



**HAL**  
open science

## Advanced Search Strategies in Europeana

Frederik Truyen

► **To cite this version:**

Frederik Truyen. Advanced Search Strategies in Europeana. Archives et Bibliothèques de Belgique - Archief- en Bibliotheekwezen in België, In press, Trust and Understanding: the value of metadata in a digitally joined-up world, ed. by R. Depoortere, T. Gheldof, D. Styven and J. Van Der Eycken, 106, pp.123-134. hal-02125052

**HAL Id: hal-02125052**

**<https://hal.science/hal-02125052>**

Submitted on 10 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ADVANCED SEARCH STRATEGIES IN EUROPEANA

Frederik TRUYEN

KU Leuven has had a long history with Europeana (now celebrating 10 years), being involved in projects such as Europeana Cloud<sup>1</sup>, and the KU Leuven library involved in Europeana inside<sup>2</sup>. In our research group Literary Theory and Cultural Studies, our research into Photography culture has led to many projects on or with Europeana, such as Europeana Photography<sup>3</sup>, Europeana Space<sup>4</sup>, the Europeana DSI<sup>5</sup> and several CEF projects such as Migration in the arts and the sciences<sup>6</sup>, Kaleidoscope: Europe in the Fifties<sup>7</sup> etc.

## The photography collection

After the project on Europeana Photography<sup>8</sup>, we co-founded Photoconsortium<sup>9</sup>, a non-profit organisation that aims at promoting early photographic heritage<sup>10</sup>. With Photoconsortium, we are the aggregator for the thematic channel on Photography in Europeana. The original collection that we contributed to Europeana with the 14 partners in the Europeana Photography project has about 450.000 images from early photography 1839-1939. Less importance has been given to photography of the Great War, since there was a parallel effort for Europeana 14-18. Together, these collections offer a very representative and comprehensive view on photography's first century.

As Photoconsortium, we broaden the scope of these collections and manage the thematic collection of Photography, which currently holds about 2 million images (numbers tend to vary due to the retrieval logic of Europeana. The thematic

---

<sup>1</sup> <https://pro.europeana.eu/post/europeana-cloud-shared-infrastructure-for-european-cultural-cont>.

<sup>2</sup> <https://pro.europeana.eu/project/europeana-inside>.

<sup>3</sup> <http://www.europeana-photography.eu>.

<sup>4</sup> <http://www.europeana-space.eu>.

<sup>5</sup> <https://pro.europeana.eu/project/europeana-dsi> - Photoconsortium started participation from DSI 2 onwards.

<sup>6</sup> <https://pro.europeana.eu/project/migration-in-the-arts-and-science>.

<sup>7</sup> <https://www.photoconsortium.net/50s-in-europe-kaleidoscope>.

<sup>8</sup> F. Truyen and S. Taes, "Trials and Tribulations, Results and Implications. Digitizing for EuropeanaPhotography", in *Image and Narrative*, 15. 4, 2014, p. 1-13.

<sup>9</sup> <http://www.photoconsortium.net>.

<sup>10</sup> F. Truyen and A. Fresa, "PHOTOCONSORTIUM: Digitizing Europe's Photographic Heritage", in M. Ioannides (Ed.), E. Fink (Ed.), A. Moropoulou (Ed.), M. Hagedorn-Saupe (Ed.), et al., *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: Vol. Part 1*. Euromed. Cyprus, October 31 – November 5, 2016, p. 409-419.

channel also holds collections from non-Photoconsortium members with very large contributions amongst others from the Deutsche Digitale Bibliothek<sup>11</sup>.

There were many reasons why partner institutions wanted to contribute to Europeana, as it offers a window to their collections. So visibility to the general audience is certainly one of the foremost considerations. Furthermore, outreach to education plays a role. Many of the institutions we represent mainly work in a B2B context or strictly in the CHI/GLAM sector, but education can play an important role in future audience development. For the thematic collections we performed together with Sinzer an impact study, which showed us the general audience is in fact layered, and an important part of the visitors of our online collections are photo amateurs and semi-professionals. Obtaining visibility outside of B2B and known client base is an extra attraction for Photoconsortium members to be present on Europeana.

When we started the project in 2012, there was a lot of uneasiness about Europeana's policy towards Open Access<sup>12</sup>. This seemed at odds with the business models of both the private as well as the public institutions in the organization, who both tend to have an important income stream out of copyright licensing. Eventually, the upside of Europeana's value proposition became clear: contrary to Google images, Europeana does not cache the images itself but refers to the source images at the provider's side; it also clearly mentions the rights holders and the licensing. This is very important for Europeana itself also since being a "trusted source" is one of its core values.

### **Issues**

Now that we have a sizable collection to manage on Europeana, it is clear that there are still a lot of issues and challenges to be addressed. The major problem for any institution that contributes to Europeana, is the sheer size of the database. With over 50 million objects, the chance that a user hits one of the images of one's collection by just searching in the search box are very low. Also, it would be quite impossible for Europeana to highlight all collections on their home page. Add to this the general findability issues in Europeana when using search box, and we have a somewhat problematic situation.

In fact, in addition to Europeana itself, three parties are affected by this issue: the contributing CHI institutions, the end user and last but not least the professional curator, as we will explain further. Except for the generic problem that is a challenge for the Europeana project and its contributors as a whole, there are also specific problems facing the photography collection. In this paper, we will

---

<sup>11</sup> <https://www.deutsche-digitale-bibliothek.de>.

<sup>12</sup> See F. Truyen and C. Waelde, "Copyright, Cultural Heritage and Photography: A Gordian Knot?", in K. Borowiecki, N. Forbes and A. Fresa (eds.), *Cultural Heritage in a Changing World*, 2016. Cham <[https://link.springer.com/chapter/10.1007/978-3-319-29544-2\\_5](https://link.springer.com/chapter/10.1007/978-3-319-29544-2_5)>.

highlight which strategies are being developed, as well at the Europeana foundation level, the European aggregator level as well as the PhotoConsortium more specifically. Let us first enumerate the specific issues facing the photography collection:

- Metadata uniformity and quality
- Photo Collection-specific issues
  - Inconsistent use of Techniques
  - Most metadata are not specific for photography
  - We do not have adequate periodisation
  - Reuse limitations

#### *Metadata uniformity and quality*

First of all, and common to other collections in Europeana, with the high number of contributing institutions, each having their own databases and descriptive traditions, the uniformity and quality of the metadata is a constant struggle and concern. During the Europeana Photography project, a metadata committee was setup to monitor the ingested metadata<sup>13</sup>, using the MINT tool<sup>14</sup>. A LIDO<sup>15</sup> intermediate layer was used to translate data from the institution-specific metadata to EDM<sup>16</sup>. This highlighted rather big discrepancies between collections: some providers had quite long descriptions with up to thirty keywords per item, others had quite minimal information. A common thesaurus was developed to support the minimal Photography EDM data, with a focus, besides the generic subject categories, on the photographic techniques and formats. The project's control mechanisms didn't prevent however serious remaining discrepancies in how the techniques are named throughout the collection. This had partly to do with the fact that some focuses on the negative, others on positive prints, the technique sometimes refers to the photographic process, sometimes to the print process. Since many partners used their own databases and the institutional procedures, differences in approach proved difficult to control and harmonize. This is actually an ongoing effort.

Another issue is that photographs do not tend to have a "title", but for search purposes this is really essential. In some cases, a shortened description was used as a title for the photograph. It is certainly one of the main lessons learned in the Europeana photography project that metadata quality control is extremely hard

<sup>13</sup> N. Van Steen, "Metadata Management in Europeana Photography", in *SCientific RESearch and Information Technology*, 4. 2, 2014, p. 127-132, <http://dx.doi.org/10.2423/i22394303v4n2p127>.

<sup>14</sup> V. Charles, A. Isaac, V. Tzouvaras and S. Henniecke, "Mapping cross-domain metadata to the Europeana Data Model (EDM)", in *Lecture Notes in Computer Science* (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, 8092, p. 484-485.

<sup>15</sup> <http://network.icom.museum/cidoc/working-groups/lido/what-is-lido>.

<sup>16</sup> <https://pro.europeana.eu/resources/standardization-tools/edm-documentation>.

and requires continued cross-institutional negotiations. In that sense it proved to be vital to have an intermediate layer, such as the MINT environment, where all the incoming collections can be reviewed before being ingested into Europeana. This is one of the reasons Europeana's idea of "Operation direct", where a direct ingestion facility was being developed, was met by a lot of scepticism among Photoconsortium members.

After the end of the Europeana Photography project, Photoconsortium was setup as a non-profit organization to continue the work on these collections in Europeana. Many, but not all, original partners joined Photoconsortium. This allowed the continuous use of MINT to monitor metadata quality<sup>17</sup>. It is through Photoconsortium that most partners continued to contribute to the Europeana DSI.

#### *Photo Collection-specific issues*

We already mentioned the inconsistent use of technique descriptions in the metadata. In fact, in the multi-lingual thesaurus we developed for Europeana Photography a rather complete list of those techniques relevant to "early" photography was elaborated<sup>18</sup>. But many issues remain, and not only because, as indicated, partner institutions often interpreted the application of these metadata descriptions differently. The main issue is that these techniques are not really very well known to the average end user, so they do not really help the search environment forward. A user with somewhat more sophisticated knowledge than "vintage photography" could have heard of "wet plate photography" given that "vintage" is the new hype, but that doesn't mean this user will be using terms as "wet collodion" or "albumen print" in their search queries. A different mechanism needs to be devised to bridge specialist "expert" knowledge with the knowledge of the user.

A more fundamental problem facing the exploration of the photo collections on Europeana is that the gist of the metadata applied consists of more generic terms and is not specific enough for photography. Most descriptions describe what is on the photo rather than describing the photographic properties of the image. They will say that it is the Trevi fountain in Rome photographed in 1906, but not that it is taken with front light and from a low angle, with a wide-angle lens. From a survey that we ran together with Sinzer for the impact task force we studied the audience that is coming to the photo collections on Europeana and discovered that photo amateurs are an important group of people consulting Europeana, certainly now that "vintage" photography has become such a hot topic – with eg in the city

---

<sup>17</sup> E. Tsiflidou and N. Manouselis, "Tools and Techniques for Assessing Metadata Quality", in E. Garoufallou and J. Greenberg (eds.), *Metadata and Semantics Research*. MTSR 2013. Communications in Computer and Information Science, vol 390.

<sup>18</sup> <http://bib.arts.kuleuven.be/photoVocabulary/en/concepts/-photoVocabulary-10000.html>.

of Brussels alone at least five studio photographers offering wet collodion portraits in tintype/ambrotype form. So, it would be useful that the collections can be searched for photographic properties.

A very pressing issue is that we do not have adequate periodisation for the Photography collection. Dates are not always precisely available with older photographs, and when they are available, they are often encoded in quite different ways. This makes it difficult to search e.g. for “photographs of the fifties” or “photography of the interbellum”, “belle époque” or “Victorian age”. Development of a facet (to the left of the search results pane) that would allow selection of at least a timeframe, e.g. 1950-1960” or so, would be very helpful to do curatorial activities on Europeana contents.

Lastly, the photographic collections have severe reusability issues given the current copyright environment. In particular, private as well as public institutions who care for photographic heritage have always image copyrights as part of their revenue model. This posed some serious issues to be able to offer open access to the wider public to highly valuable heritage collections. For the oldest collections, there was the problem that in fact these works come into the public domain – when the known author of the work has died more than 70 years ago. This would mean that if these images are given in high resolution to the public on Europeana, it is no longer possible to charge license rights on those images to commercial partners in a B2B scenario. This was a serious problem for many of the contributing organizations to Europeana Photography – and still is – in particular for the private photo agencies, whose business model is completely based on these kinds of licenses. A risk estimation needed to be made whether losing possible income from a set of images by showing them for free on Europeana would trigger enough new traffic and business for other images. Some of the partners chose to take the risk and had positive experiences, some found that there was insufficient return of investment.

To this came the related problem of Orphan works, which are works of which the creator isn't known and hence it is difficult to assess whether this possible creator has died more than 70 years ago and the work is reasonably deemed to be in the public domain.

A third problem relating to more legal issues is the privacy protection regulation (now GDPR), which makes it more and more difficult to publish more recent heritage material, as people pictured could have their privacy and/or portrait rights violated. This is a more pressing problem in the current projects we are engaged in such as photography of the Fifties in Europe.

## Solutions

Now what are the possible solutions to the issues identified? And how do they fit in the strategy of both Europeana and Photoconsortium, the aggregator for Photography? In fact, the main strategy for Photoconsortium is not technical, but curatorial, content-driven. The main idea is that you cannot simply say to a large audience: here is a bag of 3 million objects, ready for you to explore. One can only access large amounts of data when one is able to navigate its topology, to understand its structure and different layers. As the original collection of about half a million images contributed by the partners in Europeana Photography consisted of hitherto undigitized – and in most cases undescribed – archival contents – it was important to first study the new collection, something there was no time for in the original project. To enable both researchers, amateurs as well as the general public to discover this material it was necessary to do a preliminary curation, discovering patterns, common properties, salient aspects. As much of the new contents stem from Eastern and Central European countries, whose production was not part of the established canon of photography in the West, we could not rely on the standard literature on the history of European photography. We needed to let the collection speak out for themselves. This should only partly an issue for the original collection spanning 1839-1939 of course, but as this canonization in fact took place from the fifties onwards – with e.g. Magnum Photos – anyway the photography of Eastern Europe proved largely underrepresented.

### *Curation: the thematic collections*

To lay the groundwork for future use of the Europeana Photography collections, in research as well as for education and leisure, Photoconsortium developed a thematic collection of photography<sup>19</sup> on Europeana<sup>20</sup>. This collection offers access to about 3.5 million images of photography, including part of the original Europeana Photography collection, that conform to some minimal standards regarding the quality of metadata and preview images. Typically, Europeana Thematic collections consist of Virtual Exhibitions, Galleries, Browse Entry Points, and in this case also of blog contributions like the “Picture in Focus”, see e.g. “Migrants then and Now”<sup>21</sup>. The virtual exhibitions provide larger stories and background, with high quality, large “hero” images and all images in large resolution. A series was started by curator Sofie Taes, “The Pleasure of Plenty” with now already three installments published<sup>22</sup>. The galleries offer a simpler format, being short collections of images on a specific topic, such as “9 million

---

<sup>19</sup> F. Truyen F. and S. Taes, “The Europeana Thematic Collection On Photography - An Open Resource For Education From Photoconsortium”, in L. Gómez Chova, A. López Martínez and I. Candel Torres (Eds.), *EDULEARN17 Proceedings*. Edulearn. Barcelona, 2017, p. 5136-5139. Barcelona: IATED Academy.

<sup>20</sup> <https://www.europeana.eu/portal/nl/collections/photography>.

<sup>21</sup> <http://blog.europeana.eu/2018/04/pictures-in-focus-migrants-then-and-now>.

<sup>22</sup> <https://www.europeana.eu/portal/en/exhibitions/industrial-photography-in-the-machine-age>.

bicycles in Girona!”<sup>23</sup>, often composed in collaboration with experts from the Photoconsortium network. Browse Entry points are complex queries developed by the Europeana team that offer access to a specific subset of the collection, such as “Heliogravures”<sup>24</sup>. High on the wish list of Photoconsortium curators would be the possibility to add annotations to published Europeana records while doing the curation process. When literally thousands of images are reviewed to make a selection for a topic, many of the non-selected images do actually qualify for the search in question. Adding specific keywords as extra annotations to these records would greatly enhance search efficiency on Europeana, both for the curators themselves as well as for the general public.

#### *Metadata strategy*

The metadata strategy of Photoconsortium aligns with the strategy of Europeana developed in the Data Quality Committee and the Aggregator Forum. It adheres to the Europeana Publishing Framework and is in the process, in the DSI-3 and DSI-4 projects, to implement parts of the 2.0 version of this framework.

#### *Europeana overall strategy*

Europeana developed the Europeana Publishing framework in 2015, aiming to increase the quality of metadata records in Europeana<sup>25</sup>. It divides content in four different “tiers” of increasing quality, openness, and reusability. This framework is used to define the quality improvement targets in the Europeana DSI programmes. In the Data Quality Committee<sup>26</sup> continuous efforts are being made to refine and expand the publishing framework, as explained in the Europeana content strategy of 2017<sup>27</sup>. Europeana aggregators sign up to this strategy and commit in the DSI to improve their records according to the framework, for example by introducing a higher percentage of tier 3 and tier 4 objects and eliminating tier 1 as much as possible.

The “Publishing framework 2.0” as presented to the aggregator forum focusses on multilinguality and entity support, by offering cross language recall, entity based facets, browse by date or timespan based facets, browse by subjects and types, browse by Agents, browse by Places. To support this work a Europeana Publishing Guide has been published<sup>28</sup>. Besides the ability to support facets (in

---

<sup>23</sup> <https://www.europeana.eu/portal/en/explore/galleries/9-million-bicycles-in-girona>.

<sup>24</sup> See the links on the main page of: <https://www.europeana.eu/portal/en/collections/photography>.

<sup>25</sup> <https://pro.europeana.eu/post/publishing-framework>.

<sup>26</sup> <https://pro.europeana.eu/project/data-quality-committee>.

<sup>27</sup> <https://pro.europeana.eu/post/europeana-content-strategy1>.

<sup>28</sup> [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/Europeana%20Publishing%20Guide%20v1.6.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana%20Publishing%20Guide%20v1.6.pdf).



essence a column with search refinements to the left of a search results page) Europeana also works on support for “entity pages” which will allow key concepts to be defined at the Europeana results level<sup>29</sup>.

Since about 2010, Europeana engaged in an effort to support Open Linked Data and embedding Europeana object in the RDF framework<sup>30</sup>. The effort by Europeana marks a significant boost to Linked Open Data for the Cultural Heritage sector, and after many years of research we are coming to a point where we can actually connect these data among institutional resources<sup>31</sup>.

Photoconsortium aims to continue the work on the multilinguality for its collections, and wants to build a good basis for the browser facet technology by streamlining its core metadata on photographer names, subjects, photographic techniques and date information.

### *Photography Thesaurus*

During the original Europeana Photography project, a Photography Thesaurus was developed. This thesaurus holds 16 Languages including English (as the reference language), Bulgarian, Catalan, Chinese, Danish, Dutch, French, German, Hebrew, Italian, Lithuanian, Polish, Russian, Slovak, Spanish, Ukrainian and consists of over 500 concepts<sup>32</sup>. It has been published as a SKOS file<sup>33</sup> on a KU Leuven Library server. The thesaurus was the result of an extensive, methodological approach in the original project’s work package two on Themes, Collections and Content and work package 4 on Indexing and Multilingual<sup>34</sup>.

---

<sup>29</sup><https://pro.europeana.eu/post/improving-discoverability-with-new-entities-in-europeana-collections>.

<sup>30</sup> B. Haslhofer and A. Isaac, “Data.europeana.eu: The Europeana Linked Open Data Pilot”, in *International Conference on Dublin Core and Metadata Applications*, [S.l.], September 2011, p. 94-104. ISSN 1939-1366. Available at: <<http://dcpapers.dublincore.org/pubs/article/view/3625>>; A. Isaac and B. Haslhofer. “Europeana Linked Open Data–Data. Europeana. Eu”, in *Semantic Web*, 4. 3, p. 291–297, see also more recently: M. Knepper and V. Charles, “Making Library Linked Data Using the Europeana Data Model”, IFLA 2014 LLD, 2014-08-14 < [http://ifla2014-satdata.bnf.fr/pdf/7\\_iflalld2014\\_ppt\\_knepper.pdf](http://ifla2014-satdata.bnf.fr/pdf/7_iflalld2014_ppt_knepper.pdf)>.

<sup>31</sup> E. Hyvönen, “Publishing and Using Cultural Heritage Linked Data on the Semantic Web”, in *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2. 1, 2012, p. 1–159, see also M. Hallo, S. Luján-Mora, A. Maté and J. Trujillo, “Current State of Linked Data in Digital Libraries”, in *Journal of Information Science*, 42. 2, 2016, p. 117-127; L. Walker, “Linked Data for Cultural Heritage”, in E. Jones and M. Seikel for the Association for Library Collections and Technical Services. Chicago: American Library Association, 2016.

<sup>32</sup> <http://bib.arts.kuleuven.be/photoVocabulary>.

<sup>33</sup> A. Isaac and E. Summers, “SKOS Simple Knowledge Organization System”, Primer, World Wide Web Consortium (W3C), 2009, p. 44.

<sup>34</sup> <http://www.europeana-photography.eu/index.php?en/92/work-packages>.

### Photographic techniques

Lists the techniques used to create the object. It mainly distinguishes between positive and negative types of photography<sup>35</sup>.

The main sources for this part of the thesaurus are AAT and SEPIADES. The techniques are limited to the techniques available between the first photographs and the second World War (1839-1939).

### Photographic Type/Practice

In this part the consortium has tried to capture the reason for taking a specific photograph. The photographic practice describes in keywords, of which one or more can be selected, the way a photo was conceived and for which goal was intended (eg documentary photography, portraiture, advertising ...)

### Keywords

In the keywords section a large number of subjects have been regrouped for the content partners to describe the subject of their photographs: Lifestyle and leisure, Health, Work, Landscape, Conflicts, war and peace, Science and technology, Arts, culture and entertainment, Politics, Education, Human interest, Religion and belief, Disaster and accident, Society, Sport. They are mainly based on the IPTC metadata standard<sup>36</sup>, the ICA SEPIADES<sup>37</sup> and the partners' vocabularies.

### *Current work In progress*

In the context of the work in the Europeana DSI -3, Photoconsortium and Europeana developed a Data Quality Plan that aims to better integrate the Photoconsortium metadata into the overall architecture of Europeana. At this moment the Photography thesaurus is not dereferenced at the Europeana level, but at the MINT level. This means that in each record all 16 languages for the different keywords and terms are exported and ingested into Europeana. To make this procedure more efficient, it was agreed to replace as much as possible of the metadata terms and keywords with identifiers from broader and more generally accepted linked data standards such as Getty AAT<sup>38</sup>, Getty ULAN<sup>39</sup> and Wikidata<sup>40</sup>. At this moment the list of photographic terms of the thesaurus has been mapped to Getty AAT and compared to what is available in WikiData. At the

<sup>35</sup> See the full list at <http://bib.arts.kuleuven.be/photoVocabulary/en/concepts/-photoVocabulary-10000.html>.

<sup>36</sup> <https://iptc.org/standards/photo-metadata/iptc-standard>.

<sup>37</sup> <https://www.ica.org/en/sepiades-recommendations-cataloguing-photographic-collections>

<sup>38</sup> <http://www.getty.edu/research/tools/vocabularies/aat>.

<sup>39</sup> <http://www.getty.edu/research/tools/vocabularies/ulan>.

<sup>40</sup> [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page).

MINT level, a revised XML generation procedure will make sure we define classes for each term that points to different authority lists.

A Photographer's index was developed internally with Photoconsortium members covering the professional and studio photographers in the Photography collections. This index goes beyond what is already available in e.g. Getty ULAN, as many studio photography archives have been opened up for the first time during the Europeana Photography project, and also some genuinely hitherto unpublished, often private collections such as those from NALIS and ICMSS, as well as discoveries by partners such as United Archives have brought to light many new, undocumented names.

#### *Metadata for photographic attributes*

The problem with the metadata used in the Europeana Photography collection, as well as with the other records in the Thematic Channel for Photography, is that these describe, besides the photographic techniques, what is pictured in the photograph, and do not describe the photographic attributes of these images, let alone the typical photographic style. As a growing number of Europeana users are professionals and photo amateurs who are looking for specific sets of images according to lighting conditions, exposure, angle of view, perspective, colour balance etc., it would be nice to be able to search for these properties also. Adding this kind of metadata would be prohibitively expensive, and there simply is no workflow in most Cultural Heritage Institutions to support such an effort. The only option is automation. However, doing a visual similarity search on-the-fly in real time would not be very realistic, certainly not because Europeana does not host the image objects, only the thumbnails and previews. Currently KU Leuven, imec<sup>41</sup>, ECE NTUA<sup>42</sup> and Photoconsortium are involved in projects – e.g. Kaleidoscope: Europe in the Fifties<sup>43</sup> – and project proposals involving automatic tagging of images based on image recognition and deep learning algorithms. The idea is to use available, matured technology to cluster images according to attributes such as balance, fill the frame, lead room, rule of thirds, motion blur, simple, colour harmony, framed, leading lines, shallow DOF, repetition and pattern symmetry. These attributes span traditional photographic principals of color, lighting, focus, and composition, and provide a natural vocabulary for use in applications, such as auto photo editing and image retrieval. These algorithms would then be used to add extra metadata to the records ingested into Europeana. Still under discussion is whether we would embed those in the regular EDM metadata or use the newly to be introduced annotations engine for Europeana.

---

<sup>41</sup> <https://www.imec-int.com/nl/contacteer-ons/etro-imec-research-group-at-vub>.

<sup>42</sup> <https://www.ece.ntua.gr/en>.

<sup>43</sup> <https://www.photoconsortium.net/a-concept-workshop-for-50s-in-europe-kaleidoscope>.

### **Conclusion**

With a strategy solidly rooted in curation and photographic expertise, Photoconsortium aims to contribute to an enhanced search experience on Europeana by combining both the production of curated stories so that the users of Europeana can become acquainted with the richness of the collection and learn the adequate terminology to discover photographic works, as well as contributing to Europeana's efforts to enhance, in collaboration with the content aggregators, the metadata quality of the records.

For the latter, work has been done on a photo thesaurus, on integration with linked data (Getty and Wikidata), on a comprehensive photographers' index for early photography in Europe, and on developing automated tagging of images with photographic attributes. The end goal is to greatly improve the accessibility, searchability and findability of Europeana records for photography.

