja ruotsi) joukkoon, koska muut kuin englantia äidinkielenään puhuvat käyttävät sitä laajalti sosiaalisessa mediassa.

(¹) Kuten kesäkuussa 2020 julkaistussa [EUIPO:n vuoden 2020 tilanneraportissa teollis- ja tekijänoikeuksien loukkauksista](#) ilmeni.

Tutkimuksessa käytettiin sosiaalisen älykkyyden analytiikkaan (social intelligence analytics, SIA) perustuvaa menetelmää sekä määrällistä ja laadullista analyysia. Fyysisten tuotteiden ja digitaalisen sisällön teollis- ja tekijänoikeuksien loukkauksiin liittyvien keskustelujen osalta yksi tutkimuksen rajoitteista oli oikeus päästä tiettyihin keskusteluihin sosiaalisissa verkostoissa. Sen vuoksi tutkimuksessa käsiteltiin vain julkisia keskusteluja. Toinen tutkimuksessa esiin nostettu ongelma koski laillisen ja laittoman sisällön erottamisen vaikeutta. Lisäksi on huomattava, että tässä raportissa esitetyt tulokset edustavat luonnollisesti pientä otosta, ja että analyysit koskevat vain kuutta maata, seitsemää kieltä ja neljää sosiaalisen median kanavaa, jotka olivat mukana tutkimuksessa. Siitä huolimatta tämä lähteiden monipuolisuus ja tietojen laajuus merkitsevät sitä, että tehdyillä päätelmillä voidaan katsoa olevan laajempaa merkitystä. Niin ikään tämän raportin tuloksia sosiaalisessa mediassa tunnistettujen käyttäytymismallien osalta voidaan pitää merkinä teollis- ja tekijänoikeuksien loukkauksiin liittyvästä laajemmasta käyttäytymisestä, joka on tällä hetkellä vallalla sosiaalisessa mediassa.

Sosiaalisen median käyttö teollis- ja tekijänoikeuksia loukkaaviin toimiin tai niiden edistämiseen

Tutkimusta varten poimittiin sosiaalisen älykkyyden analytiikkaa käyttämällä yhteensä 3,9 miljoonaa keskustelua, jotka liittyivät valittuja fyysisiä tuotteita edustaviin ryhmiin ja tuotemerkkeihin. Kyseisten fyysisten tuotteiden ryhmien joukossa suurimmat keskustelumäärät koskivat leluja, hajuvesiä ja kosmetiikkaa. Digitaalisen sisällön osalta suurimmat keskustelumäärät koskivat elokuvia, musiikkia ja videopelejä.

Teollis- ja tekijänoikeuksien loukkaamiseen liittyväksi keskusteluksi määritettiin keskustelut, jotka koskivat teollis- ja tekijänoikeuksien rikkomista. Tutkimuksen mukaan 11 prosenttia fyysisiä tuotteita koskevista keskusteluista saattaa mahdollisesti liittyä väärennöksiin ja 35 prosenttia digitaalista sisältöä koskevista keskusteluista saattaa liittyä piratismiin. Vaikka määrällinen analyysi suoritettiin huolellisesti, tutkimuksessa korostettiin teollis- ja tekijänoikeuksien loukkauksen varman tunnistamisen vaikeutta.

Kun tarkasteltiin sitä, minkä fyysisten tuotteiden yhteydessä esiintyi suurin määrä teollis- ja tekijänoikeuksien loukkaamistapauksia, vaatteet, jalkineet ja korut olivat kolmen kärjessä. Suurin määrä digitaalista sisältöä koskevia loukkauksia havaittiin sähköisten

kirjojen, TV-ohjelmien ja musiikin aloilla. Tämän näytön perusteella voitiin selkeästi päätellä, että kyseisistä tuotteista käytiin suurin määrä keskusteluja, jotka katsottiin mahdollisesti teollis- ja tekijänoikeuksien loukkaamiseen liittyviksi keskusteluiksi.

Väärennettyihin tuotteisiin ja laittomasti kopioituun sisältöön liittyvien keskustelujen suhteellinen osuus sosiaalisessa mediassa verrattuna aitoihin tuotteisiin tai lailliseen digitaaliseen sisältöön

Tutkimuksen toisessa osassa, joka perustuu kolmen eri näkökohdan (alustat, kielet ja aikajana) analyysiin, tunnistettiin tiettyjä suuntauksia teollis- ja tekijänoikeuksien loukkaamistoimissa ja niiden edistämisyrittämissä.

Fyysisiä tuotteita koskevan teollis- ja tekijänoikeuksien loukkaamisen osalta tutkimuksessa selvisi, että sosiaalisista verkostoista Instagramissa käytyjen keskustelujen kokonaismäärä oli suurin. Instagram oli myös tuoteryhmien tasolla pääasiallinen alusta, jossa keskusteltiin kelloista, leluista, hajuvesistä ja kosmetiikasta, koruista ja jalkineista. Havaintojen mukaan Twitterissä keskusteltiin eniten vaatteista ja leluista, ja Redditä käytettiin useimmiten keskusteltaessa lääkkeistä ja päähineistä. Sitä vastoin Facebookissa ⁽²⁾ käytyjen keskustelujen määrä oli pienempi, mikä voisi selittyä kyseisen alustan tehokkaalla tavalla tunnistaa ja poistaa loukkaavaa sisältöä ⁽³⁾. Tätä ei kuitenkaan voida osoittaa todeksi tässä raportissa olevan näytön perusteella, ja Facebookissa käytyjen ja tutkimuksessa huomioon otettujen keskustelujen pieni määrä voikin johtua siitä, että teollis- ja tekijänoikeuksien loukkaamiseen liittyviä keskusteluja käydään pikemminkin yksityisesti kuin julkisesti. Vaikka kaikilla alustoilla on yksityisviestitoiminto, yksityisten ryhmien huomattava asema näyttää olevan Facebookille ominainen ilmiö. Koska tässä tutkimuksessa käsiteltiin vain julkisia keskusteluja, tätä oletusta ei voitu vahvistaa. Se olisi kuitenkin pidettävä mielessä otettaessa huomioon verrattain alhaiset sellaisten keskustelujen määrät, joiden epäiltiin liittyvän teollis- ja tekijänoikeuksien loukkauksiin ja jotka tunnistettiin Facebookissa sosiaalisen älykkyyden analytiikan avulla.

⁽²⁾ Yksityiset Facebook-ryhmät eivät sisälly tähän tutkimukseen.

⁽³⁾ Facebookiin toimenpiteisiin, joilla puututaan teollis- ja tekijänoikeuksien loukkaamiseen, kuuluu ilmianto ja poistoa koskeva maailmanlaajuinen ohjelma, oikeuksien loukkaamiseen toistuvasti syyllistyneitä koskeva tiukka menettelytapa sekä ilmianto ja poistoa tiukempia erityisiä lisätoimenpiteitä. – [Facebookin avoimuusraportti](#).

Digitaalista sisältöä koskevien teollis- ja tekijänoikeuksien loukkausten osalta havaittiin toisenlainen kaava. Redditin havaittiin olevan erityisen tärkeä alusta elokuvaan ja TV-ohjelmiin liittyville keskusteluille, kun taas Twitter oli suosituin musiikkiin ja sähköisiin kirjoihin liittyvissä keskusteluissa.

Kieliparametrin osalta tutkimuksen ensimmäisessä osassa ilmeni, että englanti oli hallitseva kieli ja että joitain kieliä, kuten puolaa ja ruotsia, käytettiin harvoin. Tästä huolimatta havaittiin, että Instagramissa käytiin ranskaksi, saksaksi, italiaksi ja espanjaksi suuri määrä fyysisiä tuotteita koskevia keskusteluja, lukuun ottamatta lääkkeisiin liittyviä mainintoja, joita esiintyi useammin Twitterissä.

Koska tutkimus kesti kuusi kuukautta (vuoden 2020 huhtikuusta syyskuuhun) ja se toteutettiin covid-19-pandemian aikana, sulkutoimien havaittiin vaikuttaneen osittain saatuihin tuloksiin. Tulosten asettamiseksi asiayhteyteen teollis- ja tekijänoikeuksiin liittyvien keskustelujen kokonaismäärää verrattiin kolmen viime vuoden (heinäkuu 2017 – heinäkuu 2020) aikana tallennettujen teollis- ja tekijänoikeuskeskustelujen yleiseen kehitykseen. Tästä kävi ilmi, että fyysisten tuotteiden osalta lääkkeitä koskevien keskustelujen määrä kasvoi huomattavasti ja että vaatteita koskevien keskustelujen määrä nousi huippuunsa sulkutoimien lopulla. Sitä vastoin digitaalista sisältöä koskevien keskustelujen määrä kasvoi asteittain, mikä voidaan tulkita oppimiskäyrän vaikutukseksi, kun sekä laillisen että laittoman digitaalisen sisällön käyttö tuli käyttäjille tutummaksi.

Sosiaalisessa mediassa käytetyt teollis- ja tekijänoikeuksia loukkaavat liiketoimintamallit

Mahdollisten teollis- ja tekijänoikeuksia loukkaavien liiketoimintamallien tunnistamiseksi suoritetuista aihemallinnuksesta ja laadullisesta analyysistä ilmeni, että teollis- ja tekijänoikeuksien loukkaamisesta sosiaalisessa mediassa oli erittäin vaikeaa havaita tiettyä kaavaa. Tämä saattaa selittyä seuraavilla havainnoilla.

- Fyysisten tuotteiden osalta analyysi osoitti, että teollis- ja tekijänoikeuksien loukkaukseen liittyvät keskustelut koskivat pääasiassa myynninedistämistä ja kaupallista toimintaa. Väärennettyjen tuotteiden tarjoajat kopioivat toimivaksi osoittautuneita ja menestyksekkäitä liiketoimintamalleja, jotka olivat jo laillisten tuotemerkkien käytössä.

- Kerätyissä keskusteluissa digitaalisen sisällön piratismi liittyi käyttäjien pyrkimykseen päästä käsiksi laittomasti kopioituun sisältöön. Kaikki keskustelut koskivat nimenomaan sitä, miten sisältöä saa laittomasti käyttöön. Tässä yhteydessä digitaalisten piraattisisältöjen tarjoajat ottivat passiivisen roolin ja luottivat siihen, että käyttäjät levittävät tietoa ja houkuttelevat uusia käyttäjiä tarjoajien sisältöön.

Vaikka räätälöityjä teollis- ja tekijänoikeuksien loukkaamiseen liittyviä liiketoimintamalleja ei voitu todentaa tässä tutkimuksessa, analyysin tuloksena laadittiin määritelmä siitä, miten tietynlaisten vihjeiden joukon avulla voidaan tunnistaa sekä fyysisten tuotteiden että digitaalisen sisällön teollis- ja tekijänoikeuksien loukkaamiseen liittyviä keskusteluja. Luotettavampi lähestymistapa keskustelujen havaitsemiseen voitaisiin saavuttaa käyttämällä sellaisia koulutettuja koneoppimismalleja, jotka voisivat tarjota ennusteita ja suosituksia.

Tutkimus osoittaa, että sosiaalisen median alustat ovat välineitä, joita käytetään toistuvasti sekä fyysisiä tuotteita että digitaalista sisältöä koskevien teollis- ja tekijänoikeuksien loukkaamiseen. Lisäksi tämän tyyppisen sisällön tunnistaminen on monimutkaista johtuen vaihtelevista ja jatkuvasti muuttuvista menettelytavoista, joita väärennettyjä tuotteita tai laittomasti kopioitua sisältöä tarjoavat tahot käyttävät, sekä tuotteiden ja sisällön tarjoamisesta laajalti eri alustoilla, monilla kielillä ja eri sisältötyypeinä.