

## **Re-exploring reuse in scientific publications**

Portfolio,  
Telma Peura

### **I Introduction**

The 18th century, the Enlightenment era, is known for the emergence of new scientific scholarship and intellectual renewal (Ford, 2003). Connected to the growing book publishing industry, the role of illustrations has been recognized as an important factor for the emerging scientific fields (Moser, 2014). The pictorial conventions reflect the scientific aspirations of the century. In addition to facilitating the classification of specimens, the illustrations reveal what was considered important for the new kind of “scientific realism” to study natural phenomena. For instance, in the case of botany, the illustrations of the Linnaean classification system largely set the way for what is now considered as typical botanical illustration. In Linnaeus’ work, relevant details, such as the stem, leaves and flowers, are depicted instead of the whole plant (Tobin, 1996).

Although the book industry grew and there was a general rise in the amount of printed publications, the production of botanical illustrations still required considerable work. Moving from woodcuts to metal plate engravings made it possible to draw more detailed pictures, but engraving metal was hard and expensive. Furthermore, both recognized and unrecognized reuse of elements from earlier illustrations was common during the century, sometimes leading to misconceptions.

In this portfolio, I explore the reuse of botanical illustrations. What can the data-driven approach reveal about the emergence of the botanical conventions?

### **II Evaluating the performance of the reuse algorithm**



















To detect image reuse in scientific publications, we developed a classification algorithm using Convolutional Neural Networks (CNN) based on the ResNet architecture<sup>1</sup>.

I ran the code using two different models, Resnet18 and Resnet50, which are presented in more detail in Vesalainen (2023). After a quick look at the results, it was evident that the

---

<sup>1</sup> See [https://huggingface.co/docs/transformers/model\\_doc/resnet#resnet](https://huggingface.co/docs/transformers/model_doc/resnet#resnet)

bigger Resnet50 model performed better. For the similarity detection, the initial similarity score was set to 0.90. Some examples from different score ranges are shown below.

 <p>Original Image 120040010101200_0.png, 1785.0</p>  <p>Similarity value 0.905 129480010201090_0.png, 1779.0</p>  <p>Original Image 068230010008220_0.png, 1773.0</p>  <p>Similarity value 0.901 099330012601230_0.png, 1759.0</p>	<p>0.9-0.91: Some exact matches (left), but mostly images similar features in them (right)</p>
 <p>Original Image 054920030000850_0.png, 1769.0</p>  <p>Similarity value 0.920 110220030200050_0.png, 1720.0</p>  <p>Similarity value 0.917 110220030200060_0.png, 1720.0</p>  <p>Similarity value 0.912 110220030300710_0.png, 1720.0</p>  <p>Original Image 049030010203770_3.png, 1715.0</p>  <p>Similarity value 0.920 064340010201120_12.png, 1764.0</p>  <p>Similarity value 0.917 064340010101570_7.png, 1764.0</p>  <p>Original Image 046120011100310_0.png, 1787.0</p>  <p>Similarity value 0.911 097220011100320_0.png, 1793.0</p>	<p>0.91-0.92 Mostly correct identifications, but sometimes many detailed pictures (as in row 2) were scored similar although the whole images were not so similar for the human eye.</p>
 <p>Original Image 049130040003690_0.png, 1799.0</p>  <p>Similarity value 0.949 060190100003660_0.png, 1796.0</p>  <p>Similarity value 0.945 064750020003930_0.png, 1788.0</p>  <p>Original Image 054890030003930_0.png, 1776.0</p>  <p>Similarity value 0.932 119700020004000_0.png, 1794.0</p>	<p>After 0.93, almost only exact matches were found in the sample</p>

Based on these results, I set the final threshold to .91 for further explorations.

## II Identifying most reused images/works

When grouped per work id, the top 5 of the publications which contents had been reused the most (measured as count of detected reuse images) were the following:

Ecco Full Title (author)	Publication year	Nr of similar pictures
Leonardi Plukenetii, M.D. Opera Omnia Botanica, in Sex Tomos Divisa; Viz. I, II, III. Phytographia, IV. Almagestum Botanicum, V. Almagesti Botanici Mantissa, VI. Amaltheum Botanicum. (Leonard Plukenet)	1720	136
Leonardi Plukenetii Amaltheum botanicum (i.e.) stirpium indicarum alterum copiae cornu millenas ad minimum & bis centum diversas species novas & indictas nominatim comprehendens, Quarum Sexcenae & insuper, Selectis Iconibus, aeneisque tabulis In gratiam Phytosoporum exquisite & summo artificio Illustrantur. Opus Tempori Sacratum. In Magnis vel Voluisse Sat est. (Leonard Plukenet)	1705	110
Plantarum historiae universalis Oxoniensis seu herbarum distributio nova, per tabulas cognationis & affinitatis ex libro naturae observata & detecta. Auctore Roberto Morison Medico & Professore Botanico Regio, nec non Inclutae & Celeberrimae Universitatis Oxoniensis P.B. ejusdemque Hort. Botan. Praefecto primo. (Robert Morison)	1715	39
The botanical Magazine; or, Flower-Garden Displayed: in which the most ornamental foreign plants, cultivated in the Open Ground, the Green-House, and the Stove, will be accurately represented in their natural colours. To which will be added, Their Names, Class, Order, Generic and Specific Characters, according to the celebrated Linnaeus; their Places of Growth, and Times of Flowering: Together with the most approved methods of Culture. A work Intended for the Use of Such Ladies, Gentlemen, and Gardeners, as wish to become scientifically acquainted with the Plants they cultivate.	1787	32
Catalogus classicus & topicus, omnium rerum figuratarum in V decadibus, seu secundo volumine Gazophylacii naturae & artis singulis ad proprias tabulas & numeros relatis. A Jacobo Petiver (James Petiver)	1711	18

However, in terms of content reused in **different** publications, the results were different:

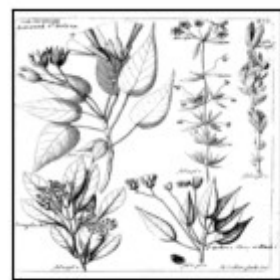
Ecco Full Title	Publication year	Nr of unique work ids with potential reuse
Plantarum historiae universalis Oxoniensis seu herbarum distributio nova, per tabulas cognationis & affinitatis ex libro naturae observata & detecta. Auctore Roberto Morison Medico & Professore Botanico Regio, nec non Inclutae & Celeberrimae Universitatis Oxoniensis P.B. ejusdemque Hort. Botan. Praefecto primo. (Robert Morison)	1715	15
The vegetable system. Or, a series of experiments, and observations tending to explain the internal structure, and the life of plants; their Growth, and Propagation; The number, proportion, and disposition of their constituent parts; with the true course of their Juices; The formation of the embryo, the construction of the Seed, and the encrease from that State to perfection. Including a new anatomy of plants. The whole from nature only. By John Hill, M.D.	1759	11
Leonardi Plukenetii Amaltheum botanicum (i.e.) stirpium indicarum alterum copiae cornu millenas ad minimum & bis centum diversas species novas & indictas nominatim comprehendens, Quarum Sexcenae & insuper, Selectis Iconibus, aeneisque tabulis In gratiam Phytosoporum exquisite & summo artificio Illustrantur. Opus Tempori Sacratum. In Magnis vel Voluisse Sat est.	1705	5
Leonardi Plukenetii, M.D. Opera omnia botanica, in sex tomos divisa; viz. I, II, III. Phytographia, IV. Almagestum botanicum, V. Almagesti botanici mantissa, VI. Amaltheum botanicum.	1720	5
Flora Londinensis: or plates and descriptions of such plants as grow wild in the environs of London: with Their Places of Growth, and Times of Flowering; their several Names according to Linnaeus and other Authors: with A particular Description of each Plant in Latin and English. To which are Added, Their several Uses in Medicine, Agriculture, Rural Oeconomy, and other Arts. By William Curtis, Demonstrator of Botany to the Company of Apothecaries.	1777	4

The difference between the two tables is explained by the fact that the first is affected by the number of pictures in the book in total, whereas the second table only counts the number of publications where any picture from the original work might be found. However, two authors, Leonard Plukenet and Robert Morison stand out from the data in both cases.

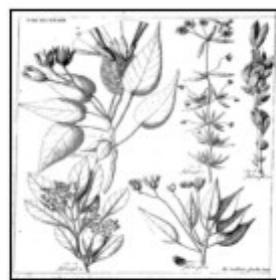
Leonard Plukenet's publications contain a lot of images, and for example his *Amaltheum botanicum* was reprinted in 1769 as well as some other of his work. Consequently, the reuse count of individual images is high. It is still interesting to notice that the version published in 1705 was in folio format, whereas the reprints were done in quarto<sup>2</sup>, although some claim that they were only a reissue of the old plates (Guedes, 1981). The difference in size required engraving new metal plates. Moreover, Plukenet lived in 1642 - 1706, so it is remarkable to notice that his work was regarded as worthy of reprinting later in the new century. Guedes

<sup>2</sup> See <https://www.abaa.org/glossary/entry/book-formats> for different format sizes

(1981) argues that Plukenet's work is particularly important since it served as a reference for Linné, and that Plukenet's influence on the development of botany has been understudied.



Original Image  
103590030002360\_0.png, 1705.0



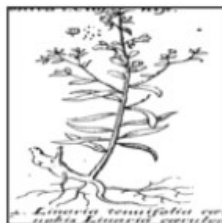
Similarity value 0.935  
110220030503110\_0.png, 1720.0



Similarity value 0.962  
048590050002030\_0.png, 1769.0

*Reprints of Plukenet's illustration.*

As to Robert Morison, the reuse detected by the algorithm is rather stylistic. No exact matches can be found, but the way he depicts plants is similar between the images:



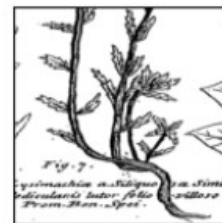
Original Image  
049030010207390\_4.png, 1715.0



Similarity value 0.927  
129480040000720\_2.png, 1800.0



Similarity value 0.912  
064340010200190\_1.png, 1764.0



Similarity value 0.911  
054920020003080\_1.png, 1769.0



Similarity value 0.939  
115830010006090\_4.png, 1741.0

*Morison reused as detected by our model.*

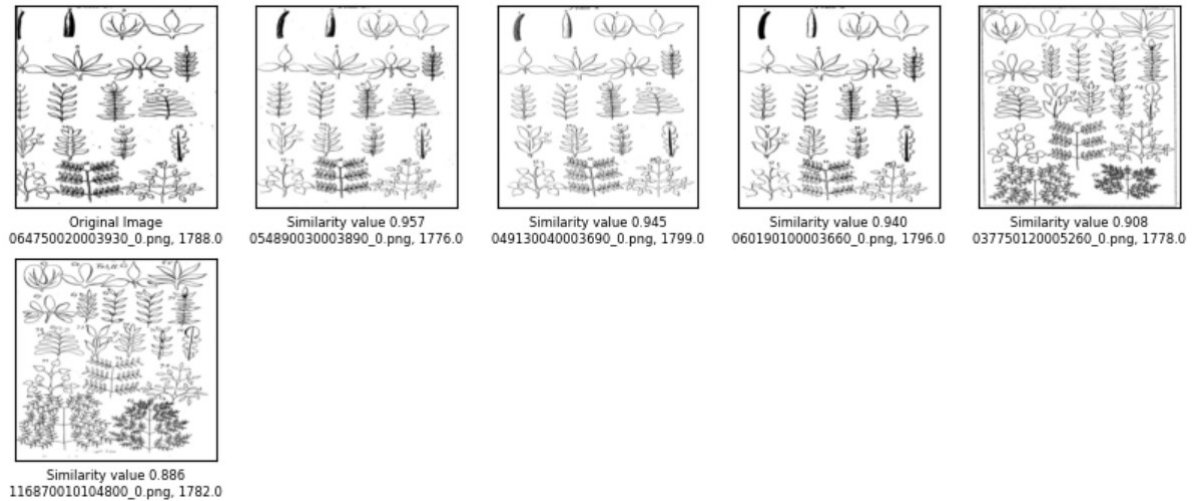
Morison, a professor in Oxford, was a pioneer in the field of botany, and one of the first ones to use metal engravings instead of woodcut in botanical illustrations (Mandelbrote, 2015). However, *Plantarum historiae universalis Oxoniensis* was not as successful as he had hoped. Morison's classification system exhibited inconsistencies, and, against his ambitions, the engravers had used shortcuts and copied material from other earlier illustrations. Perhaps the failures in the printing process discussed in Mandelbrote (2015) partly explain why his illustrations were not reused directly so much.

### III Case Study: Linné's plate of compound leaves

Carl von Linné was an influential botanist in the 18th century, and his classification system served as the basis of the taxonomy we consider typical for botany today. Therefore, I wanted to carry out a short case study on his work. After exploring the data, I selected one plate depicting different compound leaves. The earliest version of the plate is found in the data in *An introduction to botany. Containing an explanation of the theory of that science; extracted from the works of Dr. Linnæus; with twelve copper-plates, two explanatory tables, an appendix, and glossary* by James Lee, published in 1776.

Interestingly, when I looked for similar illustrations for the whole plate or for only one leaf from the same plate, the search yielded slightly different results.

### 1) *The whole plate*



Where the images come from (from left to right):

**1-4:** *An introduction to botany. Containing an explanation of the theory of that science; extracted from the works of Dr. Linnæus; with twelve copper-plates, two explanatory tables, an appendix, and glossary.* (Various editions in 1760, 1776, 1788, 1796, 1799) By James Lee. 8vo.

In the preface, Lee writes that his book offers a complete scientific account of the study of botany in English (Lee, 1788, p. v), that has previously been unavailable in its completeness or too costly to the English reader. Lee goes thoroughly through Linné's classification system, emphasizing the Sexual System of Botany. The illustrative plates are at the end of the book.

**5:** *A botanical dictionary: or elements of systematic and philosophical botany. Containing Descriptions of the Parts of Plants-an Explanation of the scientific Terms used by Morison, Ray, Tournefort, Linnaeus, and other eminent Botanists-a brief Analysis of the principal Systems in Botany-a critical Enquiry into the Merits and Defects of the Linnaean Method of Arrangement, and Distribution of the Genera-Descriptions of the various Tribes, or natural Families of Plants, their Habit and Structure, Virtues, sensible Qualities, and oeconomical Uses-an impartial Examination of the Doctrine of the Sex of Plants-With a Discussion of several curious Questions in the Vegetable Oeconomy, connected with Gardening. The whole forming a complete system of botanical knowledge.* (1778) By Colin Milne, LL.D. 8vo.

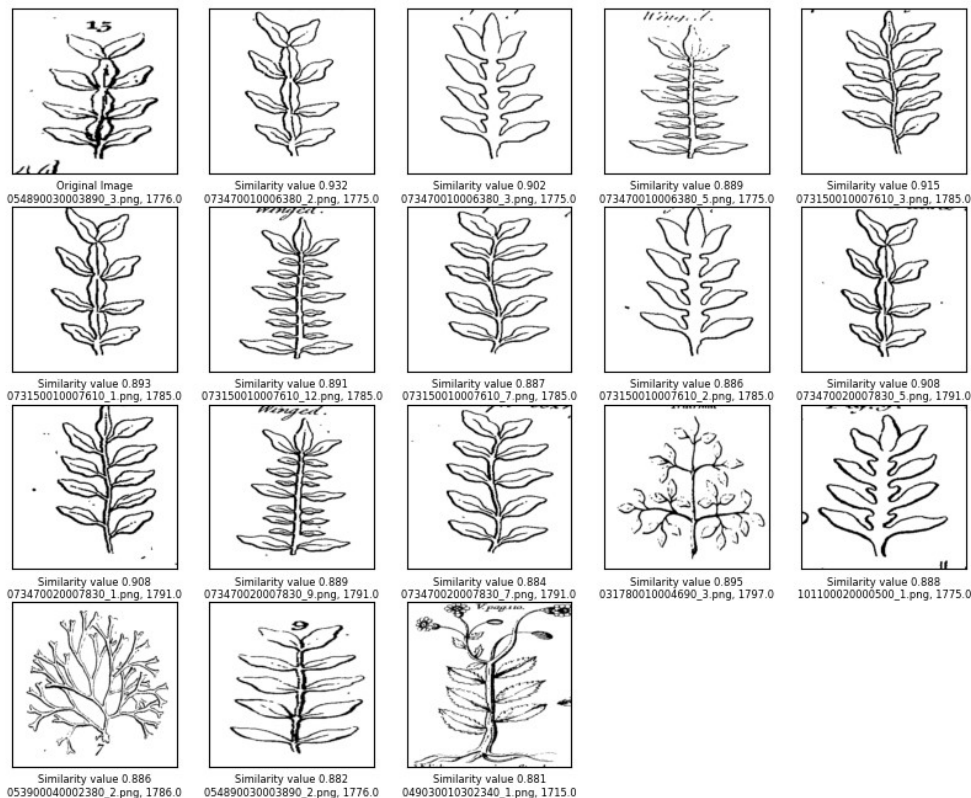
As the name says, this book is a dictionary, where botanical terms are explained in alphabetical order. As differences in the layout of the plate show, this is not an exact copy of

Lee's plate. Whereas in Lee's work, the two first leaves in the upper left corner are the only simple leaves depicted, Milne's book contains a whole plate dedicated to simple leaves, and the plate here only illustrates compound leaves. Also other plates showing Linné's method have similar differences.

6: *A system of vegetables, according to their classes genera orders species with their characters and differences. In two volumes. Translated from the thirteenth edition (as published by Dr. Murray) of the Systema vegetabilium of the late Professor Linneus; and from the Supplementum plantarum of the present Professor Linneus. (1780). By a botanical society, at Lichfield. 8vo.*

Similarly to Lee, the aim of this book is to bring Linné's work to the English audience. The preface refers to Lee and Milne for their contributions in English but the authors point out that this is the first time that Linné's Systema Vegetabilium is fully translated into English. Thus, the reuse of the same illustrations is not surprising. Curiously, when I looked for digitized versions of the original Systema Vegetabilium in Latin, I could not find any edition with illustrations. As this illustration is similar to Milne's dictionary of Linnaean terms, it is likely that they are closer to Linné's original illustrations in their layout.

## 2) A smaller leaf from the same page



Here, I ignored pictures 14 and 15 that were not correctly identified reuse cases.

**1:** *An introduction to botany. Containing an explanation of the theory of that science; extracted from the works of Dr. Linnæus; with twelve copper-plates, two explanatory tables, an appendix, and glossary.* (1776) By James Lee. 8vo.

**2-13:** *A new medical dictionary; or, general repository of physic. Containing an explanation of the terms, and a description of the various particulars relating to anatomy, physiology, physic, surgery, materia medica, chemistry, &c. &c. &c.* (1775, 1785, 1791) by George Motherby. 2fo.

Motherby's dictionary continues the succession of physical and medical dictionaries but is nevertheless ambiguous when it comes to citing his sources (McConchie, 2009). According to McConchie (2009), Linné is mentioned only three times in the dictionary. Botanical knowledge is considered relevant thanks to the "medical virtues" of different plants, which can thanks to chemical and pharmaceutical advancements be applied in the most successful ways (Motherby, 1795, pp. x-xi). As to the accompanying botanical plates, I could not find direct reference to Linné.

**14:** *The gardeners kalendar; directing the necessary works to be done every month in the kitchen, fruit, and pleasure-gardens, As also in the Conservatory and Nursery: Shewing I. The particular Seasons for propagating Esculent Plants and Fruits, with the time when each sort is proper for the table. II. The proper seasons for transplanting all sorts of Trees, Shrubs, and Plants, with the time of their flowering.* (1775) Philip Miller. 12mo.

This publication has a more practical approach, as contains instructions and information related to gardening. The publication size, 12mo, also reflects the nature and target audience of the publication, compared to Motherby's medical dictionary in folio. In the preface of an edition I could find online, the Linnaean influence becomes evident: "In this edition the titles of many plants are altered, so as to correspond with those mentioned in the last edition of the Gardeners Dictionary, which was altered to correspond with Linnaeus's system" (Miller, 1765, xi). A short introduction to the science of botany, following the Linnaean system, is added because "having some skill in the knowledge of plants is now become a necessary qualification for those who have a desire to advance themselves in their profession" (Miller, 1765, xiv). Miller writes that the Linnaean system is the preferable system of those that have yet appeared. This reflects the popularity Linné had gained in England.

## Final thoughts

My goal was 1) to evaluate the accuracy of the reuse results, and 2) to explore and identify some influential or widely reused botanical illustrations from a data-driven approach. As to 1) my experiments with the data showed that the results were not perfect but that the larger Resnet50 model performed better, and a cutoff around 0.90 for the similarity score was enough to detect reuse. The results of the algorithm in my case study about Linné showed that the results were different between a whole illustration plate and an extracted detail from



it. This could be further improved so that details could be matched with full illustrations, and vice versa. It would make the reuse search more flexible to trace small changes over time.

For the second part, I would say that the approach was rather data-inspired than data-driven in the end. I based my experiment on Linné on previous research in the history of botany. However, the case study was encouraging in its results, so that the same method could be tried out with less studied authors. It was interesting to notice that similar illustrations were used from folio to twelvemo, indicating the importance and authority Linné had. Compared to the illustration and their reuse of Plukenet's and Morison's work, Linné was clearly more popular and directly copied. It could be interesting to combine the study of the illustrations with the texts to see whether and how the reuse of illustrations is connected with textual references to Linné or other botanists.

## References

Ford, B. J. (2003). Scientific illustration in the eighteenth century. *The Cambridge history of science*, 4, 561-83.

Guedes, M. (1981). Les oeuvres de Leonard Plukenet et leurs réimpressions. *Archives of Natural History*, 10(1), 67-76.

Lee, J. (1788). An Introduction to Botany: containing an explanation of the theory of that science, and an interpretation of its technical terms. Extracted from the works of Dr. Linnæus, and calculated to assist such as may be desirous of studying that author's method and improvements. With twelve copper plates, and two explanatory tables. To which is added an appendix; containing upwards of two thousand English names of plants, referred to their proper titles in the Linnæan system. (4th Ed.) J.F. and C. Rivington.

Lichfield Botanical society. (1783). A system of vegetables, according to their classes genera orders species with their characters and differences. In two volumes. Translated from the thirteenth edition (as published by Dr. Murray) of the *Systema vegetabilium* of the late Professor Linneus; and from the *Supplementum plantarum* of the present Professor Linneus. [urn:oclc:record:1085601030](https://nla.gov.au/urn:oclc:record:1085601030)

Mandelbrote, S. (2015). The publication and illustration of Robert Morison's *plantarum historiae universalis oxoniensis*. *Huntington Library Quarterly*, 78(2), 349-379.

McConchie, R. W. (2009). Propagating what the Ancients taught and the Moderns improved: The sources of George Motherby's *A New Medical Dictionary; or, a General Repository of Physic* (1775). In *McConchie, RW et al*, 123-133.

Miller, P. (1765). The gardeners kalendar : directing what works are necessary to be performed every month in the kitchen, fruit, and pleasure-gardens, as also in the conservatory and nursery: Shewing I. The particular seasons for propagating all sorts of esculent plants and fruits, with the time when each sort is proper for the table. II. The proper seasons for transplanting all sorts of trees, shrubs, and plants, with the time of their flowering. John Rivington ..., H. Woodfall. <https://doi.org/10.5962/bhl.title.32480>

Milne, C. (1803) A botanical dictionary: or elements of systematic and philosophical botany. (3rd Ed.) Bye & Law for H.D. Symonds. <https://doi.org/10.5962/bhl.title.125381>

Moser, S. (2014). Making expert knowledge through the image: connections between antiquarian and early modern scientific illustration. *Isis*, 105(1), 58-99.

Motherby, G. (1791). A New Medical Dictionary; Or, General Repository of Physic: Containing an Explanation of the Terms, and a Description of the Various Particulars Relating to Anatomy, Physiology, Physic, Surgery, Materia Medica, Chemistry, &c. &c. &c.... J. Johnson. <https://wellcomecollection.org/works/zvkmtp6x>

Vesalainen, A. (2023) Image similarity for the 18th-century illustrations. [https://github.com/vesalaia/Image\\_similarity](https://github.com/vesalaia/Image_similarity)

Tobin, B. F. (1996). Imperial designs: Botanical illustration and the British botanic empire. *Studies in Eighteenth-Century Culture*, 25(1), 265-292.