

### Species Classification Automation for Microfossil Photomicrograph Images (Scampi)

David Wade\*, Sissa Stefanowicz, Alex Cullum & Erik Anthonissen





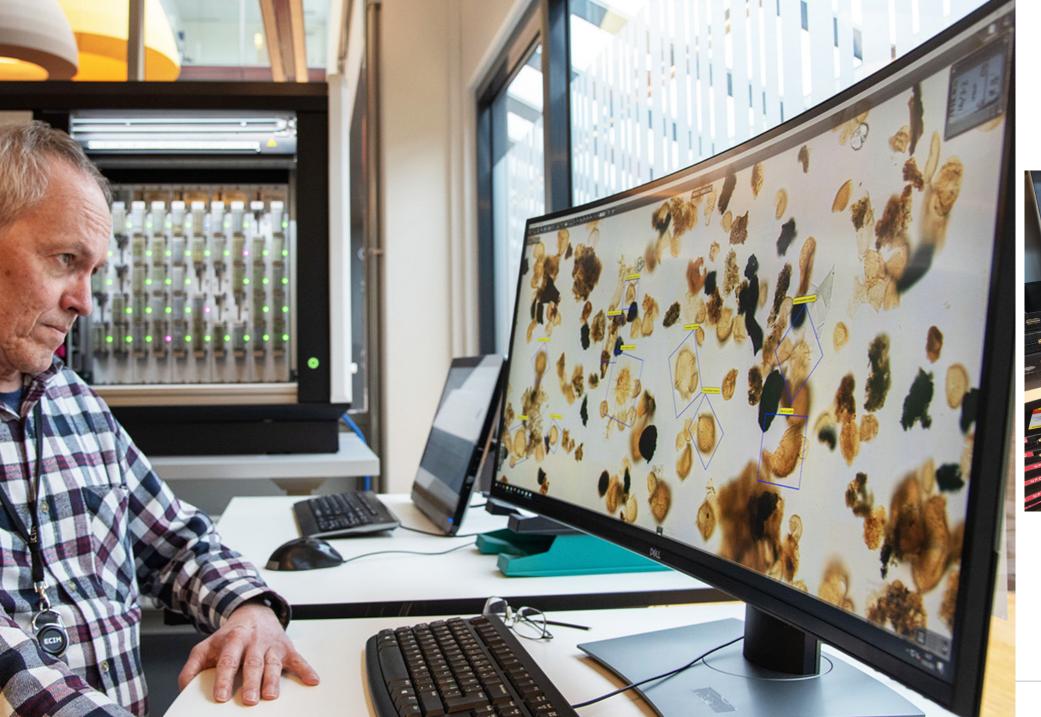
#### Palynology Slide Analysis

Traditional approach:

- Long hours with a microscope
- > Make species counts
- > Few practitioners









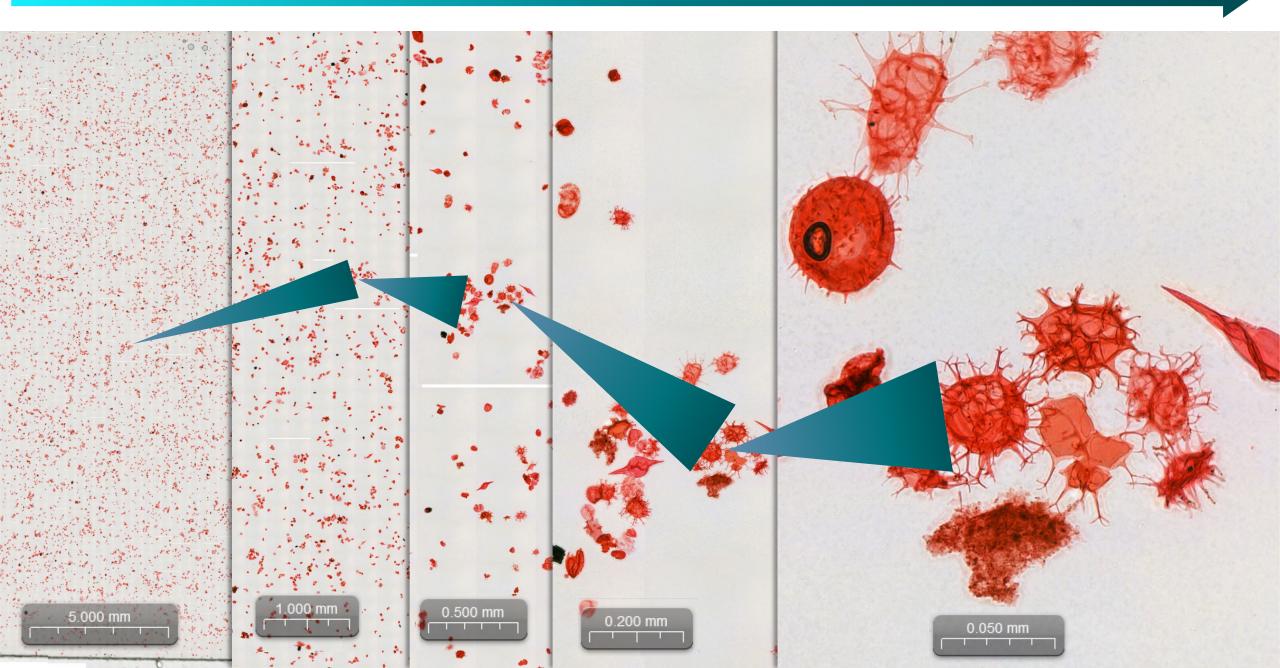




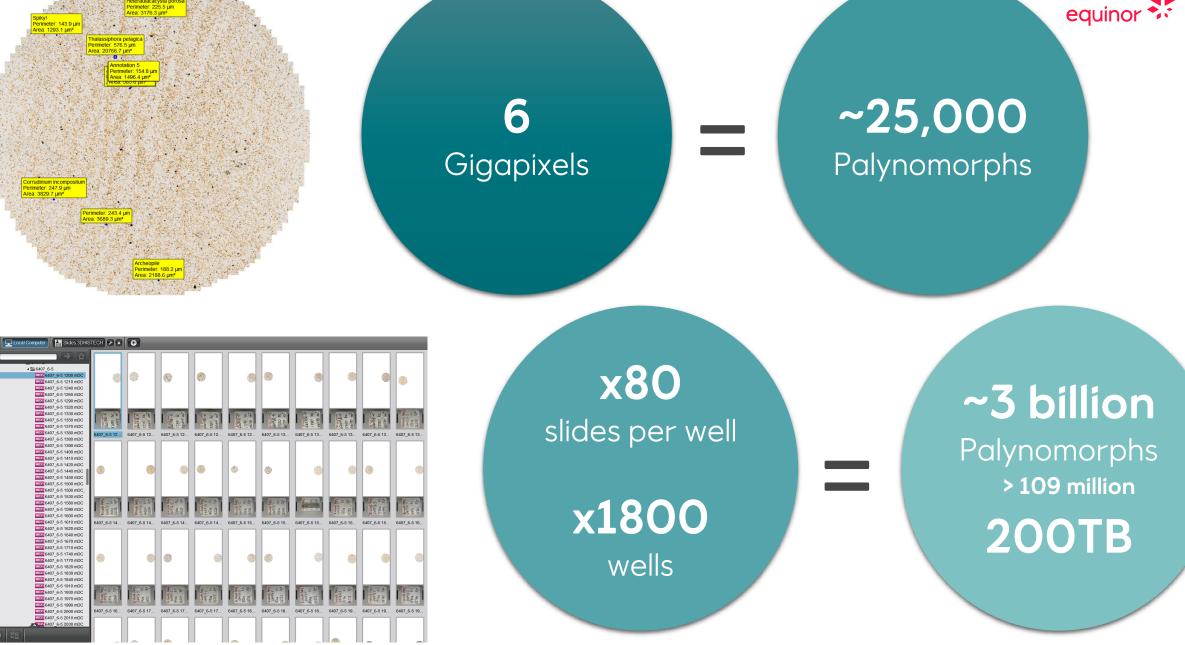
#### Zoom : Whole slide

... to ...

#### Individual fossils







Open



#### **Computer-Vision Origins**

• **1966** – ah, we should be able to crack this over the summer...

MASSACHUSETTS INSTITUTE OF TECHNOLOGY PROJECT MAC

Artificial Intelligence Group Vision Memo. No. 100. July 7, 1966

#### THE SUMMER VISION PROJEC

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".



#### **Computer-Vision Origins**

• **1966** – ah, we should be able to crack this over the summer...

• **2012** – major breakthrough using DNNs to classify images

≡ Google Scholar								
	Alex Krizhevsky <u>University of Toronto</u> Verified email at cs.toronto.edu - <u>Homep</u> Machine Learning		<b>EOLLOW</b>					
ARTICLES CITED BY								
TITLE		CITED BY	YEAR					
Imagenet classification with deep convolutional neural1124702012networksA Krizhevsky, I Sutskever, GE HintonAdvances in neural information processing systems 25								

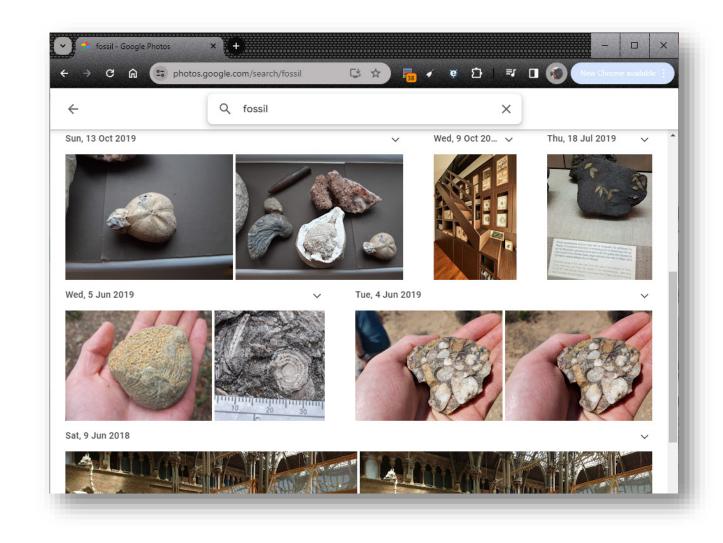


#### **Computer-Vision Origins**

• **1966** – ah, we should be able to crack this over the summer...

• **2012** – major breakthrough using DNNs to classify images

• **Now** – readily available commodity





#### Example results from supervised learning



### Correct prediction on image from outside training set

- ✓ Palaeocystodinium is here!
- ✓ No false positives

Incomplete prediction on image from outside training set

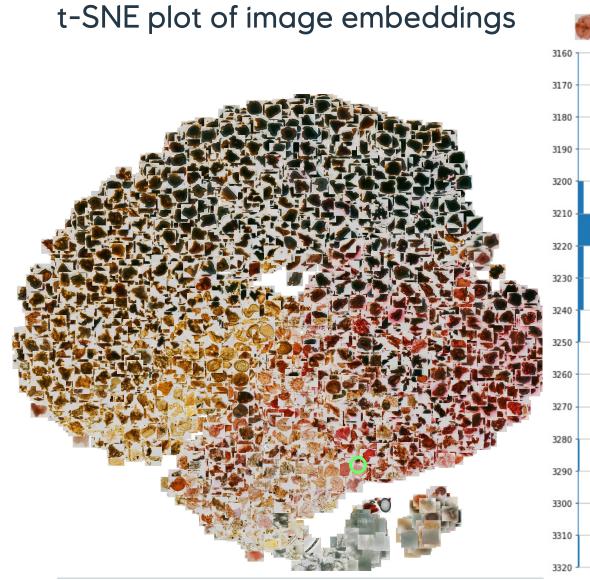
- ✓ Deflandrea phosphoritica here
- ✓ Second *Deflandrea* missing?

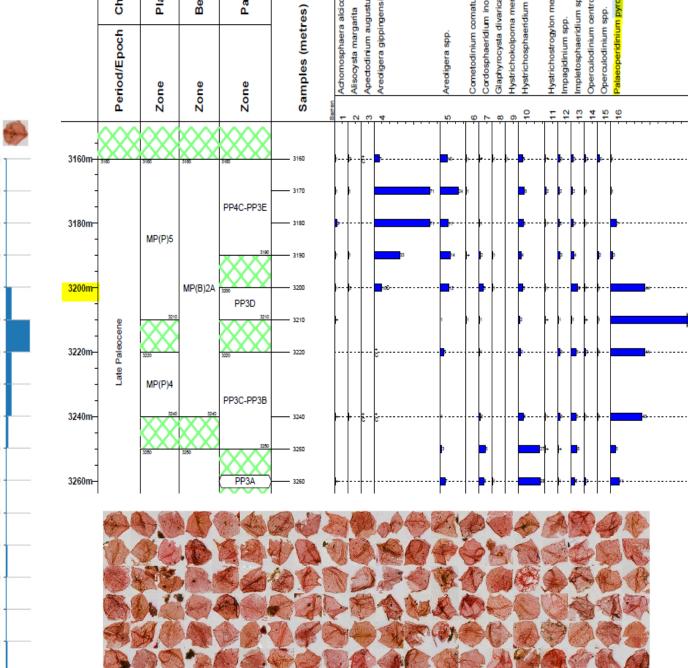
#### Confident and wrong

✓ Not a Deflandrea

Open

✓ Ignored pollen



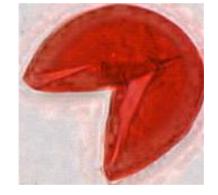


Open

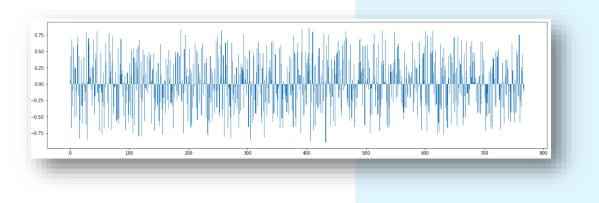


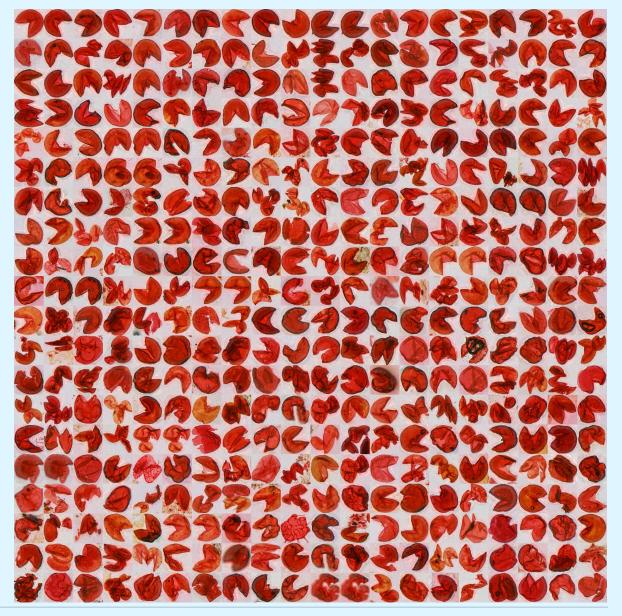
#### Content-Based Image Retrieval

Query :



Result :

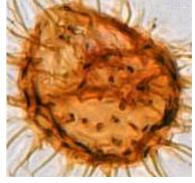




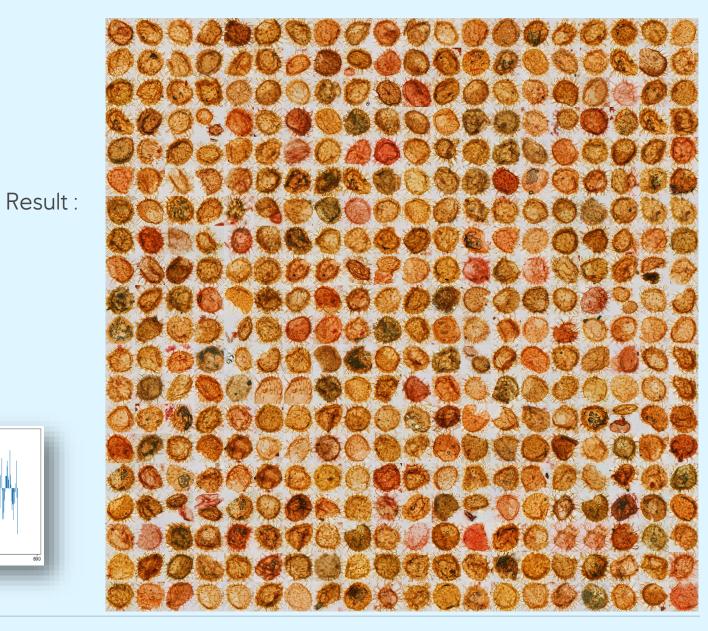


#### **Content-Based** Image Retrieval

Query :



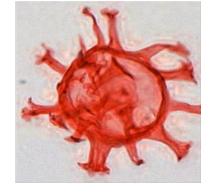
0.75 0.50 -0.25 0.00 -0.25 -0.50 --0.75 -1.00 100 200 300 400 500 600 700





#### Content-Based Image Retrieval

Query :



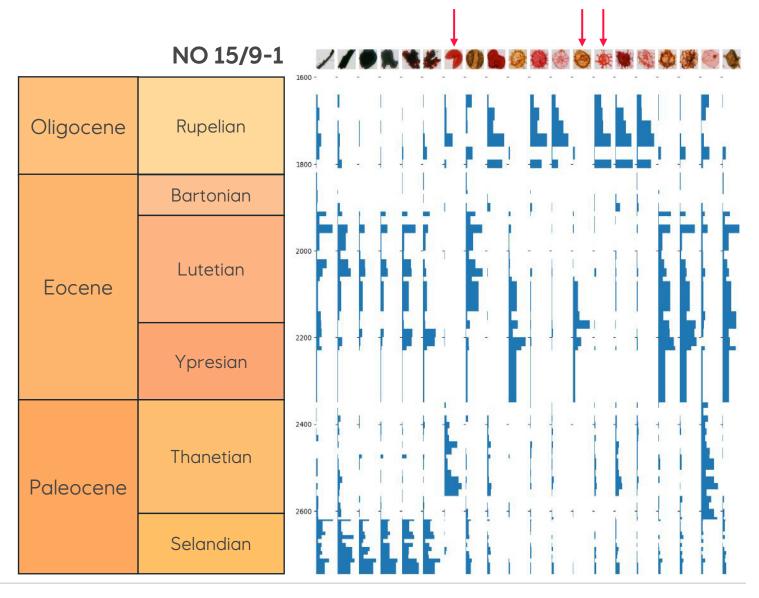
0.75 0.50 -0.25 0.00 --0.25 -0.50 --0.75 --1.00 100 200 300 ó 400 500 600 700 800





#### Chronostratigraphy

- Define multiple search crops
- Find depths of matching crops
- Plot histograms of matchs
- Compare with well picks





#### Use Case : Formation Zonation

#### Iteration #1: Queries from not-a-palynologist

This works surprisingly well:

... finds things which look like the queries

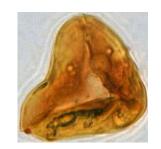
... but what do the results mean?!

### Well A – query results



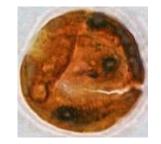


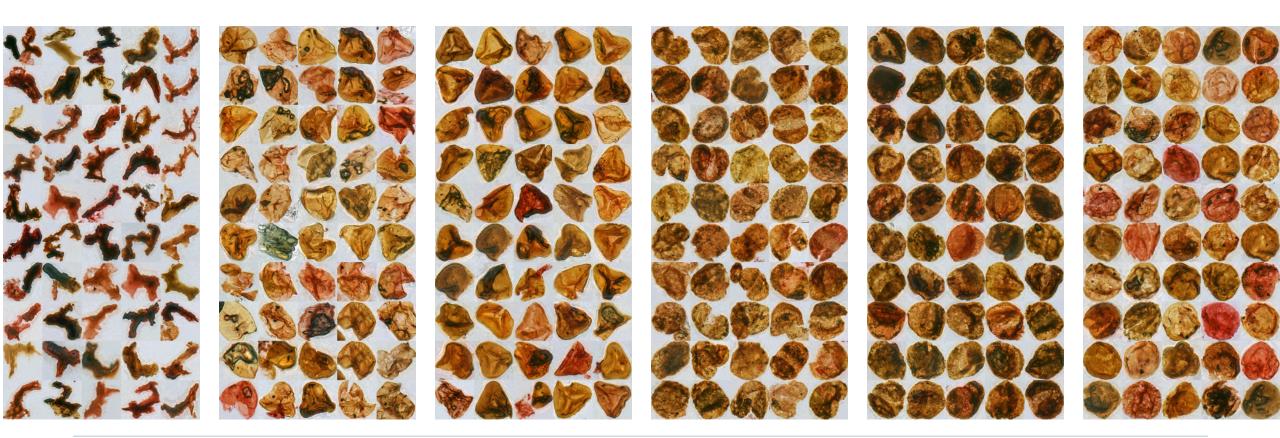










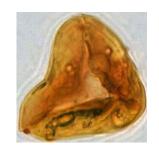


### Well B – query results

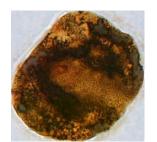


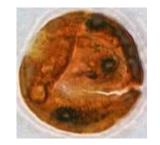


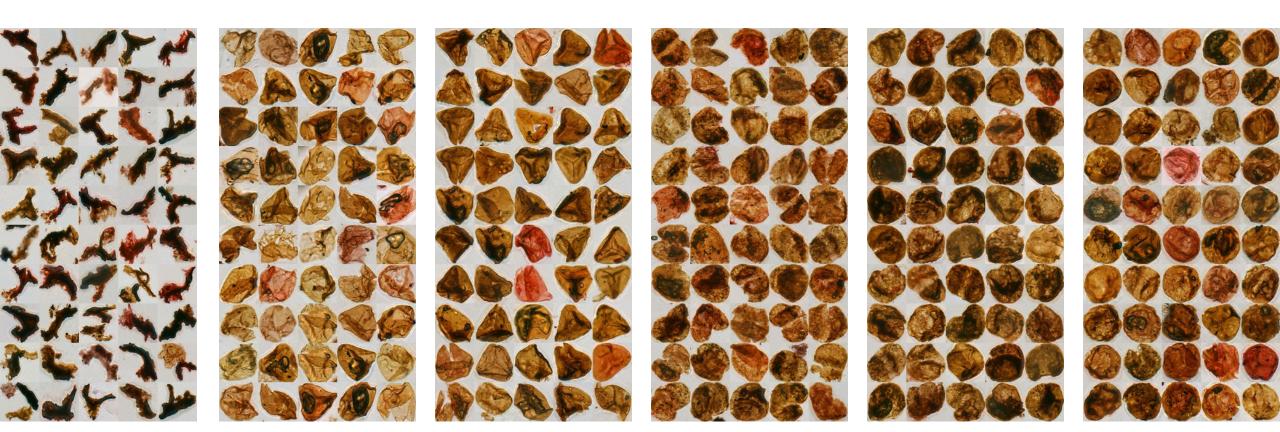










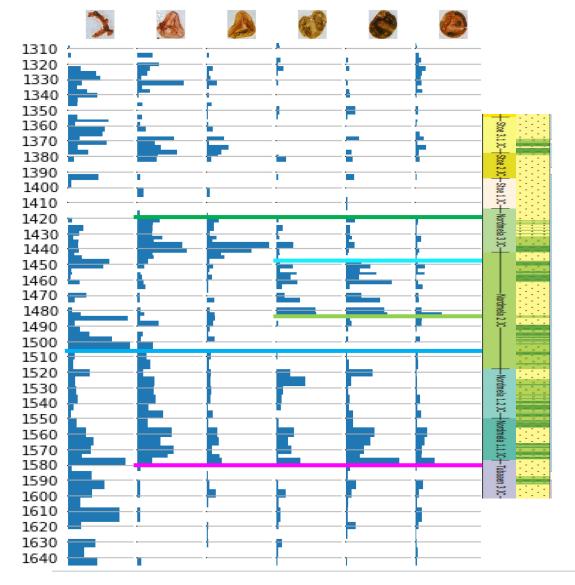


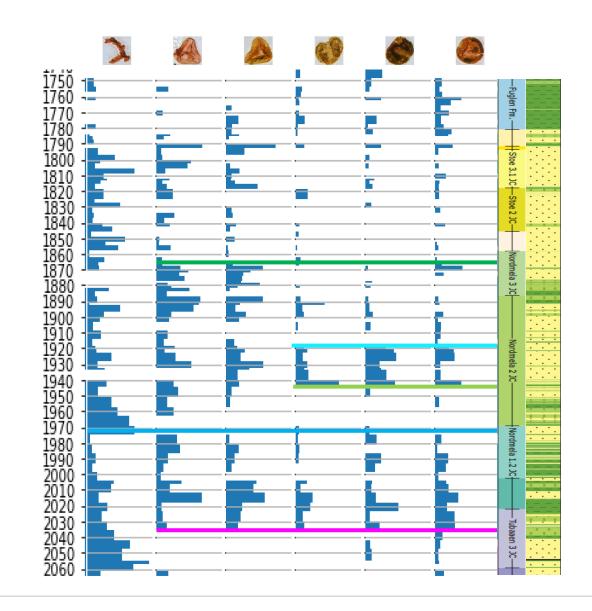
#### Scampi charts



Data scientist finds some interesting images









#### Use Case : Formation Zonation

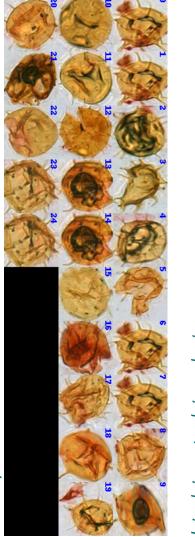
#### Iteration #2 : One real palynologist runs queries

#### Find exactly what we're looking for:

- ... interpretable results, fast!
- ... consistent patterns! (which nearly agree with official Equinor zonation)
- ... useful information for the asset!

### Micrhystridium

20 |



## Limbicysta bjaerkei

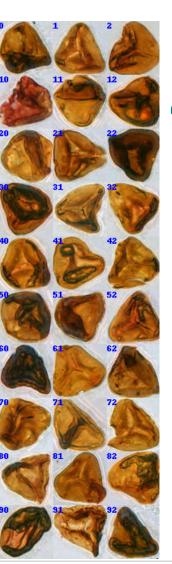


Well A – query results

# Vannoceratopsis spp.

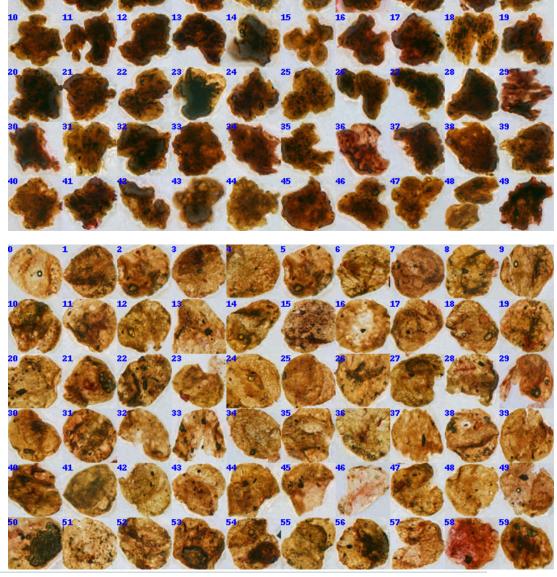


# Deltoidospora toralis



Bisaccate

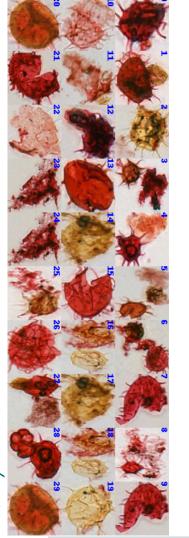
Botryococcus





Open

## Micrhystridium

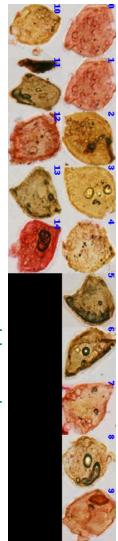


## Limbicysta bjaerkei

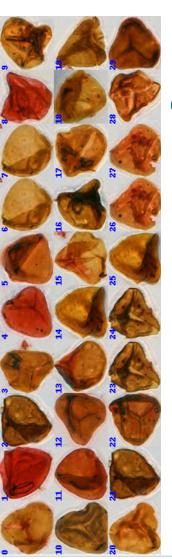


Well B – query results

# Vannoceratopsis spp.

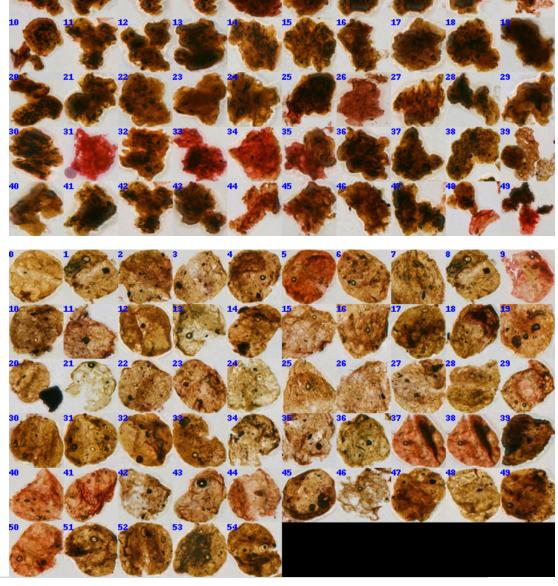


# Deltoidospora toralis



### Bisaccate

Botryococcus





Open



2040

2045 -



#### 1 1835 -Stoe 1400 1840 Toarcian Late 1405 Ы 1410 Toarcian Early Toarcian Early 1435 1875 -1445 1450 1895 R 1910 -1915 ------J. 1925 -1935 · 1525 1970 1535 1.2 Ы 1545 1995 2000 2005 2010 -Ы 1575 2015 1585 2025 2030 ien 3 JC-

Open

Rhaetian

1595

1600 -

1605

Scampi charts



#### Implementation

#### One webapp for all:

- Logging contractors
- Equinor biostrat team in Shared Services
- Equinor generalists

Select query crop	Library
Choose a file	
Drop your file here	
Search	

24 |

Select query crop

Library

MultiQuerySearch and LineProjection



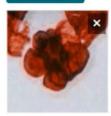
1634 1652 1670

0

0.1

0.2

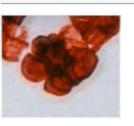
Choose a file



Selected file:

foram\_lining.JPG

Search

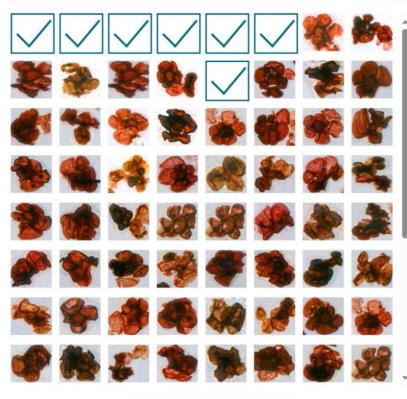


00

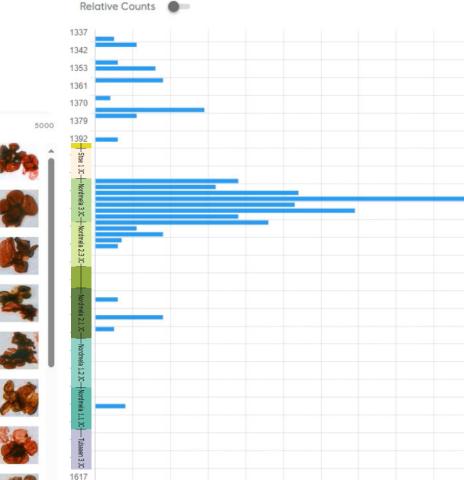
100

number of result slides to show is 100

× •



Number of crops loaded: 100/100



0.4

0.5

0.6

0.3

05 March 2024

0.8

0.9

1.0

0.7

Select query crop	NO 16/2-10	Î
	NO 16/2-12	
No well selected	NO 33/9-6	
	NO 34/10-30	
	NO 34/2-2	
	NO 34/2-3	_

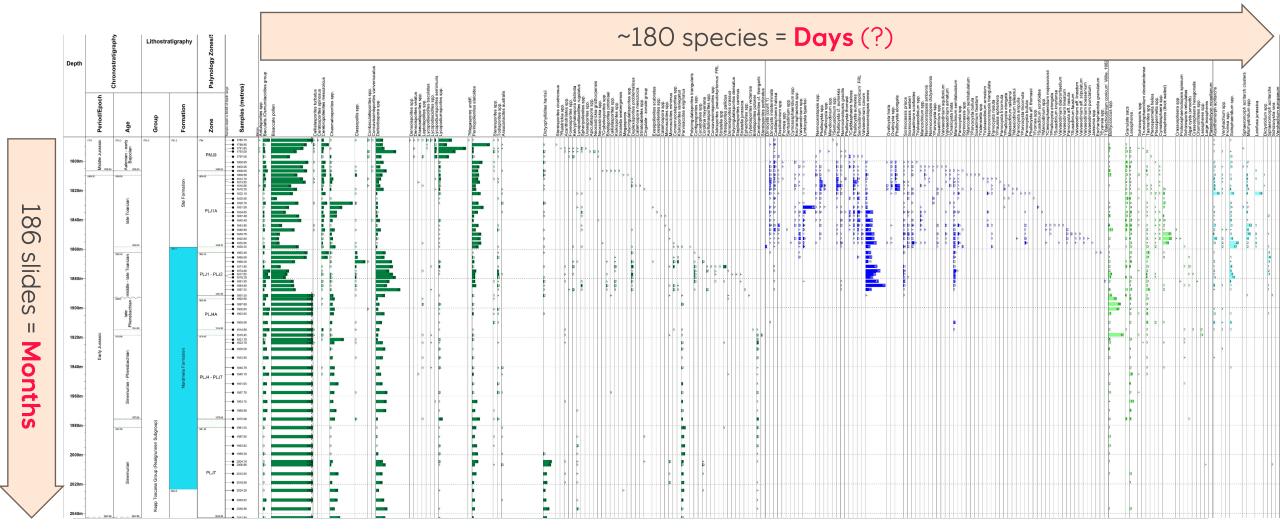
					(BIOS) AGES: P	Palynology	GENERA: DC	GENERA: SP			
	Litho-				(Cv excluded)	*	(Cv excluded)	(Cv excluded)		-	
	stratigr-	Chro	onostratigra	phy	*	*	*	*			
	aphy						DC Palaeocystodinium				
							DC Palaeoperidinium				Absolute Counts
							DC Phthanoperidinium	SP Azolla			460
							■DC ?Phthanoperidinium	SP Baculatisporites/O			520
							DC Polysphaeridium	SP Caryapollenites			580
							DC Reticulatosphaera	SP Cupuliferoipollenites			640
							DC Rhombodinium	SP Deltoidospora			700
		σ					DC Rottnestia	SP llexipollenites			760
		<u>.ŏ</u>	C L				DC Samlandia DC Saturnodinium	SP Inaperturopollenites SP Liquidambar			820
		er	Epoch		*DC 01.Pleist-Plio	DC	DC Saturnodinium DC Selenopemphix	SP Liquidambar SP Lycopodiumsporites	E a		880
		Period/Sub Period			*DC 06.Mio. E.	ALPR	DC Spiniferites	SP Momipites	t <b>mples</b> Palaeocystodinium Palaeoperidinium		940
6	u	Sul	Sub		*DC 07.Olig. L.	AC	DC Subtilisphaera	SP Nyssapollenites	stoc		1000
Samples	Formation	6/6	Epoch/Sub		*DC 08.Olig. E.	ALIN	DC Svalbardella	SP Schizosporis	Samples Palaeocys Palaeopen	Svalbardella Samples	1060
d	Ĕ	<u>io</u>	5 C	<i>a</i>	*DC 09.Eoc. L.	AL		■SP Sciadopityspollenites	n p alae alae	np	1120
ar	o	er	d	Age	*DC 10.Eoc. M.	DF	DC Tectatodinium	SP Tiliaepollenites		ar s	1180
0	ш	<u>п</u>	ш	< <	*DC 11.Eoc. E.	■ FU	DC Thalassiphora	SP Tsuga			1240
1300	1	1070	1210 1	1210	■*DC 12.Pal. L.	SP	DC Wetzeliella	SP bisaccate		□	1300
1310									1310 1320 <mark>_</mark> 5	1310	
1350			arly						— 1350		1100
1360 1370			Oligocene Early	an					— 1360 <mark>—</mark> 6 — — — — — — — — — — — — — — — — — —	13	
			ene	Rupelian							
1390 1400			Ö	Rul						12	
1410			Olig						—1410		1747
1420 1430		ē							— 1420 <mark>1</mark> 3	1420	1771
1440		Paleogene							——1440 💻 7 — — — — — — — — — — — — — — — — — —	1,40	
		leo	1460	1460					— 1450 — 1460 <mark>-</mark> 3	+ 1450	1807
		Ра	1470 Eocene Middle	<sup>1470</sup> Bartonian 1480							
1490			1490	1490 Lutetian					1490	1490	1002
	1500.0			1500					— 1500 +	1500 1510	
1520	Balder Fm.		Early Early Early	Ypresian						1520	1011
	Sele Fm.		1530.0	1530.0					— 1530 — 1540 — <b>—</b> 7		
1550	1540.0		Paleocene Late	Thanetian							
	Lista Fm.		1570	-20202020					— 1560 12 12 —	1560 1570	1990
			1580 Paleocene	1580	<					1580	
1590 1600		1620	Middle 1620	Selandian 1620					1590 1600	1590 1600	J

Select query crop	Library	MultiQuerySearch and LineProjection × • Refine search	
~	1520_mDC		
~	1540_mDC		Relative Counts
~	1560_mDC		520
^	1580_mDC	number of result slides to show is 1000	640 700
ee 🐄 🧶 🖤	💵 🥌 🔤 . D	0 1000 5000	760
19 🍆 🥦 🍘	🕌 🚄 🐂		820
			940
📥 🔬 🚵 🐜	<b>N</b>		1060 1120
<ul> <li>Ø</li> <li>Ø</li> </ul>	🛁 🐟 🎪 i		1180 1240
			1300
19 🗢 🐴 🔎	🧠 🦓 📢 I		1420
🔎 🎯 😹 🖌			1540
			1639
🔮 🍋 🐔 🦄	🤫 🚄 💓	i i i i i i i i i i i i i i i i i i i	1771 1789
			1807
			1825 1882
~	1600_mDC	<b>A A B 6 A 6 A A A</b>	1898
~	1620_mDC	Number of crops loaded: 100/1000	1931 1948
~	1633_mDC		0 5 10 15 20 25 30 35

					(BIOS) AGES: P	Palynology	GENERA: DC	GENERA: SP				
	Litho-				(Cv excluded)	*	(Cv excluded)	(Cv excluded)	-			
	stratigr- aphy	Chro	onostratigra	phy	-		*					
	~6)						DC Palaeocystodinium					Relative Counts
							<ul> <li>DC Palaeoperidinium</li> <li>DC Phthanoperidinium</li> </ul>	SP Azolla				460
							<ul> <li>DC Phthanoperidinium</li> <li>DC ?Phthanoperidinium</li> </ul>	SP Azolia				520
							DC Polysphaeridium	SP Caryapollenites				580
							DC Reticulatosphaera	SP Cupuliferoipollenites				640
							DC Rhombodinium DC Rottnestia	SP Deltoidospora				700
		ро	<del>,</del>				DC Samlandia	SP Inaperturopollenites				760
		eri	Epoch				DC Saturnodinium	SP Liquidambar		2		820
		<u>с</u>	Ш		*DC 01.Pleist-Plio *DC 06.Mio. E.	■ DC ■ ALPR	DC Selenopemphix DC Spiniferites	SP Lycopodiumsporites SP Momipites		iniur nium		940
s	u	Period/Sub Period	Epoch/Sub		*DC 07.Olig. L.	AC	DC Subtilisphaera	SP Nyssapollenites	s	Palaeocystodinium Palaeoperidinium Svalbardella	s	1000
Samples	Formation	s/p	Vų		*DC 08.Olig. E.		DC Svalbardella	SP Schizosporis	Samples	Palaeocysto Palaeoperid Svalbardella	Samples	1060
	nm	, jrio	00	g	*DC 09.Eoc. L. *DC 10.Eoc. M.	■ AL ■ DF	DC Tanyosphaeridium DC Tectatodinium	SP Sciadopityspollenites SP Tiliaepollenites		Pala Pala Sval		1120
S	ЪС	Ъ	Ш	Age	*DC 11.Eoc. E.	FU	DC Thalassiphora	SP Tsuga	S S	0.0	လိ	1180
1300 1310		1070	1210 1	1210	*DC 12.Pal. L.	SP	DC Wetzeliella	SP bisaccate	1300 1310		1300 1310	1240
1320									1320	<b>5</b>	— 1320	1360
			>								— 1330 — 1340	1420
			Earl	c l								1480
			Oligocene Early	Rupelian							— 1370	1540
			oce	dny	$\leftarrow$							
			Olig	<u> </u>							— 1410	1747
		e L	Ŭ						1420 		— 1420 — 1430	1771
		Paleogene							1440 1450		— 1440 — 1450	1789
1460 1470		ale	1460	1470 P. 4							— 1460 — 1470	1807
			Eocene Middle 1490	<sup>1470</sup> Bartonian <sub>1480</sub>					1480 1490			1825
	1500.0		1500 @ > 1	1500					1500	+		1882
	Balder Fm.		Early Early Early Early	Ypresian						11		1911
	Sele Fm. 1540.0		1540 1	1540					1540	7	1540	1931
			Paleocene Late	Thanetian						2 2	1550 1560	1948
	Lista Fm.		1580 Paleocene	-2020 2020 2020 2020 2020 2020 2020 202						+	1570 1580	0 5 10 15 20 25 30 35
1590 1600		1620	Middle 1620	Selandian 1620					1590 1600	69·	1590 1600	

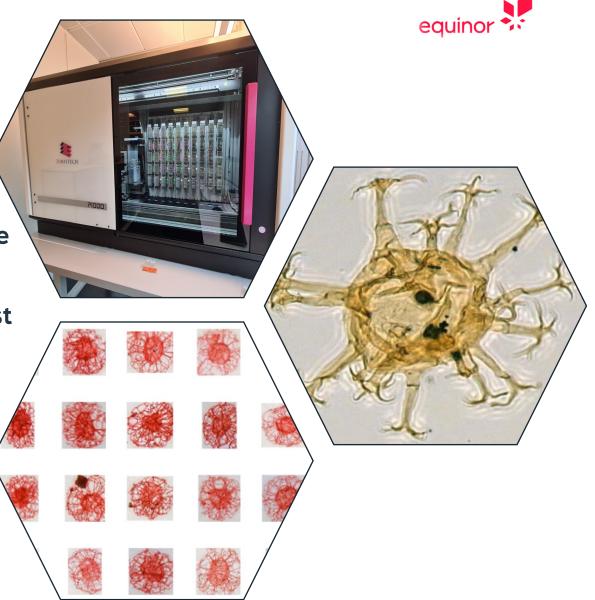


#### Imagine creating a distribution chart for a new well...



#### Summary

- > NOD's Digital Palynology is an amazing resource
- > Scampi enables us to maximize value from it, fast
- > Totally new way to view and use biostrat data
- > Improves communication with non-specialists





### Species Classification Automation for Microfossil Photomicrograph Images (Scampi)

David Wade<sup>\*</sup>, Sissa Stefanowicz, Alex Cullum & Erik Anthonissen Muhammad Gibran Alfarizi, Håkon Ruud, Thomas Elvestad<sup>‡</sup>, Einar Salomonsen<sup>‡</sup>, Thomas Karlsson<sup>‡</sup> <sup>‡</sup> Bouvet ASA

#### © Equinor ASA

This presentation, including the contents and arrangement of the contents of each individual page or the collection of the pages, is owned by Equinor. Copyright to all material including, but not limited to, written material, photographs, drawings, images, tables and data remains the property of Equinor. All rights reserved. Any other use, reproduction, translation, adaption, arrangement, alteration, distribution or storage of this presentation, in whole or in part, without the prior written permission of Equinor is prohibited. The information contained in this presentation may not be accurate, up to date or applicable to the circumstances of any particular case, despite our efforts. Equinor cannot accept any liability for any inaccuracies or omissions.