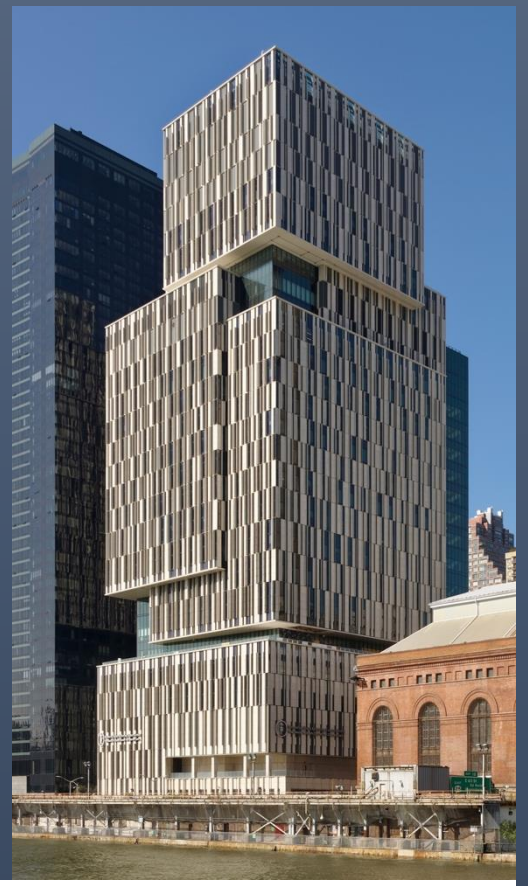
From Principles to Practice: Managing a Medical Research Data Repository

Nicholas Kurtansky

Memorial Sloan Kettering Cancer Center

Self-introduction – Nicholas Kurtansky

- MSKCC Dermatology Service
- Sr. Data Analyst
- Statistics
- 6+ years ISIC Archive admin
- Skin cancer research



Memorial Sloan Kettering Cancer Center

Morphology and risk factors

Distribution

> J Invest Dermatol. 2022 Dec;142(12):3274-3281. doi: 10.1016/j.jid.2022.06.011. Epub 2022 Jul 14.

Analyzing the Spatial Randomness in the Distribution of Acquired Melanocytic Neoplasms

Emmanouil Chousakos [1], Kivanc Kose [2], Nicholas R Kurtansky [2], Stephen W Dusza [2], Allan C Halpern [2], Ashfaq A Marghoob [2]

> J Eur Acad Dermatol Venereol. 2022 Dec;36(12):2379-2387. doi: 10.1111/jdv.18470. Epub 2022 Aug 5.

The long-term evolution of melanocytic nevi among high-risk adults

O Reiter [1 2], N R Kurtansky [1], S T Musthaq [1], S Dusza [1], A C Halpern [1], M A Marchetti [1], A A Marghoob [1], A Scope [1 2 3], V Rotemberg [1]

Dermoscopy

> J Am Acad Dermatol. 2025 Mar;92(3):480-486. doi: 10.1016/j.jaad.2024.10.070. Epub 2024 Nov 9.

Differentiating seborrheic keratosis from melanoma among lesions exhibiting blue-white veil: A retrospective study

Isabella N Dana [1], Nicholas R Kurtansky [2], Larissa M Pastore [3], James R Xu [4], Zaeem H Nazir [5], Stephen W Dusza [2], Emmanouil Chousakos [6], Ofer Reiter [7], Cristian Navarrete-Dechent [8]

> J Eur Acad Dermatol Venereol. 2021 May;35(5):1111-1118. doi: 10.1111/jdv.17133. Epub 2021 Feb 23.

The differences in clinical and dermoscopic features between in situ and invasive nevus-associated melanomas and de novo melanomas

O Reiter [1 2], N Kurtansky [1], J K Nanda [1], K J Busam [3], A Scope [1 2], S Musthaq [1], A A Marghoob [1]

> J Eur Acad Dermatol Venereol. 2021 Apr;35(4):892-899. doi: 10.1111/jdv.17035. Epub 2020 Dec 23.

Association between the dermoscopic morphology of peripheral globules and melanocytic lesion diagnosis

O Reiter [1 2], E Chousakos [1], N Kurtansky [1], J K Nanda [1], S W Dusza [1], M A Marchetti [1], N Jaimes [3], A Moraes [4], A A Marghoob [1]

Epidemiology of melanoma

J Invest Dermatol. 2022 Jul;142(7):1804-1811.e6. doi: 10.1016/j.jid.2021.12.003. Epub 2021 Dec 11.

An Epidemiologic Analysis of Melanoma Overdiagnosis in the United States, 1975–2017

Nicholas R Kurtansky [1], Stephen W Dusza [1], Allan C Halpern [1], Rebecca I Hartman [2], Alan C Geller [3], Ashfaq A Marghoob [1], Veronica M Rotemberg [1], Michael A Marchetti [4]

Br J Dermatol. 2022 Sep;187(3):430-432. doi: 10.1111/bjd.21251. Epub 2022 May 15.

Risk of nonacral cutaneous melanoma after the diagnosis of acral melanoma

Nicholas R Kurtansky [1], Rachel N Manci [1], Danielle M Bello [2], Adewole S Adamson [3], Alexander N Shoushtari [1], Mackenzie R Wehner [4], Charlotte E Ariyan [2], Ashfaq A Marghoob [1],

Primary diagnosis

> NPJ Digit Med. 2023 Jul 12;6(1):127. doi: 10.1038/s41746-023-00872-1.

Prospective validation of dermoscopy-based open-source artificial intelligence for melanoma diagnosis (PROVE-AI study)

Michael A Marchetti [#1], Emily A Cowen [1], Nicholas R Kurtansky [1], Jochen Weber [1], Megan Dauscher [1], Jennifer DeFazio [1], Liang Deng [1], Stephen W Dusza [1], Helen Haliasos [1], Allan C Halpern [1], Sharif Hosein [1], Zaeem H Nazir [1], Ashfaq A Marghoob [1], Elizabeth A Quigley [1], Trina Salvador [1], Veronica M Rotemberg [#2]

> J Am Coll Surg. 2024 Jan 1;238(1):23-31. doi: 10.1097/XCS.0000000000000886. Epub 2023 Oct 23.

Clinical Utility of Melanoma Sentinel Lymph Node Biopsy Nomograms

Harrison M Drebin [1], Sharif Hosein [2], Nicholas R Kurtansky [2], Emily Nadelmann [2], Andrea P Moy [3], Charlotte E Ariyan [1], Danielle M Bello [1], Mary S Brady [1], Daniel G Coit [1],

> Lancet Digit Health. 2022 May;4(5):e330-e339. doi: 10.1016/S2589-7500(22)00021-8.

Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge

Marc Combalia [1], Noel Codella [2], Veronica Rotemberg [3], Cristina Carrera [1], Stephen Dusza [4], David Gutman [5], Brian Helba [6], Harald Kittler [7], Nicholas R Kurtansky [4], Konstantinos Liopyris [8], Michael A Marchetti [4], Sebastian Podlipnik [1], Susana Puig [1], Christoph Rinner [9]

Clinical applications of risk-prediction models

> Ann Surg Oncol. 2025 Mar;32(3):1463-1472. doi: 10.1245/s10434-024-16698-4. Epub 2024 Dec 16.

Declining Clinical Utility of Tools for Predicting Sentinel Lymph Node Biopsy Status: A Single Institution Experience from 2000 to 2021

Harrison M Drebin [1 2], Nicholas R Kurtansky [3], Sharif Hosein [3], Emily Nadelmann [3], Andrea P Moy [4], Charlotte E Ariyan [1], Danielle M Bello [1], Mary S Brady [1], Daniel G Coit [1],

> J Invest Dermatol. 2025 May 29:S0022-202X(25)00539-1. doi: 10.1016/j.jid.2025.05.019. Online ahead of print.

Noninvasive Diagnosis of Melanoma Using Machine Learning and Reflectance Confocal Microscopy

Jonathan Kentley [1], Nicholas Kurtansky [2], Manu Jain [2], Miguel Cordova [2], Jochen Weber [2], Ucalene Harris [2], Anabel Alfonso [2], Allan C Halpern [2], Veronica Rotemberg [2],

> J Surg Oncol. 2023 Jun;127(7):1167-1173. doi: 10.1002/jso.27231. Epub 2023 Mar 11.

Are the MIA and MSKCC nomograms useful in selecting patients with melanoma for sentinel lymph node biopsy?

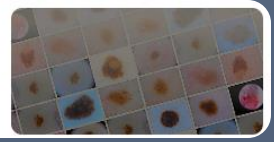
Sharif Hosein [1], Harrison M Drebin [2], Nicholas R Kurtansky [1], Roger Olofsson Bagge [3 4 5],

Sentinel lymph node biopsy

kaggle ISIC Grand Challenges

SIIM-ISIC Melanoma Classification

Identify melanoma in lesion images



2020

> Sci Data. 2021 Jan 28;8(1):34. doi: 10.1038/s41597-021-00815-z.

A patient-centric dataset of images and metadata for identifying melanomas using clinical context

Veronica Rotemberg [#1], Nicholas Kurtansky [#2], Brigid Betz-Stablein [3], Liam Caffery [3], Emmanouil Chousakos [2,4], Noel Codella [5], Marc Combalia [6], Stephen Dusza [2], Pascale Guitera [7], David Gutman [8], Allan Halpern [2], Brian Helba [9], Harald Kittler [10], Kivanc Kose [2], Steve Langer [11], Konstantinos Liopyris [4], Josep Malvehy [6]

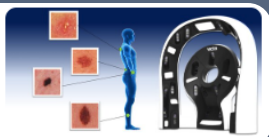
> J Eur Acad Dermatol Venereol. 2025 Aug;39(8):1489-1499. doi: 10.1111/jdv.20479. Epub 2024 Dec 8.

Effect of patient-contextual skin images in human- and artificial intelligence-based diagnosis of melanoma: Results from the 2020 SIIM-ISIC melanoma classification challenge

Nicholas R Kurtansky [1], Clare A Primiero [2], Brigid Betz-Stablein [2], Marc Combalia [3], Pascale Guitera [4], Allan Halpern [1], Jonathan Kentley [1,5], Harald Kittler [6]

ISIC 2024 - Skin Cancer Detection with 3D-TBP

Identify cancers among skin lesions cropped from 3D total body photographs



2024

> Sci Data. 2024 Aug 14;11(1):884. doi: 10.1038/s41597-024-03743-w.

The SLICE-3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection

Nicholas R Kurtansky [1], Brian M D'Alessandro [2], Maura C Gillis [3], Brigid Betz-Stablein [4], Sara E Cerminara [5], Rafael Garcia [6], Marcela Alves Girundi [7], Elisabeth Victoria Goessinger [5], Philippe Gottfrois [5], Pascale Guitera [7,8], Allan C Halpern [3], Valerie Jakrot [7], Harald Kittler [9], Kivanc Kose [3], Konstantinos Liopyris [10], Josep Malvehy [11,12], Victoria Mar [13,14]

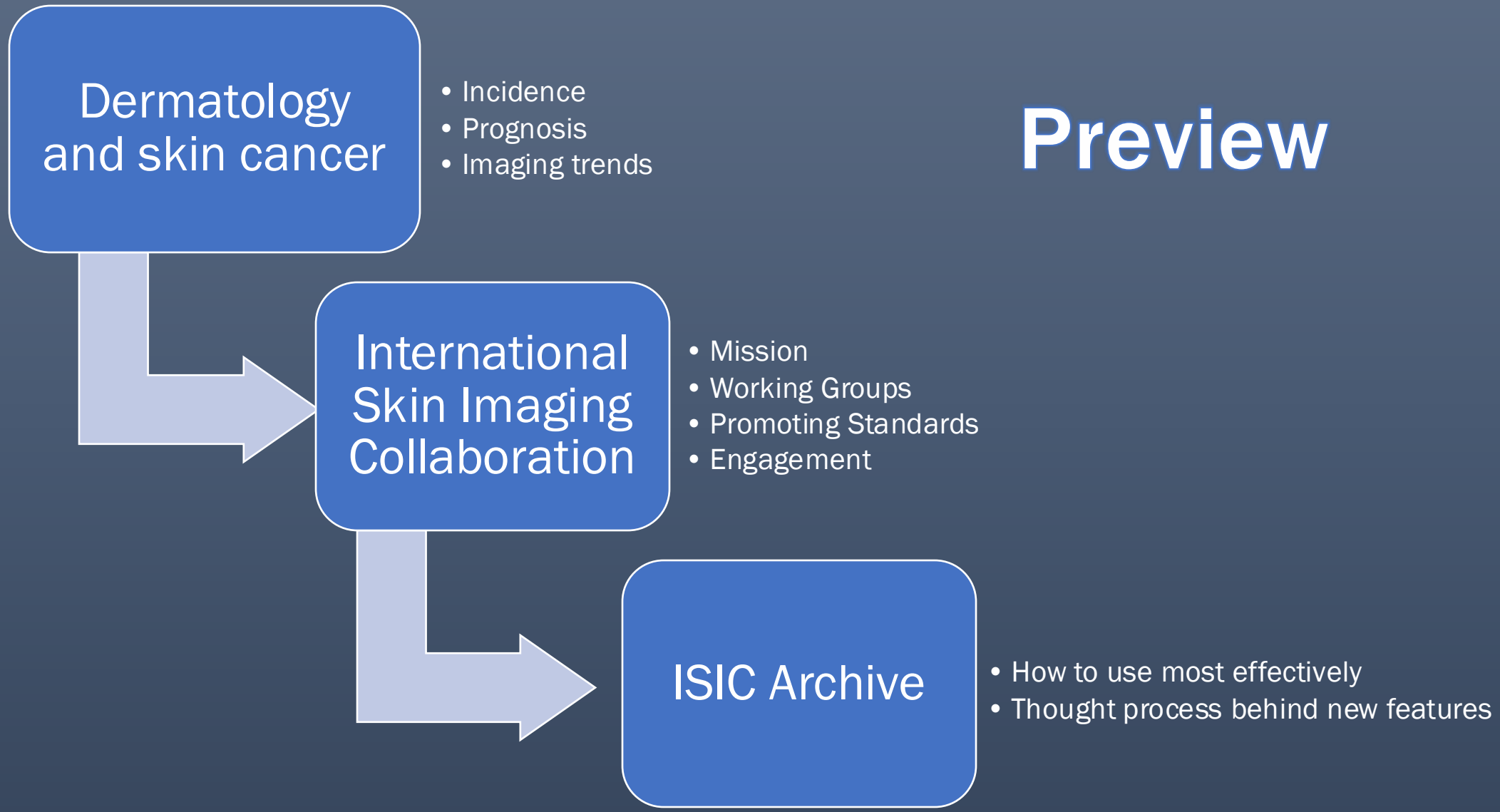
SPRINGER NATURE SNAPP | npj | digital medicine

CURRENT STATUS

Your revised submission is in peer review

Automated triage of cancer-suspicious skin lesions with 3D total-body photography

Preview


Dermatology and skin cancer
• Incidence
• Prognosis
• Imaging trends
International Skin Imaging Collaboration
• Mission
• Working Groups
• Promoting Standards
• Engagement
ISIC Archive
• How to use most effectively
• Thought process behind new features

1-in-5 Americans will develop skin cancer in their lifetime[1]

Type	Survival[1]
Melanoma	Highly treatable if detected early 75% 5yr survival if spread to regional lymph nodes 35% 5yr survival if spread to distant lymph nodes
Basal cell carcinoma	Highly treatable if detected early
Squamous cell carcinoma	Highly treatable if detected early

[1] https://www.aad.org/media/stats-skin-cancer

Digital imaging techniques in dermatology

Jenny L. Stone, MD,[a] Robert L. Peterson, MD,[b] and John E. Wolf, Jr., MD[a]
Houston, Texas

Digital imaging is a versatile technique that has been infrequently used in dermatology to record visual images. We have used this technology for 10 patients to follow cutaneous lesions, including alopecia mucinosa, psoriasis, and dysplastic nevi. The setup included a personal computer, digitizer board, monitor, video camera, and lights. An introduction to electronic (digital) imaging is given and some of the many possible applications in dermatology are discussed. (J Am Acad Dermatol 1990;23:913-7.)

JAAD - 1990

>80% of dermatologists capture digital photos

J Am Acad Dermatol. 2013 Nov;69(5):837-838. doi: 10.1016/j.jaad.2013.07.010.

The use of digital cameras by US dermatologists

Peter Accetta [1], Julia Accetta [2], James Kostecki [3]

Table I. Physician and practice characteristics by use of photography in practice

Characteristic	Photographers	Non-photographers	Total
Gender			
Male	80.1%	19.9%	56.0%
Female	84.9%	15.1%	44.0%
Age (yr)			
< 40	90.1%	9.9%	24.1%
40-49	80.6%	19.4%	26.0%
50-59	80.9%	19.1%	29.2%
60+	76.9%	23.1%	20.7%
Practice setting			
Solo	74.0%	26.0%	34.7%
Dermatology group	85.7%	14.3%	42.7%
Multispecialty group	86.3%	13.7%	13.5%
Academic	93.9%	6.1%	8.8%
NR			0.3%
Total	82.2%	17.8%	100.0%

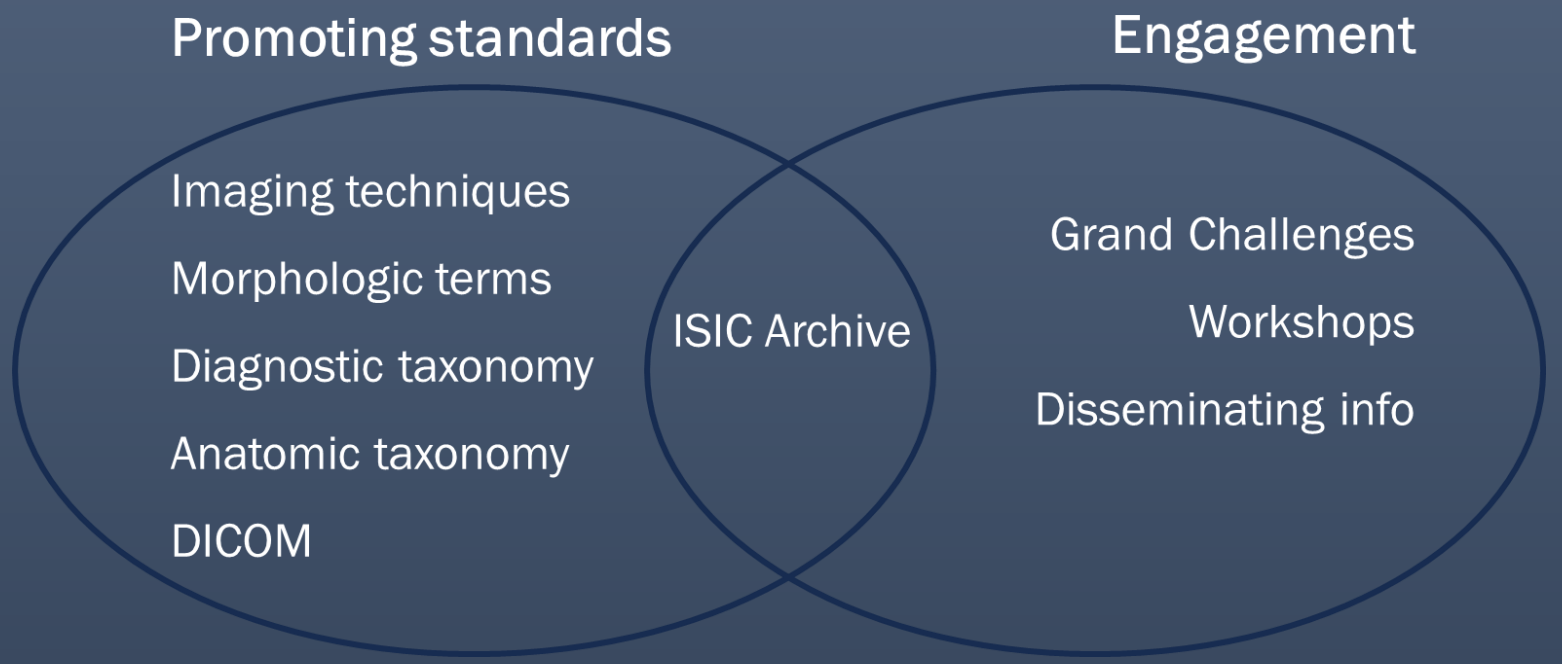
NR, No response.

JAAD - 2013

International Skin Imaging Collaboration

Founded in 2013

Mission: Advance digital skin imaging to reduce skin cancer burden.



Working Groups: Technology – Technique – Terminology – Privacy – DICOM – AI – Education – Community Engagement

What is the ISIC Archive?

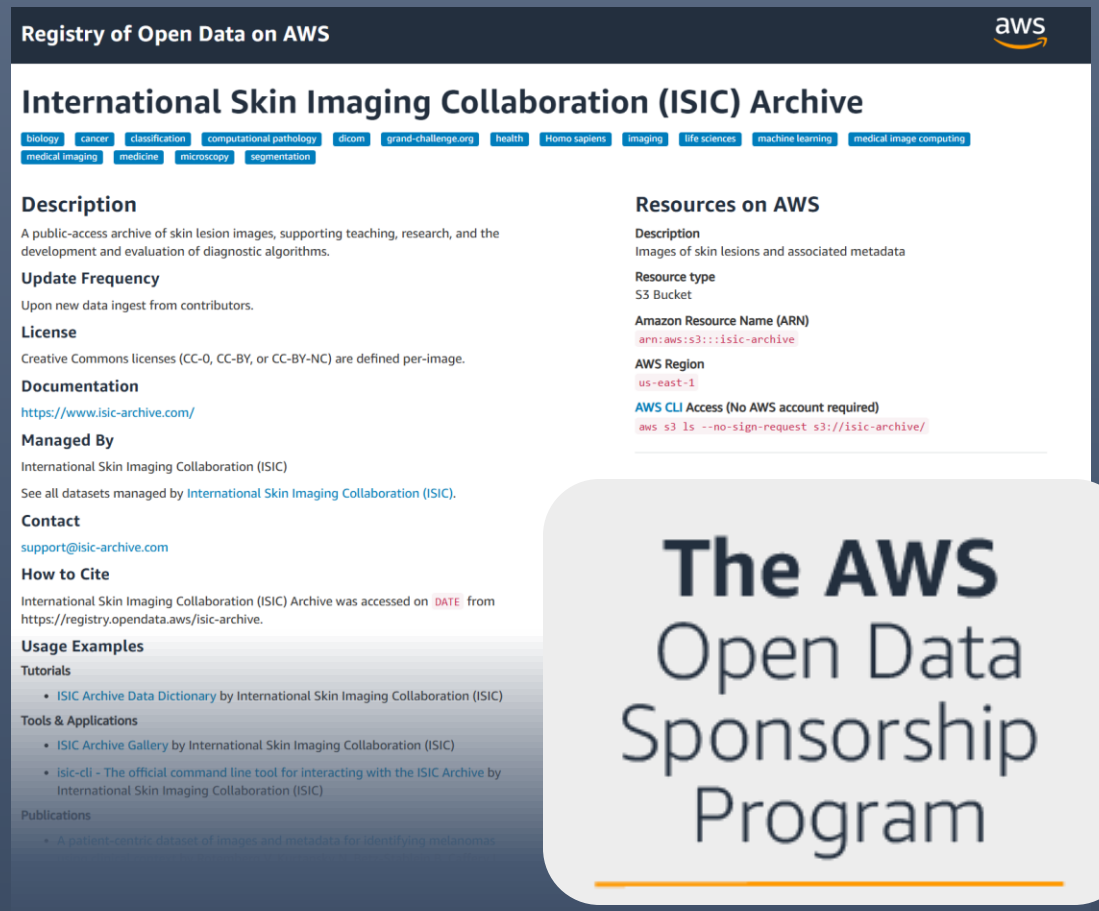
1. Archive of individual skin images

With structured data elements (data dictionary)

2. Repository of skin image datasets

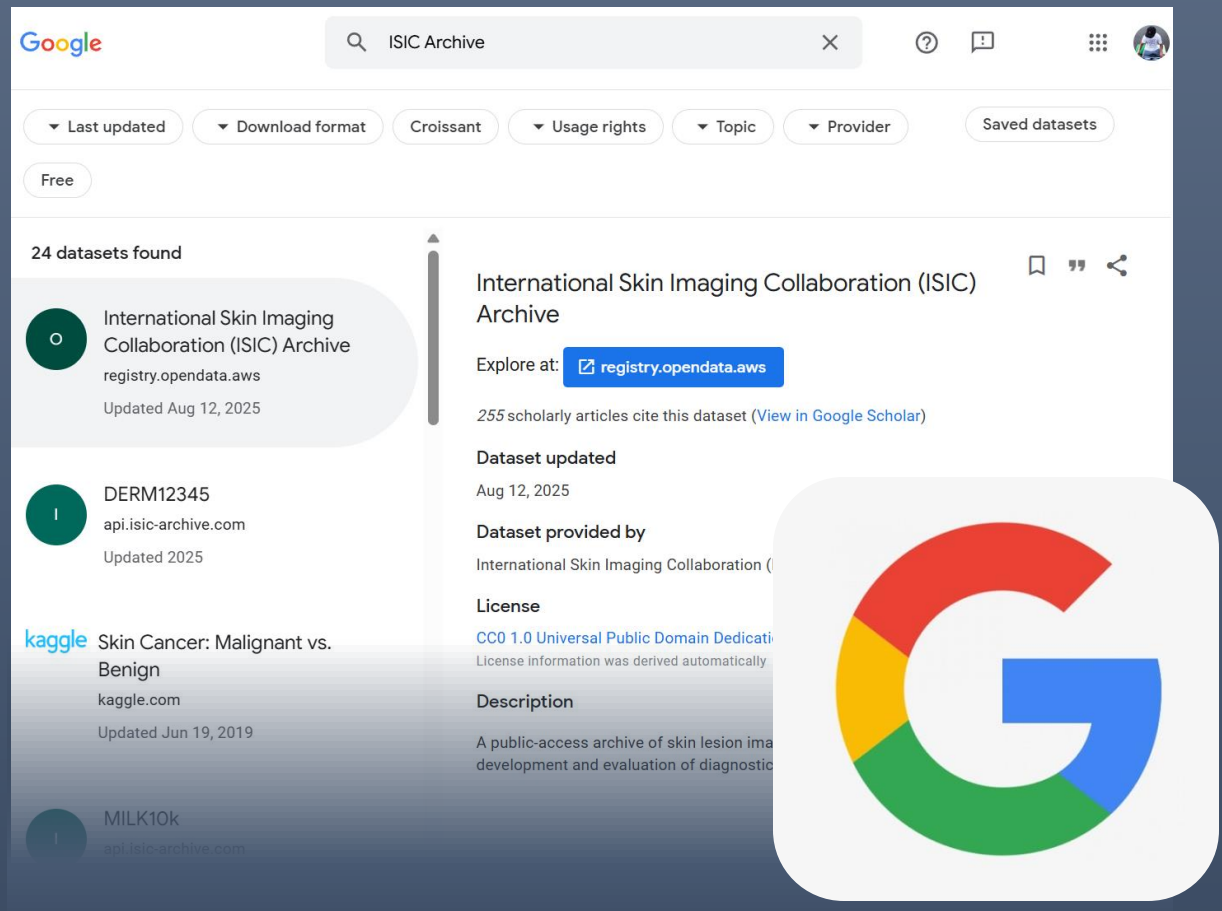
May contain unstandardized/study-specific data

Image total (public)	549,575
Image total (public + embargoed)	1,203,229
Public datasets	53
Datasets DOIs	17
Total users	10,592
Sessions per month	20,000



The AWS Open Data Sponsorship Program

Sustainability reassured by AWS sponsorship



Indexed in Google dataset-search

3 questions
for when considering
NEW FEATURES

1. Why is it important?
2. Does it break anything?
3. Who is it for?

1. Why is it important?

- User feedback & demand
- Working group initiatives
- Scalability
- Research vision

- Funding agent vision

Mission: Advance digital skin imaging to reduce skin cancer burden.

1. Promote standards 2. Engagement

2. Does it break anything?

- *Align with built infrastructure?*
- *Maintainability*
- *Scalability*

3. Who is it for?

Data Users

Model developers, educators, & presenters

- Looking for existing benchmark datasets
- Looking for certain types of cases

Data Publishers

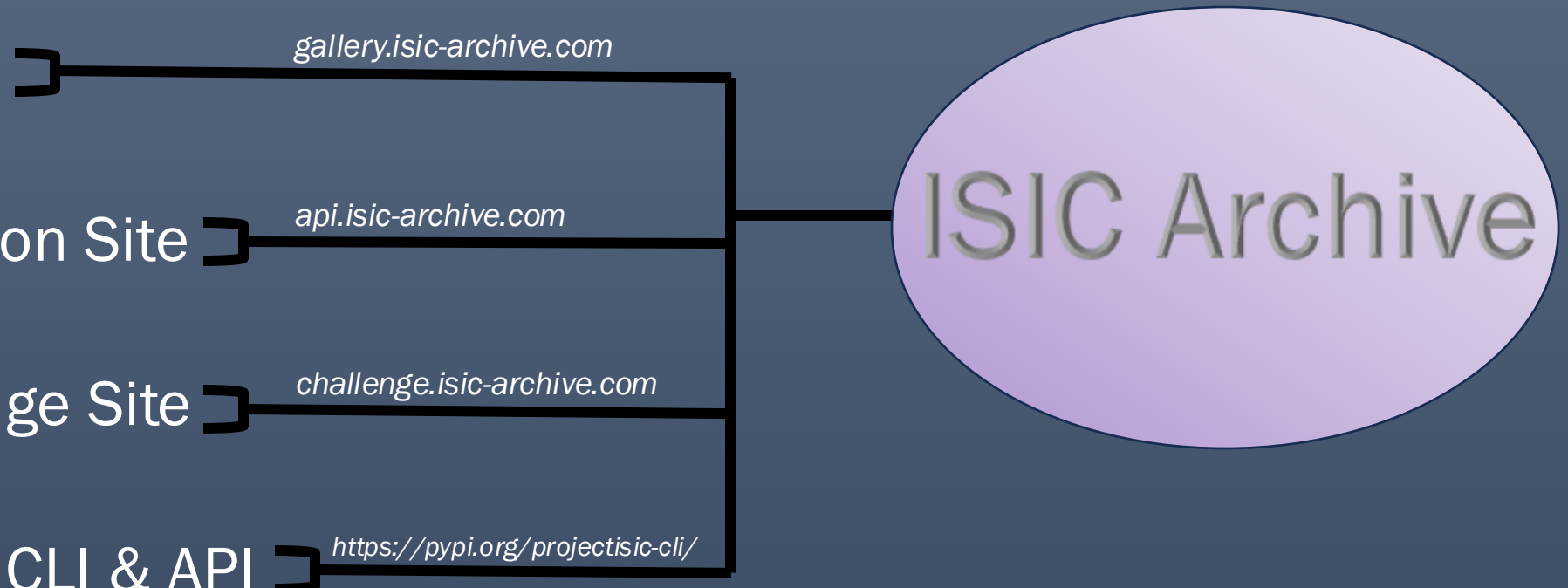
Researchers & image-contributors

- Uploading new images
- Publishing a study
- Using a custom subset of existing data

Examples

Multiple ways to access

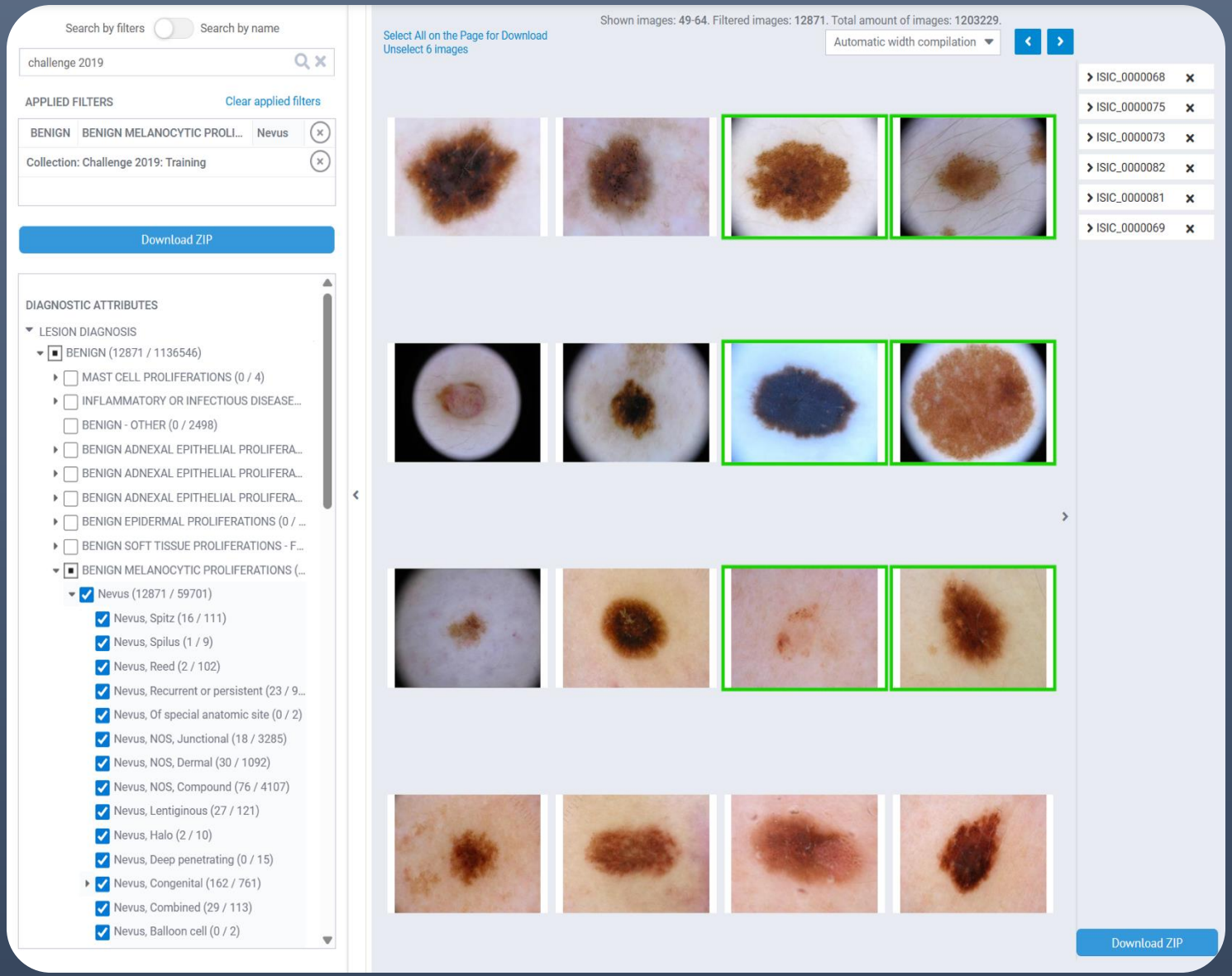
1. Gallery
2. Collection Site
3. Challenge Site
4. Official CLI & API



Multiple ways to access:

1. Gallery

Browsing images



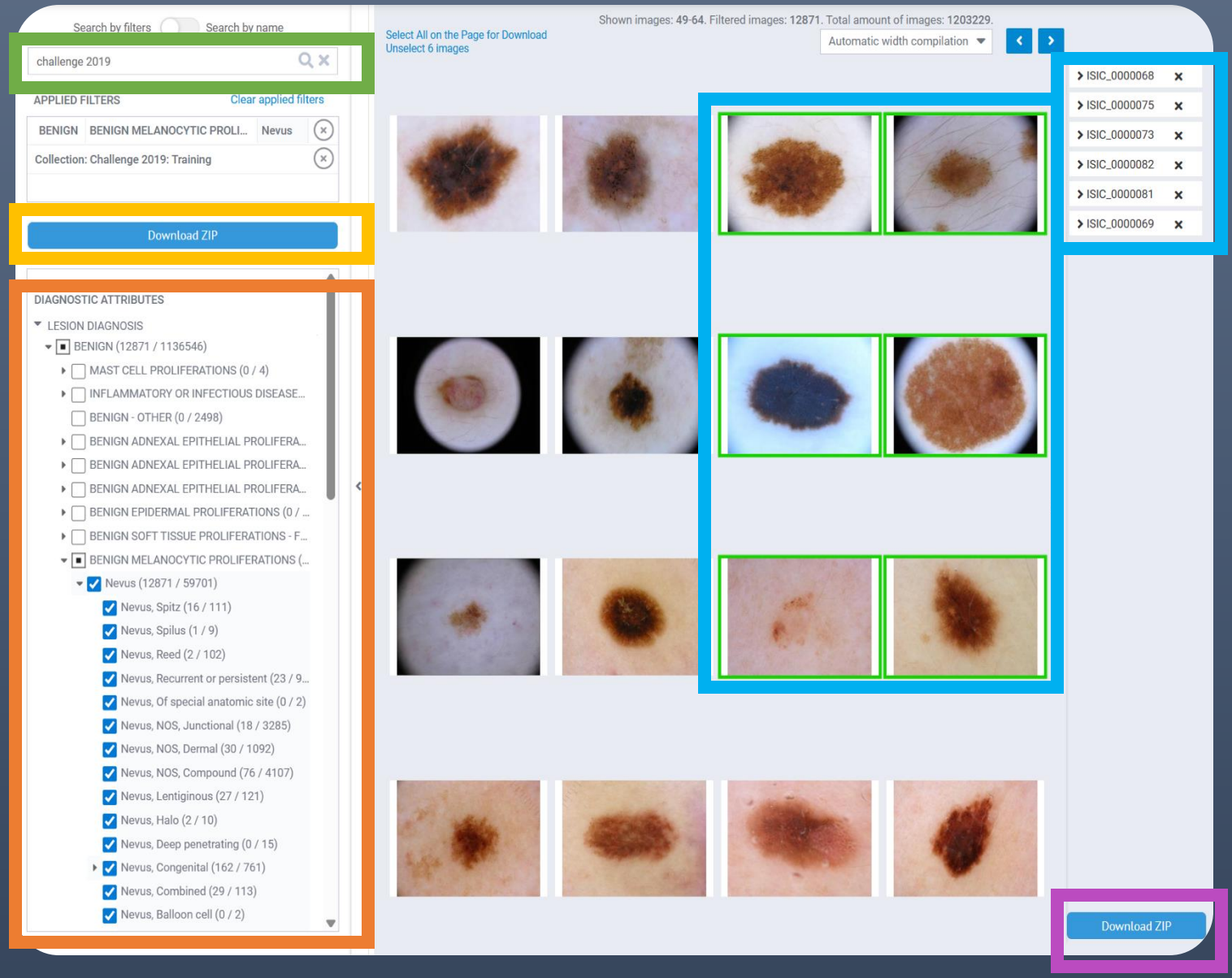
Faceted search

Faceted filter

Download filter

Image selection

Download selection



J Eur Acad Dermatol Venereol. 2025 Jan;39(1):117-125. doi: 10.1111/jdv.20055. Epub 2024 May 11.

International Skin Imaging Collaboration–Designated Diagnoses (ISIC-DX): Consensus terminology for lesion diagnostic labeling

Alon Scope [1 2], Konstantinos Liopyris [2 3], Jochen Weber [2], Raymond L Barnhill [4], Ralph P Braun [5], Clara N Curiel-Lewandrowski [6], David E Elder [7], Gerardo Ferrara [8], Jane M Grant-Kels [9 10], Thiago Jeunon [11], Aimilios Lallas [12], Jennifer Y Lin [13],

1. Why is it important?
2. Does it break anything?
3. Who is it for?

"Nevus" branch of taxonomy tree

- Diagnosis
  - Benign
    - Benign melanocytic proliferations
      - Nevus
        - Blue nevus
          - Blue nevus, Cellular
          - Blue nevus, Common
          - Blue nevus, Epithelioid
          - Blue nevus, Plaque type
          - Blue nevus, Sclerosing
        - Nevus, Acral
        - Nevus, Agminated
        - Nevus, Atypical, Dysplastic, or Clark
          - Nevus, Atypical
          - Nevus, Dysplastic
          - Nevus, Clark
        - Nevus, BAP-1 deficient
        - Nevus, Balloon cell
        - Nevus, Combined
        - Nevus, Congenital
          - Nevus, Congenital, by history
          - Nevus, Congenital, by histopathological pattern
          - Nevus, Congenital, by history and histopatholog
        - Nevus, Deep penetrating
        - Nevus, Halo
        - Nevus, Lentiginous
        - Nevus, Meyerson
        - Nevus, NOS, Compound
        - Nevus, NOS, Dermal
        - Nevus, NOS, Junctional
        - Nevus, Of special anatomic site
        - Nevus, Recurrent or persistent
        - Nevus, Reed
        - Nevus, Spilus
        - Nevus, Spitz
      - Dermal melanocytosis
        - Mongolian spot
        - Nevus of Ito
        - Nevus of Ota
      - Lentiginous melanocytic proliferation
      - Lentigo simplex
      - Pigmented epithelioid melanocytoma
      - Proliferative nodule in congenital melanocytic nevi wit

Filter with new diagnostic taxonomy

DIAGNOSTIC ATTRIBUTES

- LESION DIAGNOSIS
  - ☐ BENIGN (0 / 1136546)
  - ■ INDETERMINATE (387 / 3798)
    - ■ INDETERMINATE MELANOCYTIC PROLIFERATIONS (387 / 631)
      - ☑ Atypical melanocytic neoplasm (387)
      - ☐ Atypical intraepithelial melanocytic proliferation (0 / 235)
      - ☐ Atypical Spitz tumor (0 / 5)
      - ☐ Atypical proliferative nodules in congenital melanocytic nevus (0 / 4)
    - ☐ INDETERMINATE EPIDERMAL PROLIFERATIONS (0 / 3160)
  - ☐ MALIGNANT (0 / 31528)
- TYPE OF DIAGNOSIS (387 / 358234)
- MELANOCYTIC
- MELANOMA MITOTIC INDEX
- MELANOMA THICKNESS (MM)
- MELANOMA ULCERATION

1. Why is it important?
2. Does it break anything?
3. Who is it for?

Incorporating multimodal imaging

Dermoscopy

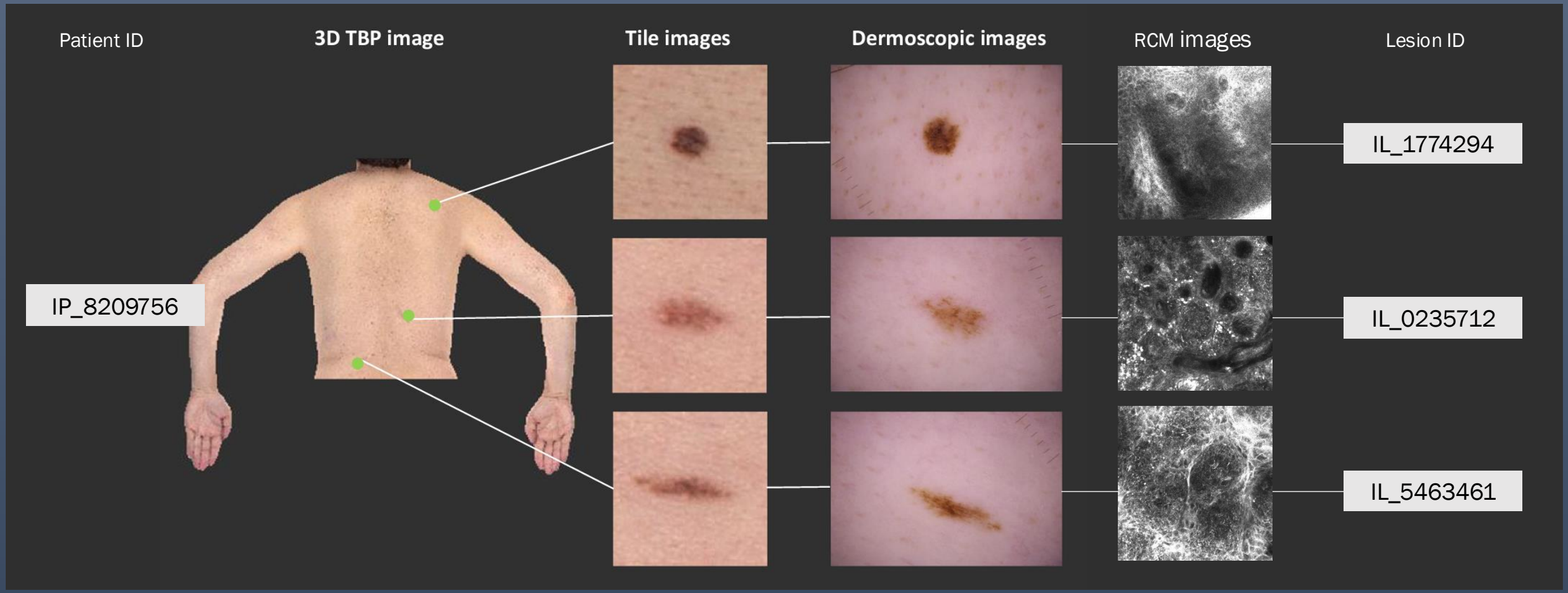
Clinical Close-up & Overview

3D TBP

Reflectance Confocal Microscopy

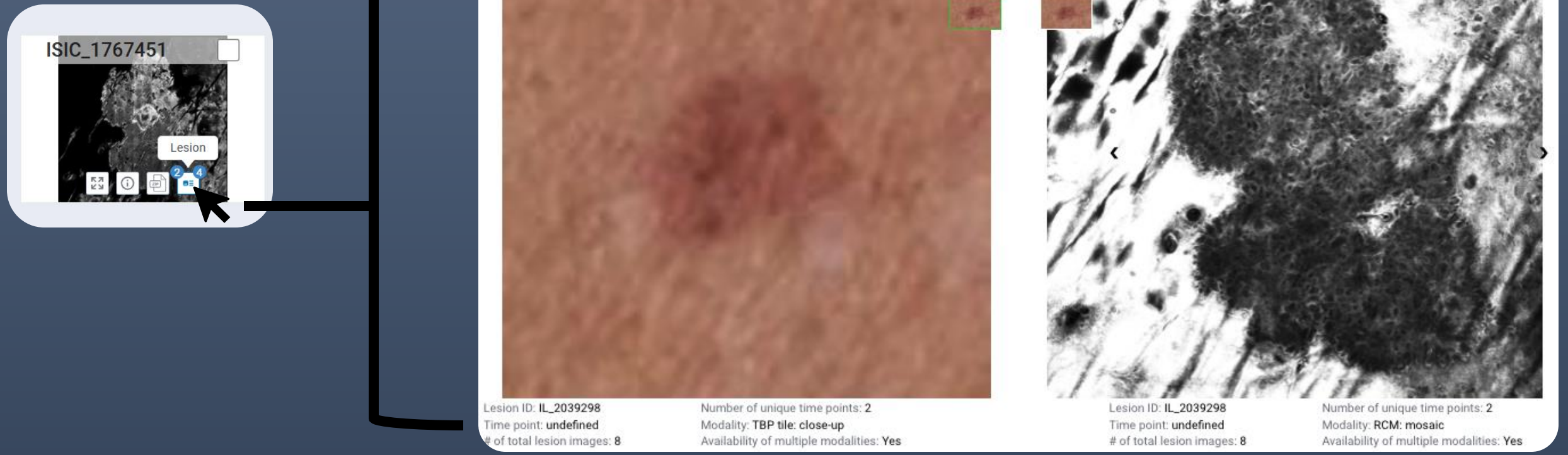
1. Why is it important?
2. Does it break anything?
3. Who is it for?

Supporting image relations



1. Why is it important?
2. Does it break anything?
3. Who is it for?

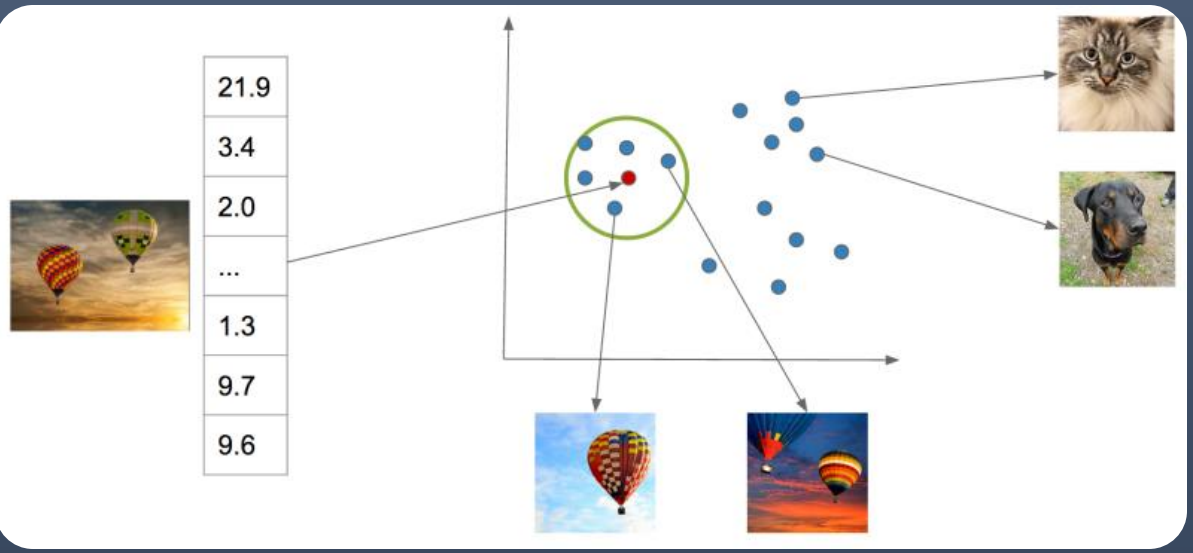
Multi-image case viewer

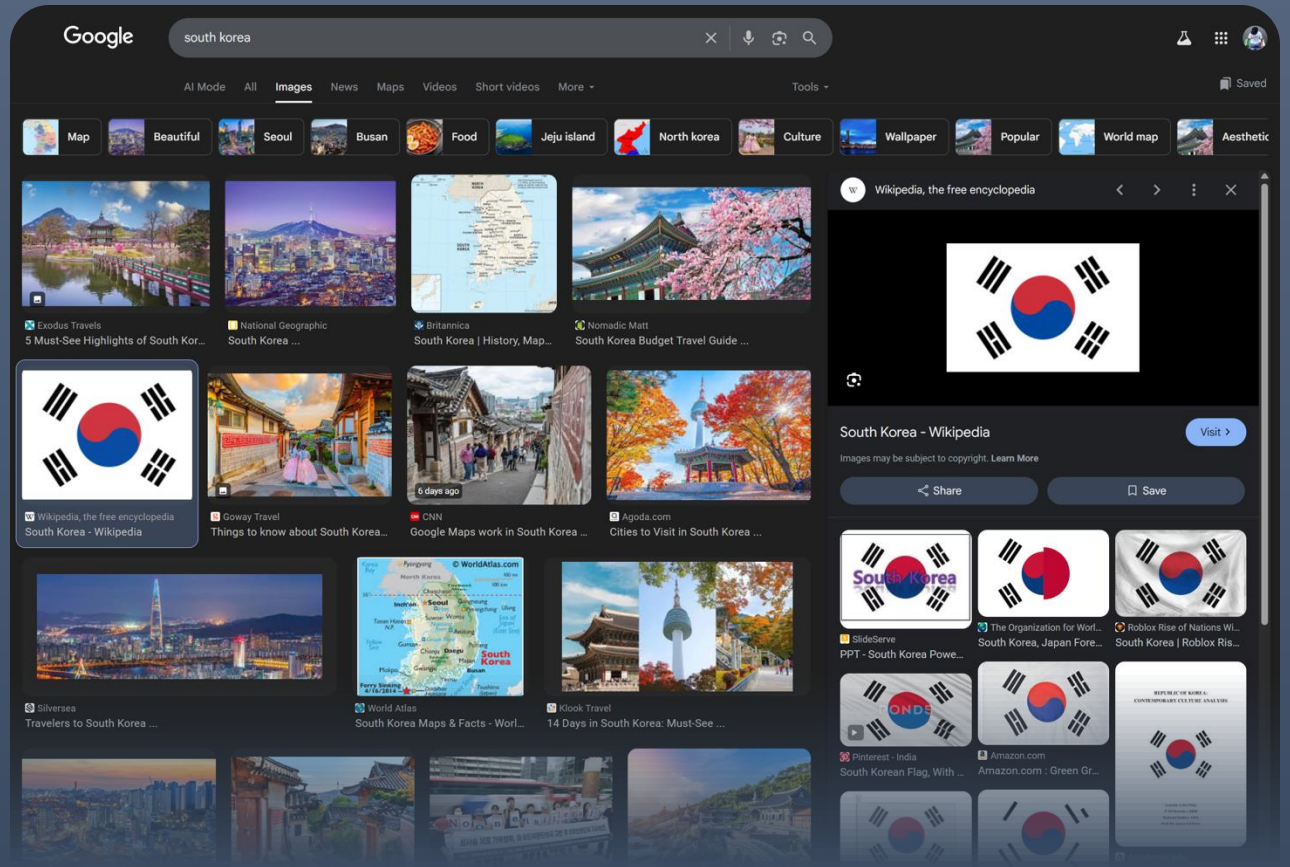


1. Why is it important?
2. Does it break anything?
3. Who is it for?

Visual similarity search

- *Future capability*
- Storing vector representations
- More control over Gallery search



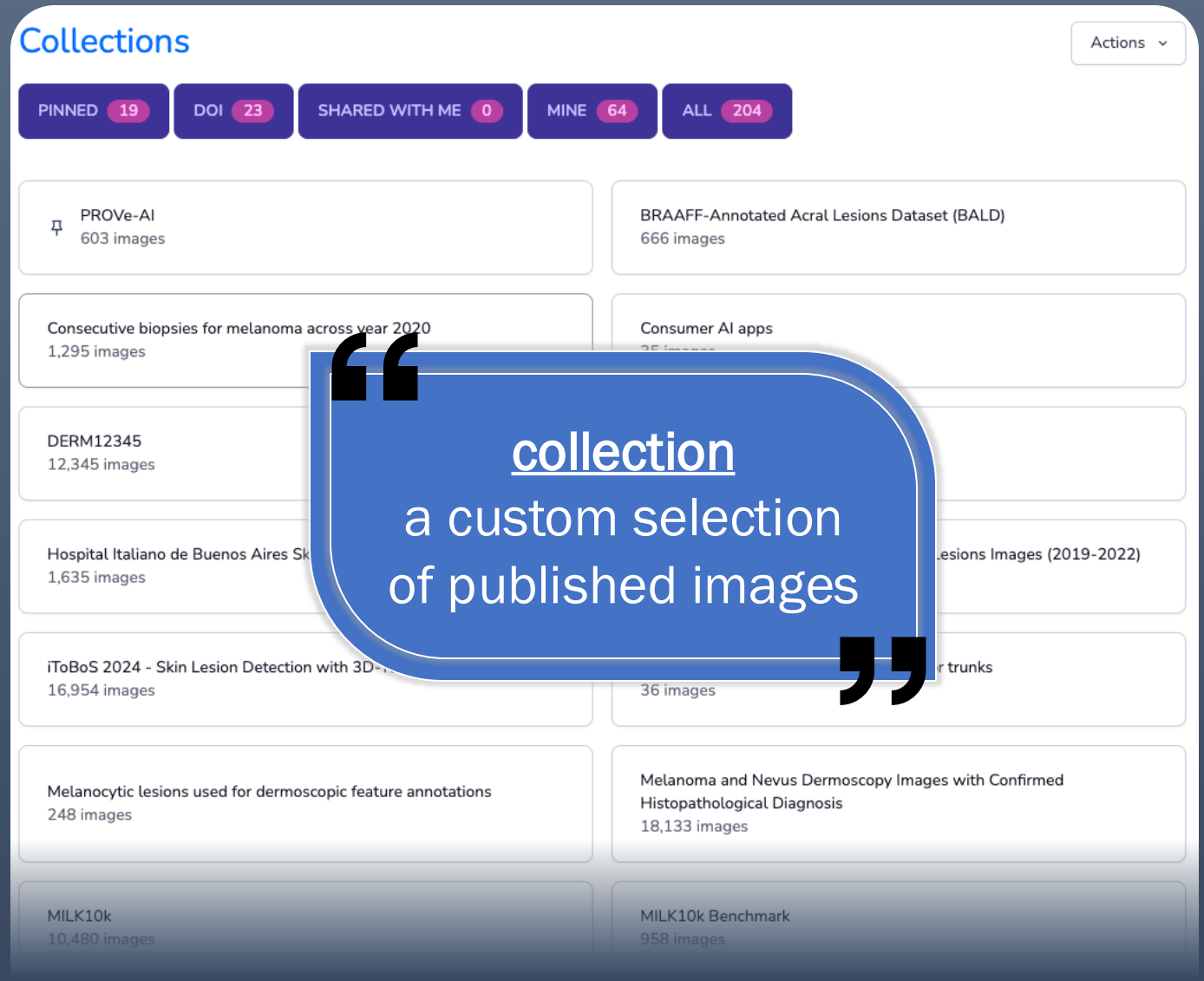


Google Images

Multiple ways to access:

1. Gallery
2. Collection Site

Finding & creating datasets



Curated datasets

- PROVe-AI
  603 images
- BRAAFF-Annotated Acral Lesions Dataset (BALD)
  666 images
- Consecutive biopsies for melanoma across year 2020
  1,295 images
- Consumer AI apps
  35 images
- DERM12345
  12,345 images
- HIBA Skin Lesions
  1,635 images
- Hospital Italiano de Buenos Aires Skin Lesions
  1,635 images
- Hospital Italiano de Buenos Aires - Skin Lesions Images (2019-2022)
  1,616 images
- iToBoS 2024 - Skin Lesion Detection with 3D-TBP
  16,954 images
- Longitudinal overview images of posterior trunks
  36 images
- Melanocytic lesions used for dermoscopic feature annotations
  248 images
- Melanoma and Nevus Dermoscopy Images with Confirmed Histopathological Diagnosis
  18,133 images

From Research Studies

- BCN20000
  18,946 images
- Challenge 2016: Test
  379 images
- Challenge 2016: Training
  900 images
- Challenge 2017: Test
  600 images
- Challenge 2017: Training
  2,000 images
- Challenge 2017: Validation
  150 images
- Challenge 2018: Task 1-2: Test
  1,000 images
- Challenge 2018: Task 1-2: Training
  2,594 images
- Challenge 2018: Task 1-2: Validation
  100 images
- Challenge 2018: Task 3: Test
  1,512 images
- Challenge 2018: Task 3: Training
  10,015 images
- Challenge 2018: Task 3: Validation
  193 images
- Challenge 2019: Test
- Challenge 2019: Training

From Grand Challenges

https://api.isic-archive.com/collections/

Conducting research?

- ❑ Annotating an existing dataset?
- ❑ Focused selection of images?

Consider creating your own collections!

- ✓ Save for future downloads
- ✓ Get a DOI link for your manuscript

Collection DOI Page

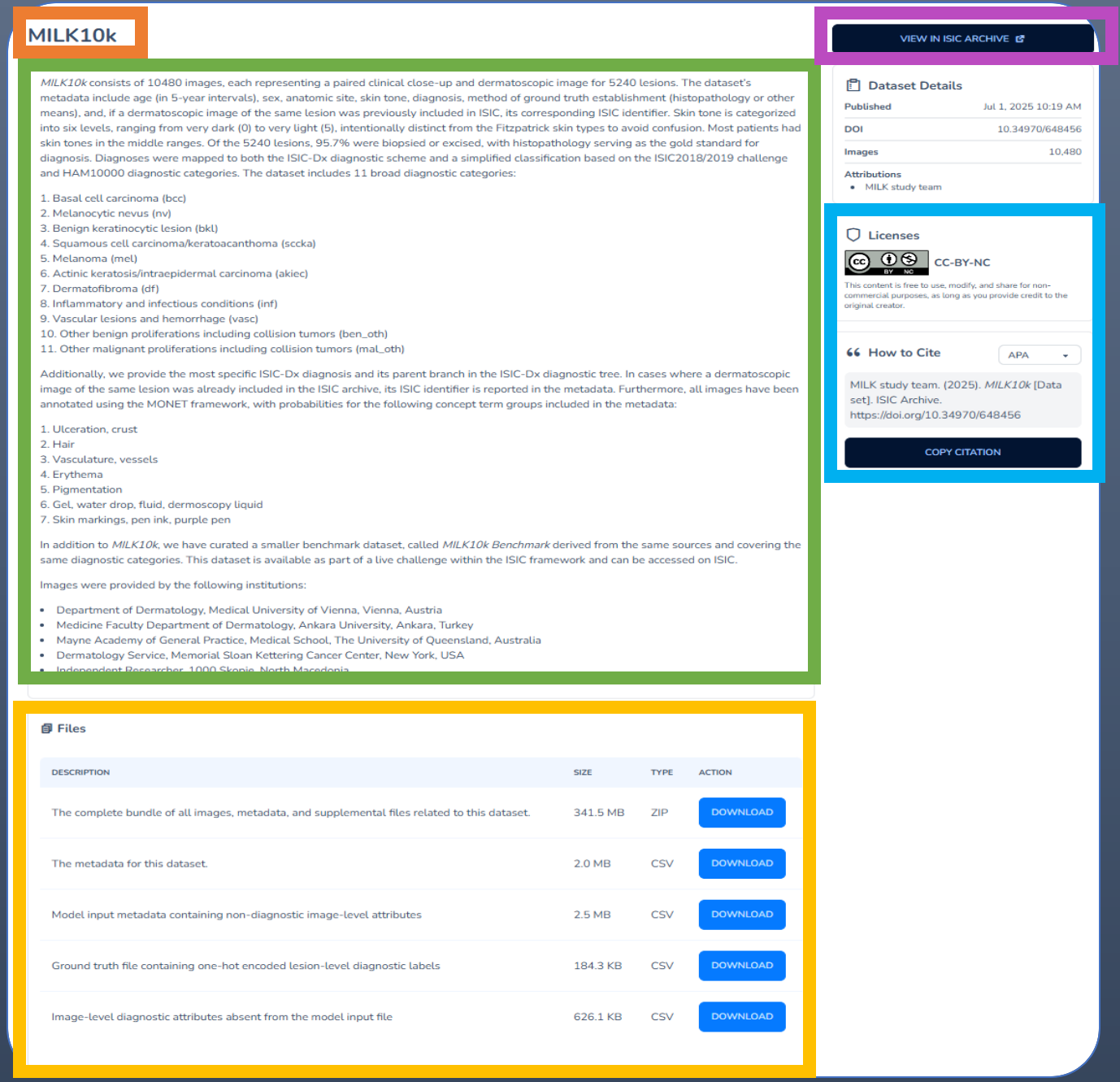
Collection name

Link to view images

Custom description

License and citation

File download links



[1] MILK study team. (2025). *MILK10k* [Data set]. ISIC Archive. https://doi.org/10.34970/648456

Benefits of Collections DOIs

1. Why is it important?
2. Does it break anything?
3. Who is it for?

Discoverability	• Indexed on Google dataset search
Machine readability	• Related identifiers - standard DataCite property • Facilitates links to related study article, descriptor, or GitHub repo
Supplemental data	• Study-specific annotations
Data consistency	• “Frozen ZIPs” ensure data won’t be impacted by future changes
Draft pages	• Adjust the hidden page prior to study publication

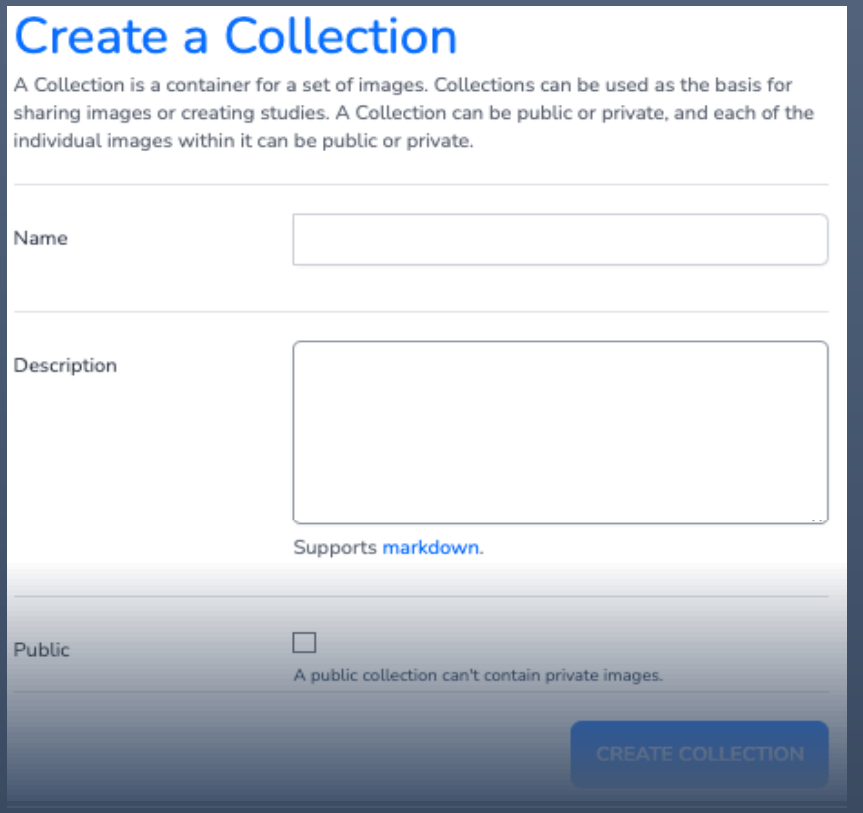
1. Why is it important?
2. Does it break anything?
3. Who is it for?

Create custom collection from existing data:

1. Initialize on the Collection Site
2. Create a line-separate file of select ISIC_IDs
   - ➢ *Utilize image search functionality*
3. Add images to the collection using ISIC-CLI
   - ➢ *Mentioned later*

If minting a DOI:

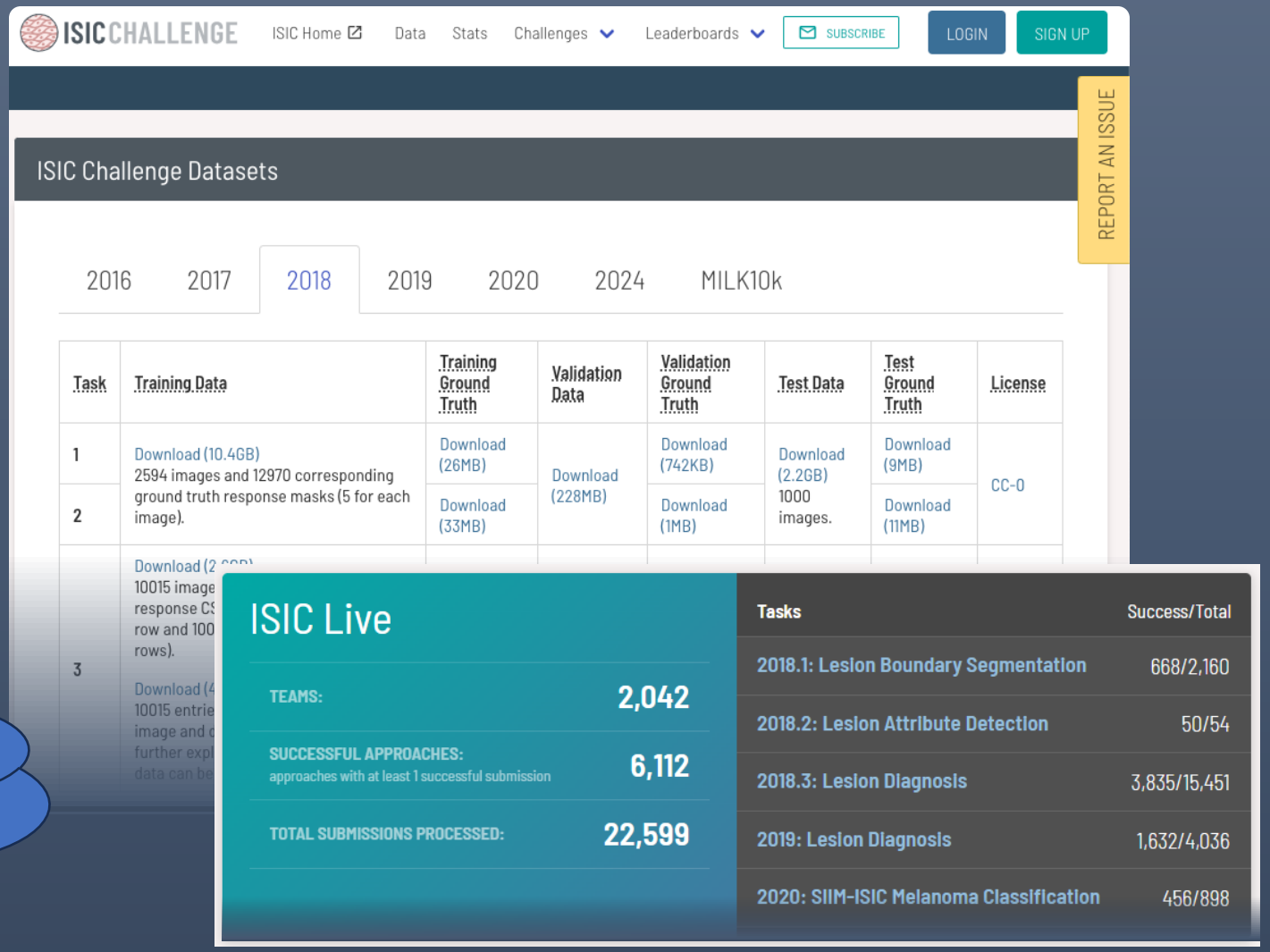
4. Prepare supplemental files (if necessary)
5. Contact us at support@isic-archive.com



Multiple ways to access:

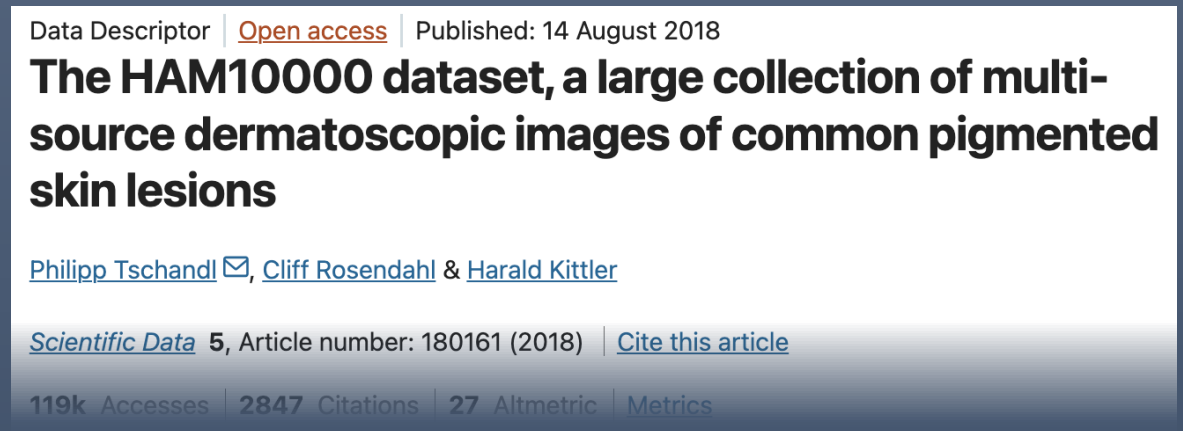
1. Gallery
2. Collection Site
3. Challenge Site

Benchmarking models & downloading competition data



HAM10000 dataset released in 2018

- 10,000 dermoscopic images
- Austria & Australia
- 7 diagnostic classes
- 2018 & 2019 Grand Challenges
- Thousands of citations



Data Descriptor | Open access | Published: 14 August 2018

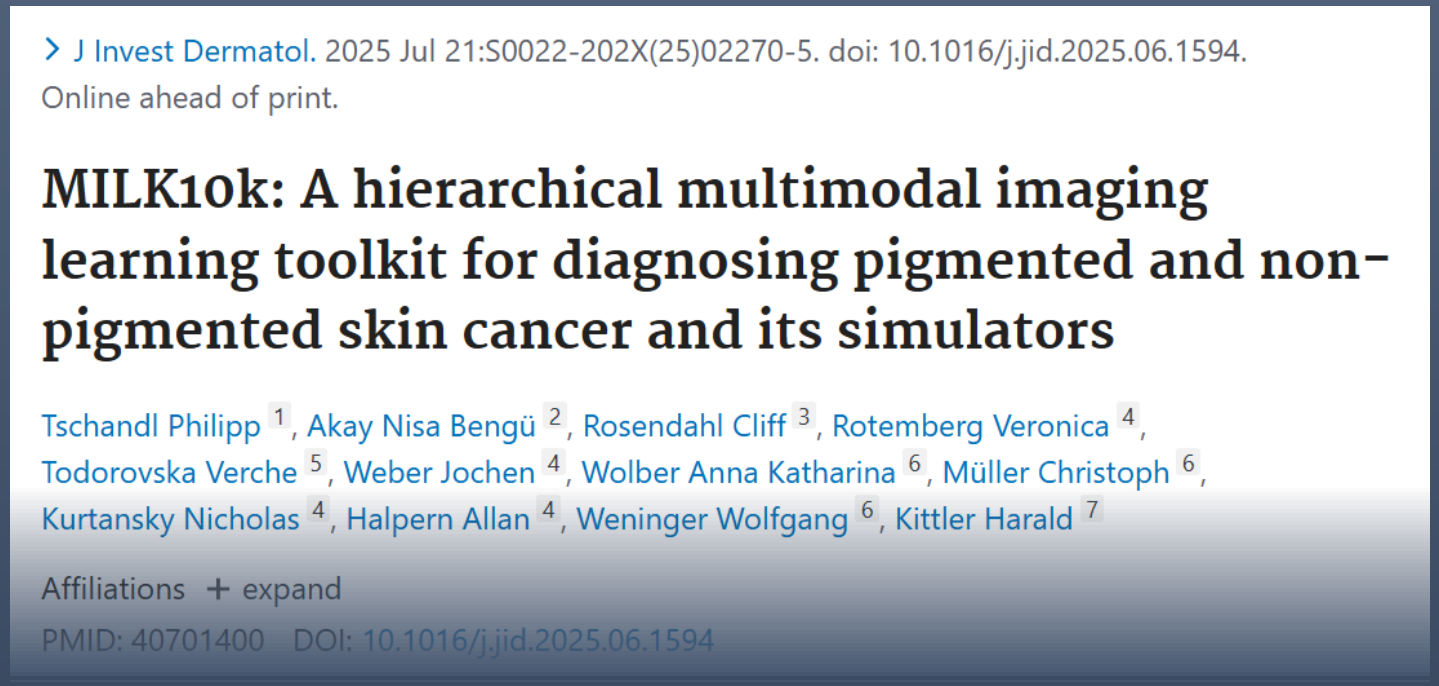
The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions

Philipp Tschandl ✉, Cliff Rosendahl & Harald Kittler

Scientific Data **5**, Article number: 180161 (2018) | Cite this article

119k Accesses | 2847 Citations | 27 Altmetric | Metrics

“MILK10k”
A new benchmarking platform



> J Invest Dermatol. 2025 Jul 21:S0022-202X(25)02270-5. doi: 10.1016/j.jid.2025.06.1594.
Online ahead of print.

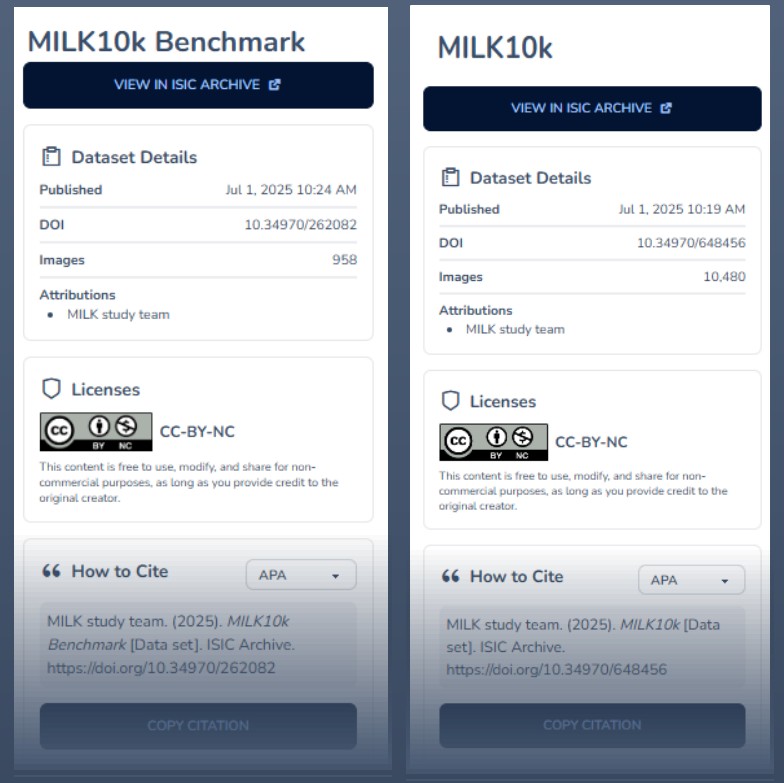
MILK10k: A hierarchical multimodal imaging learning toolkit for diagnosing pigmented and non-pigmented skin cancer and its simulators

Tschandl Philipp [1], Akay Nisa Bengü [2], Rosendahl Cliff [3], Rotemberg Veronica [4], Todorovska Verche [5], Weber Jochen [4], Wolber Anna Katharina [6], Müller Christoph [6], Kurtansky Nicholas [4], Halpern Allan [4], Weninger Wolfgang [6], Kittler Harald [7]

Affiliations + expand

PMID: 40701400 DOI: 10.1016/j.jid.2025.06.1594

Data descriptor



DOI Collections

“MILK10k”

Multi-modal dataset

Two parts

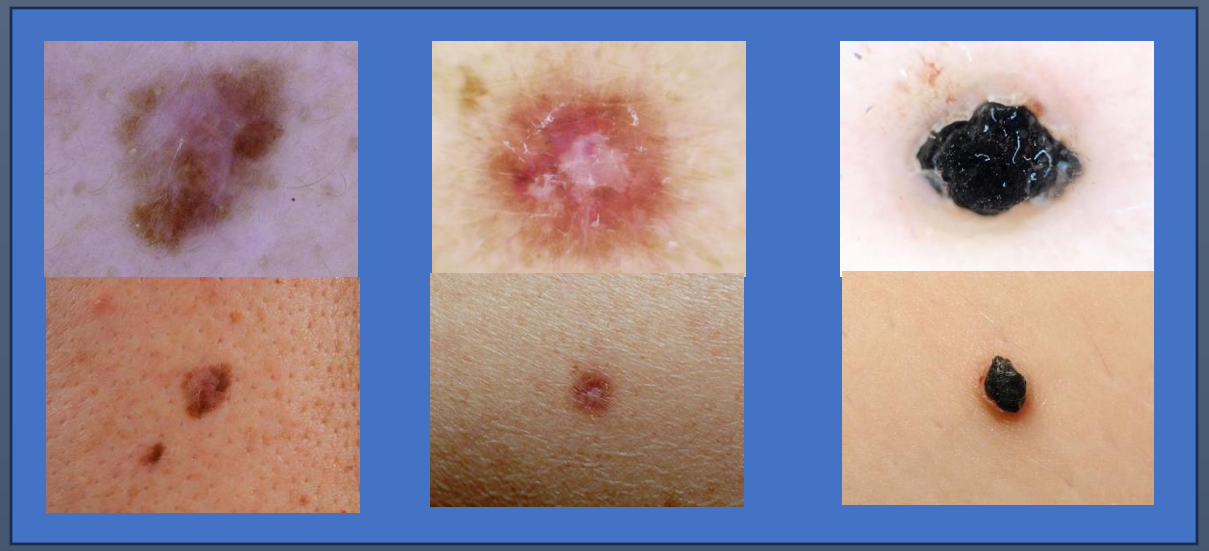
1. Training[1]
2. Test[2] (dx labels hidden)

Lesions contain two images

- Dermoscopic
- Clinical close-up

Metadata

- Diagnosis (ISIC-DX schema[3])
- Age, sex, anatomical site
- Skin tone
- MONET[4] framework annotations
  (1) ulceration/crust,
  (2) hair,
  (3) vasculature,
  (4) erythema,
  (5) pigmentation,
  (6) gel/fluid,
  (7) pen/skin markings

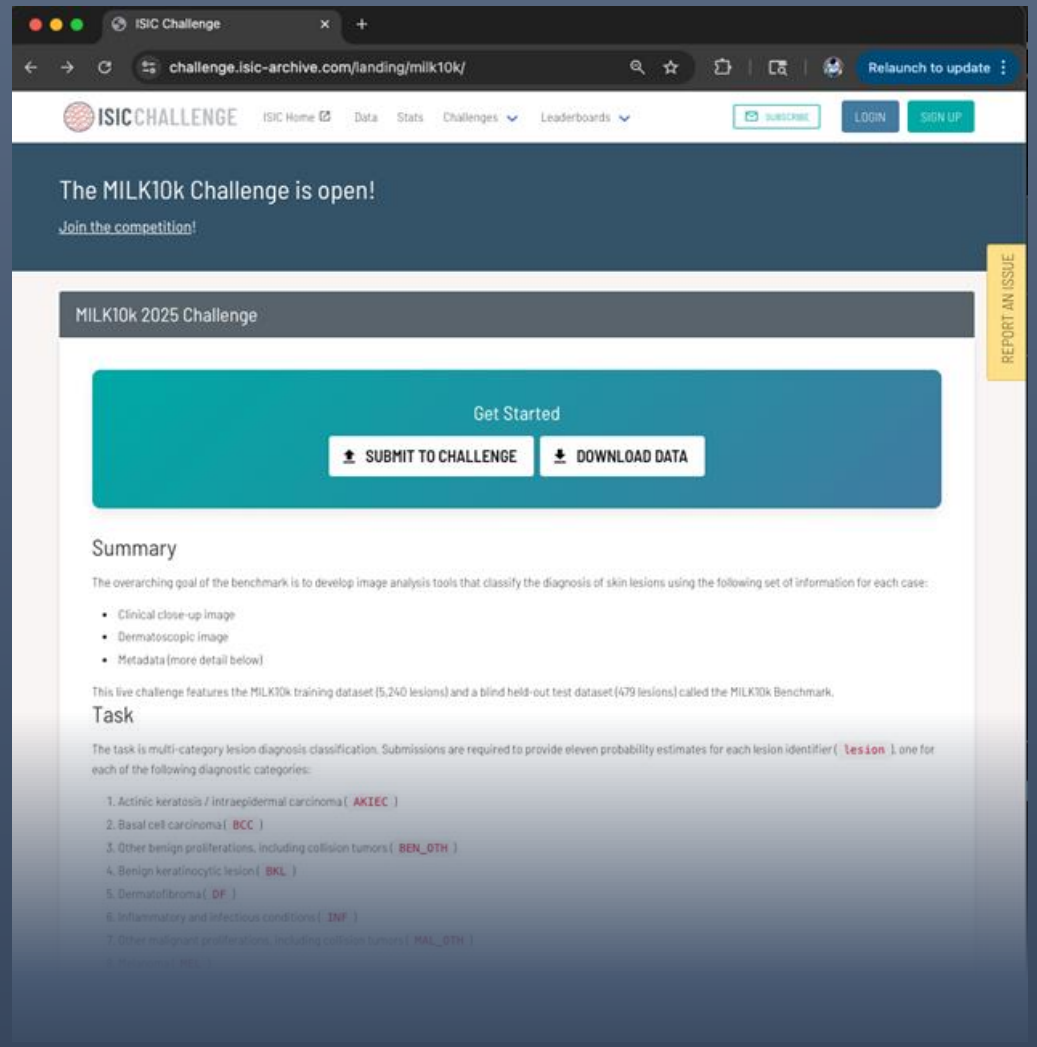


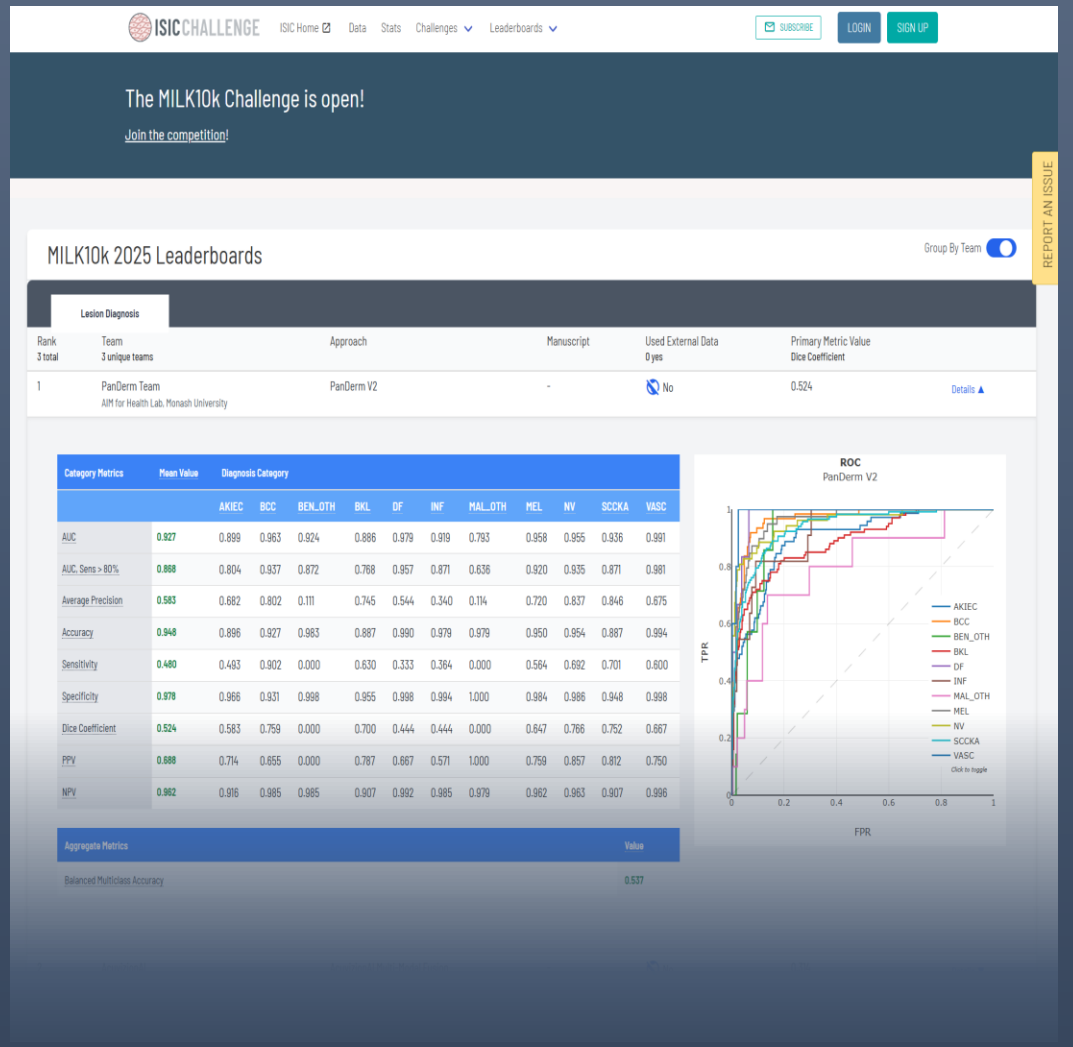
	Train *MILK10k*	Test *MILK10k Bench*
Lesion count	5,240	479
Image count	10,480	958

[1] MILK study team. (2025). *MILK10k* [Data set]. ISIC Archive. https://doi.org/10.34970/648456
[2] MILK study team. (2025). *MILK10k Benchmark* [Data set]. ISIC Archive. https://doi.org/10.34970/262082
[3] Scope A, Liopyris K, Weber J, et al. International Skin Imaging Collaboration-Designated Diagnoses (ISIC-DX): Consensus terminology for lesion diagnostic labeling. *J Eur Acad Dermatol Venereol*. 2025;39(1):117-125.
[4] Kim C, Gadgil SU, DeGrave AJ, et al. Transparent medical image AI via an image-text foundation model grounded in medical literature. *Nat Med*. 2024;30(4):1154-1165.

"MILK10k" Challenge Submission Platform

1. Why is it important?
2. Does it break anything?
3. Who is it for?



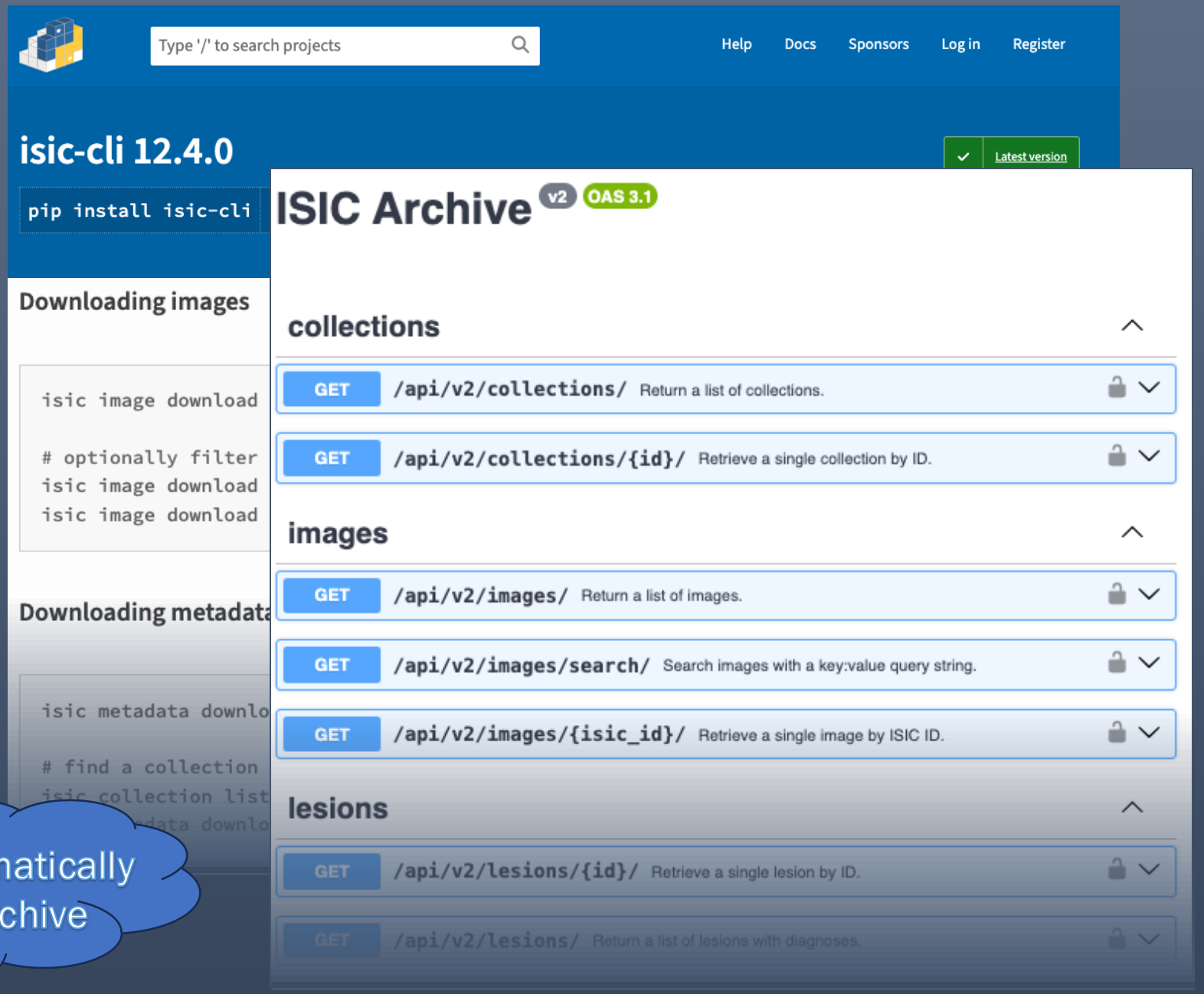


https://challenge.isic-archive.com/landing/milk10k/

Multiple ways to access:

1. Gallery
2. Collection Site
3. Challenge Site
4. Official CLI & API

Interact programmatically with the ISIC Archive



In Summary

Mission of ISIC

- Creating & promoting standards
- Engaging clinical and computer vision communities

How to find existing data on the ISIC Archive

- ISIC Gallery
- Pre-built collections
- Benchmarking platform

Benefits of creating your own DOI Collection

- How to create a custom image-selection
- Supplemental data,
- Draft DOIs
- Linking/machine-readability

Acknowledgments

National Institute of Health (NIH)

- [P30] Cancer Center Support Grant
- [U24] M-ISIC: Multimodal ISIC Informatics Platform for Skin Cancer Detection
- [U24] ISIC-REPO: Repository Enhancements for Promoting Interoperability

Melanoma Research Alliance (MRA)

- Team Science Award

Department of Defense (DOD)

- Melanoma Scholar
- MRP Team Science Award

Send us a message or visit our website

support@isic-archive.com

https://www.isic-archive.com/