Focus on AI

Alex de Maria Antolinos
Software Engineer
Data Manager@Data Automation Unit
Software Group
ESRF

Index

- What is the current status of AI tool usage in the daily operations of SWG?
- How can AI be used to enhance user experience?

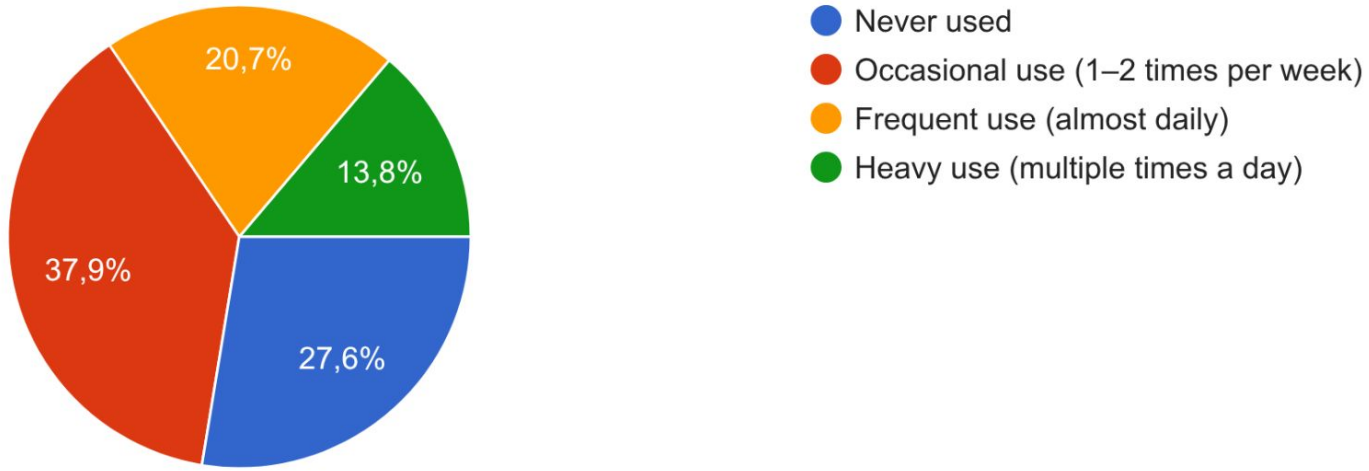
Survey

- **Software Group**
  - Inhouse and collaborative projects
  - Open source projects
    - Working on public repositories
  - No confidential data
  - No personal data

Survey: Usage Frequency

How frequently do you use generative AI tools (ChatGPT, Copilot, Claude, etc.)?

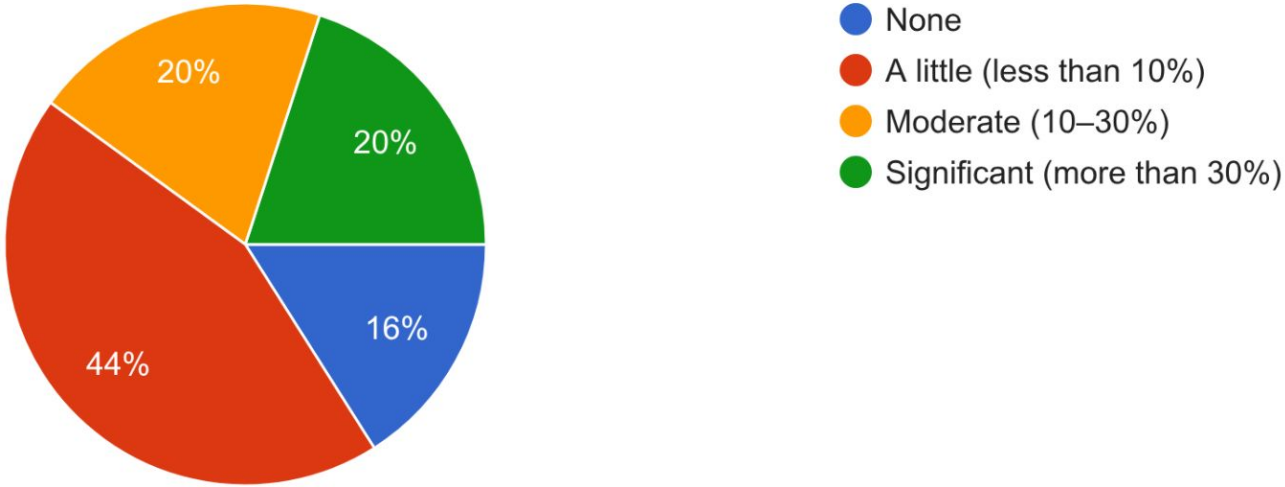
29 respuestas



Survey: Time saving

How much time would you say you save by using these tools?

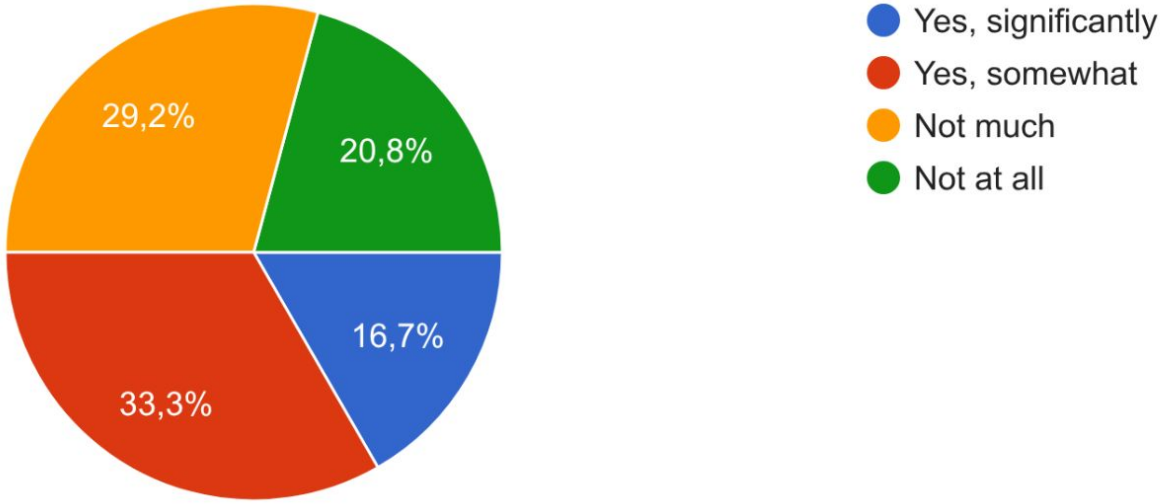
25 respuestas



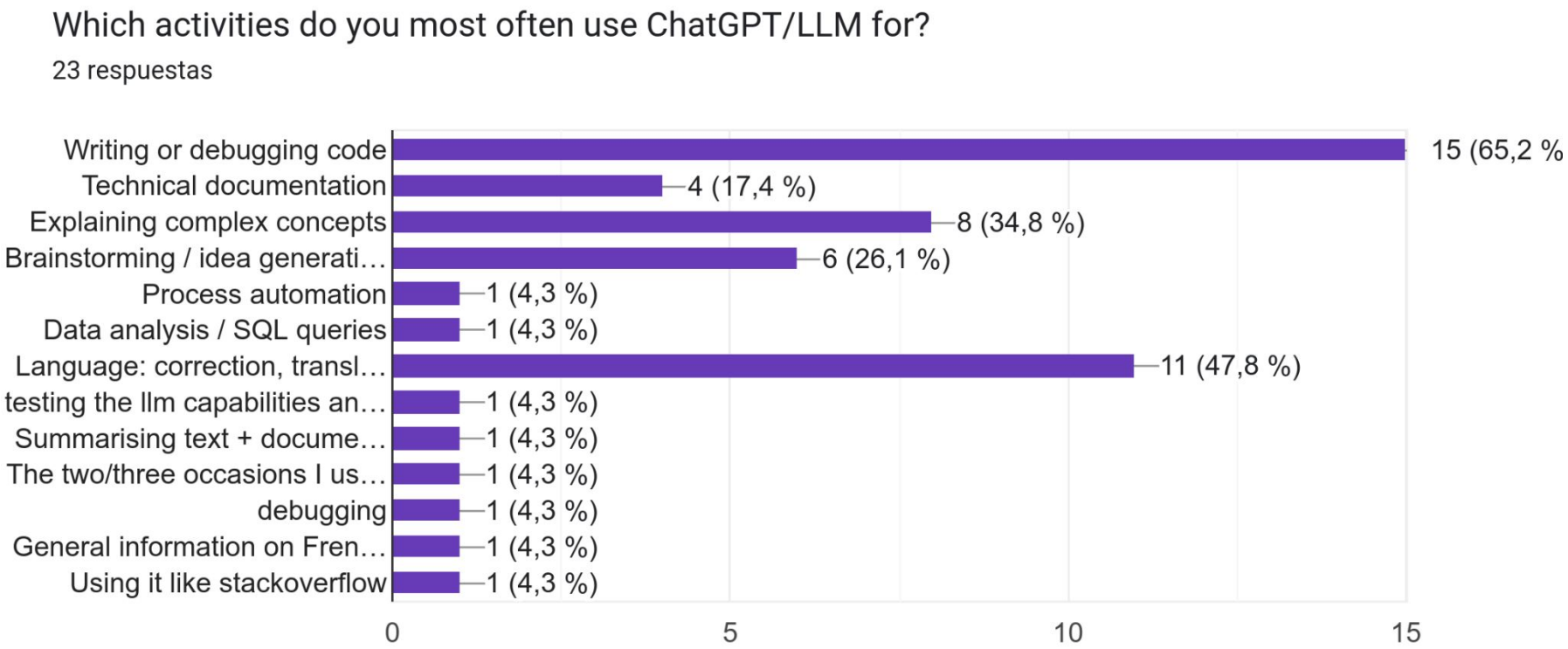
Survey: Quality Improvement

Has AI improved the quality of your work (e.g., fewer errors, more clarity, better ideas)?

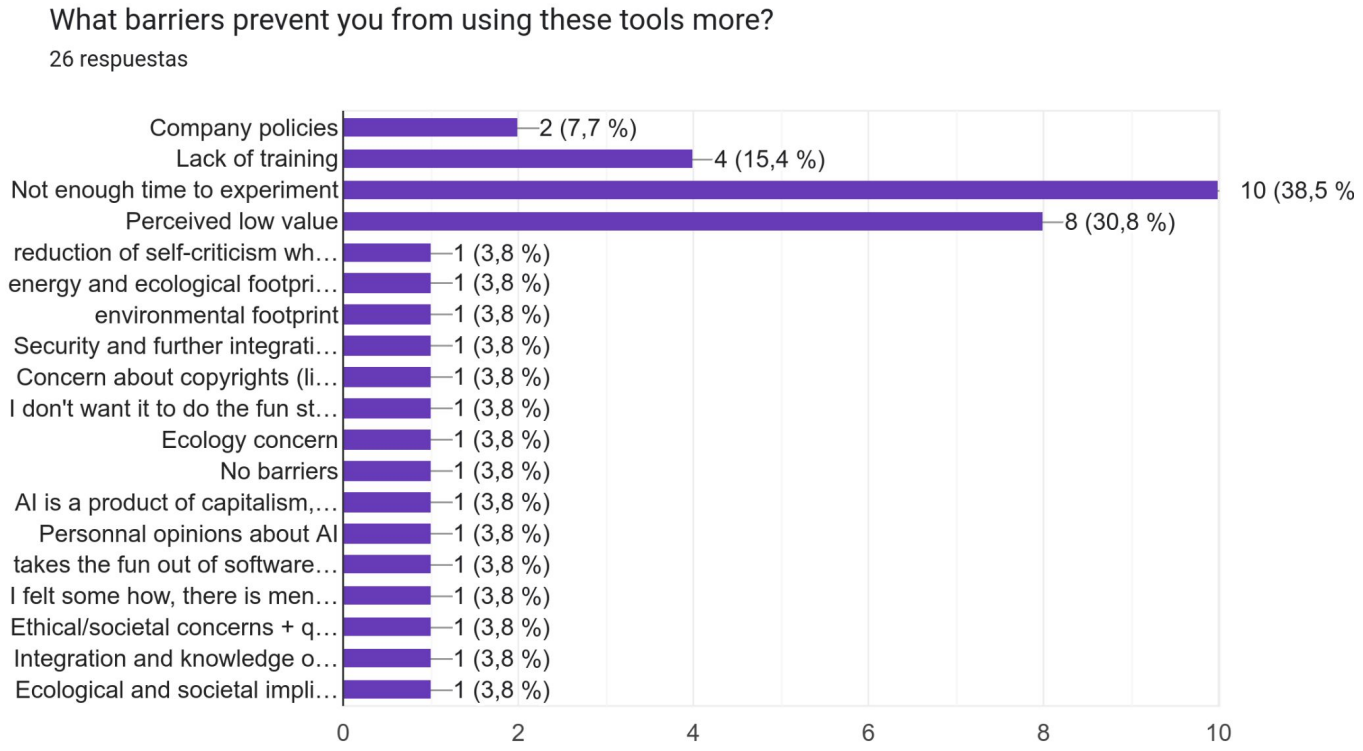
24 respuestas



Survey: Quality Improvement



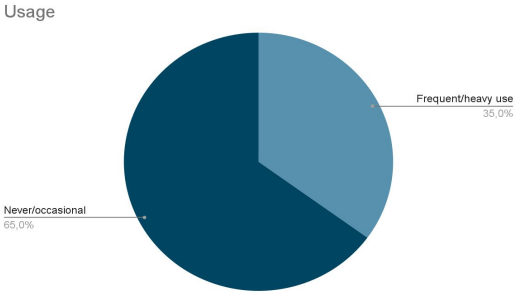
Survey: Barriers

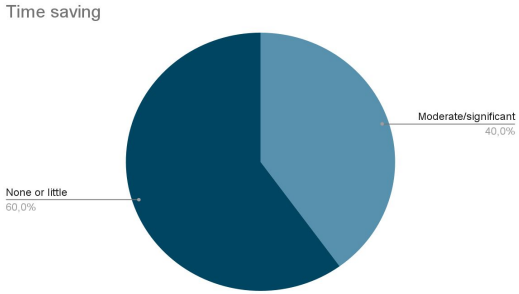


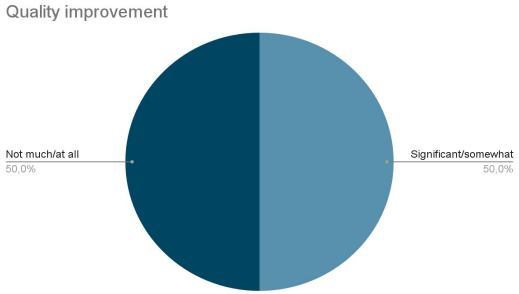
Survey: Barriers

- Ecological, carbon foot concerns (30%)
- Copyrights and license authorship unknowns (10%)
- Unknown medium/long term future business model (10%)
- Risk of losing ability to solve engineering problems (10%)

Survey: Wrap up







Main usage:

- Writing and debugging code
- Language correction
- Explaining complex ideas
- Brainstorming
- Producing technical documentation

Barriers:

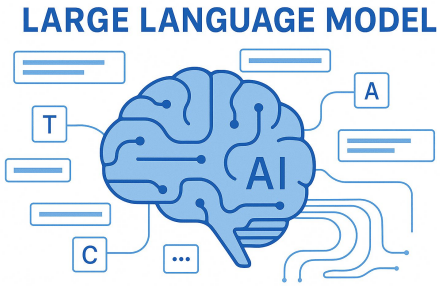
- Not enough time to experiment
- Ecological/ethical concerns
- Perceived low value
- Lack of training

In what ways can AI enhance the user experience?

DRAC

Data Repository
for Advancing open science

+



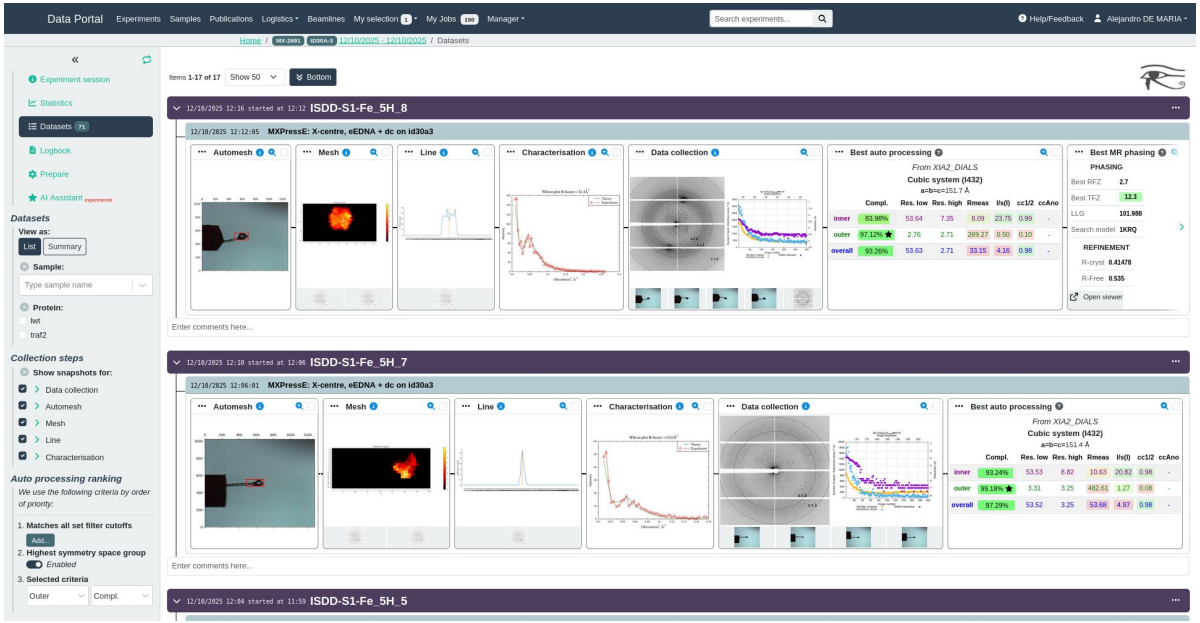
=

Catalog

- 40k Experiments
- 6M Datasets
- 170M exp. params
- ~30PB of data

Logbook

DRAC: https://data.esrf.fr



Example of the Data portal for a MX experiment at ID30A-3

Data portal

Automatic capture of the metadata
Automatic processing and analysis
Real time quality assessment

MX Experiment

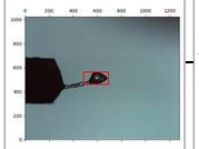
~2 minutes/sample
Hundreds samples/experiment
Thousand of datasets

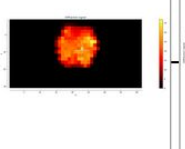
DRAC: Metadata Avalanch

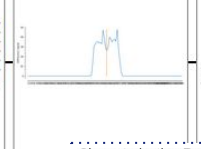
12/10/2025 12:16 started at 12:12 **ISDD-S1-Fe_5H_8**

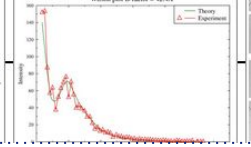
12/10/2025 12:12:05 **MXPressE: X-centre, eEDNA + dc on id30a3**

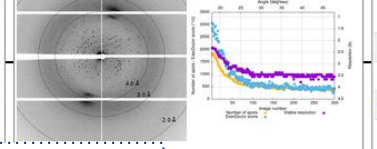
Automesh | Mesh | Line | Characterisation | Data collection











Best auto processing

From XIA2_DIALS

Cubic system (I432)

a=b=c=151.7 Å

	Compl.	Res. low	Res. high	Rmeas	I/s(I)	cc1/2	ccAno
inner	83.98%	53.64	7.35	8.09	23.75	0.99	-
outer	97.12% ★	2.76	2.71	269.27	0.50	0.10	-
overall	93.26%	53.63	2.71	33.15	4.16	0.98	-

Best MR phasing

PHASING

Best RFZ	2.7
Best TFZ	12.3
LLG	101.988
Search model	1KRQ

REFINEMENT

R-cryst	0.41478
R-Free	0.535

Open viewer

Grid size

Steps x	Steps y	Size x [mm]	Size y [mm]
41	20	0.342	0.175

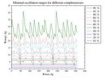
Characterisation Results

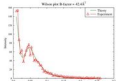
Indexing summary

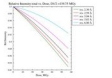
Selected spacegroup	a [Å]	b [Å]	c [Å]	alpha [°]	beta [°]	gamma [°]
I23	151.160	151.160	151.160	90.000	90.000	90.000

Collection plan strategy: resolution limit is set by the radiation damage

Wedge	Subwedge	Start (°)	Width (°)	No images	Exp time (s)	Max res (Å)	Rel trans (%)	Distance (mm)
1	1	18.00	0.10	300	0.167	2.30	10.00	169.95


compl


B


I2D

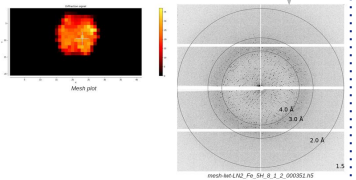

phi_overlap

Diffraction Plan

Forced space group	Anomalous data	Aimed multiplicity	Aimed completeness	Aimed I/sigma at highest res.	Aimed resolution (Å)	Min osc. width
None	False	Default (optimized)	Default (>= 0.99)	1.00		

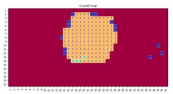
Program	a,b,c (Å)	α,β,γ (°)		Compl.	Res. low	Res. high	Rmeas	I/s(I)	cc1/2	ccAno
12/10/2025 11:21 XIA2_DIALS I432	151.7	90.0	inner	83.98%	53.64	7.35	8.09	23.75	0.99	-
	151.7	90.0	outer	97.12% ★	2.76	2.71	269.27	0.50	0.10	-
	151.7	90.0	overall	93.26%	53.63	2.71	33.15	4.16	0.98	-
12/10/2025 12:17 EDNA_proc I432	150.4	90.0	inner	84.20%	40.21	12.51	10.20	15.50	0.99	-
	150.4	90.0	outer	95.10% ★	3.34	3.23	126.20	2.00	0.55	-
	150.4	90.0	overall	93.00%	40.21	3.23	35.90	6.00	0.97	-
12/10/2025 12:18 grenades_parallelproc I432	151.8	90.0	inner	87.60%	107.34	8.88	6.50	21.10	0.99	-
	151.8	90.0	outer	94.60% ★	3.35	3.14	88.70	2.60	0.68	-
	151.8	90.0	overall	93.20%	107.34	3.14	25.50	7.80	0.98	-
12/10/2025 12:18 trimmed_grenades_fastproc I432	151.4	90.0	inner	86.90%	107.03	8.64	7.70	19.10	0.99	-
	151.4	90.0	outer	94.10% ★	3.27	3.06	99.10	1.80	0.58	-
	151.4	90.0	overall	93.10%	107.03	3.06	27.30	6.60	0.98	-
12/10/2025 12:23 autoPROC I432	151.9	90.0	inner	65.90%	62.03	6.94	12.00	8.40	0.97	-
	151.9	90.0	outer	89.90% ★	2.60	2.56	382.90	0.90	0.36	-
	151.9	90.0	overall	82.40%	62.03	2.56	36.30	3.50	0.96	-
12/10/2025 12:17 grenades_fastproc I432	151.9	90.0	inner	87.50%	107.43	8.76	7.30	20.20	0.99	-
	151.9	90.0	outer	89.30% ★	3.31	3.10	99.90	1.70	0.56	-
	151.9	90.0	overall	92.10%	107.43	3.10	26.60	6.90	0.98	-
12/10/2025 12:23 autoPROC_staraniso I432	151.9	90.0	inner	65.80%	107.44	7.07	12.00	8.30	0.97	-
	151.9	90.0	outer	29.70% ★	2.63	2.6 2.6 2.6	190.40	1.50	0.43	-
	151.9	90.0	overall	70.90%	107.44	2.6 2.6 2.6	30.50	3.90	0.97	-
	151.8	90.0	inner	85.40%	107.34	8.57	6.40	18.10	0.99	-
	151.8	90.0	outer	92.10% ★	3.50	3.24	80.60	2.30	0.70	-
	151.8	90.0	overall	92.70%	107.34	3.24	25.30	6.50	0.98	-

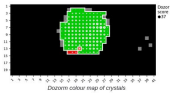
Results of mesh scan



dozorm results

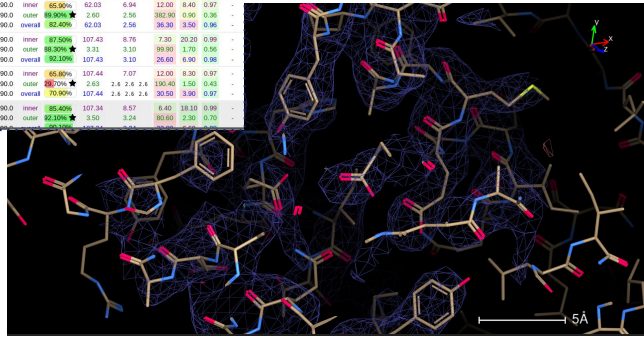
Crystal number	Aperture	Resolution	Central Image	X	Y	Int/ Sig	Images all	Images dX	Images dY	Images sum	Score Helical	Start X	Start Y	End X	End Y	Helical Int/Sig
1	15	2.44	386	23	9.5	35	143	14	13	2626.9	Yes	18	2	24	14	273.7
2	7	2.69	512	20	13	14	4	4	2	43.6	Yes	17	14	18	14	20.7


Dozorm map of crystals

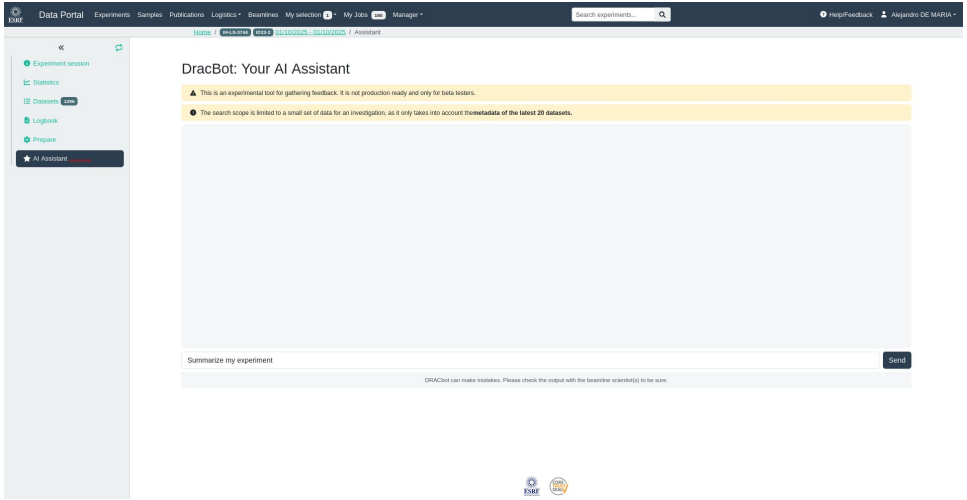

Dozorm colour map of crystals

All positions

No	X	Y	Image file	Dozor score	Dozor resolution	Motor positions for moving to sample position
1	1	70	line-list-LN2_Fe_5H_8_1_3_000070.h5	47.921	2.3	sampx = -0.264 mm, sampy = 0.616 mm, phiy = -0.349 mm
2	1	59	line-list-LN2_Fe_5H_8_1_3_000059.h5	47.334	2.3	sampx = -0.210 mm, sampy = 0.616 mm, phiy = -0.349 mm
3	1	64	line-list-LN2_Fe_5H_8_1_3_000064.h5	39.991	2.5	sampx = -0.234 mm, sampy = 0.616 mm, phiy = -0.349 mm
4	1	60	line-list-LN2_Fe_5H_8_1_3_000060.h5	39.27	2.5	sampx = -0.215 mm, sampy = 0.616 mm, phiy = -0.349 mm
5	1	68	line-list-LN2_Fe_5H_8_1_3_000068.h5	38.007	2.5	sampx = -0.254 mm, sampy = 0.616 mm, phiy = -0.349 mm
6	1	65	line-list-LN2_Fe_5H_8_1_3_000065.h5	37.973	2.5	sampx = -0.239 mm, sampy = 0.616 mm, phiy = -0.349 mm
7	1	67	line-list-LN2_Fe_5H_8_1_3_000067.h5	37.406	2.5	sampx = -0.249 mm, sampy = 0.616 mm, phiy = -0.349 mm
8	1	69	line-list-LN2_Fe_5H_8_1_3_000069.h5	35.611	2.5	sampx = -0.259 mm, sampy = 0.616 mm, phiy = -0.349 mm
9	1	55	line-list-LN2_Fe_5H_8_1_3_000055.h5	35.595	2.5	sampx = -0.190 mm, sampy = 0.616 mm, phiy = -0.349 mm



DRACBot: AI Assistant



POC

- A proof of concept deployed on 16th Oct
- Accessible to ESRF beamline staff only

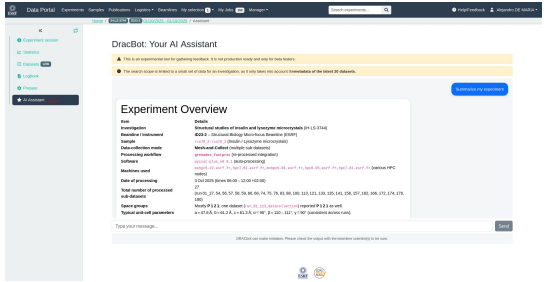
RAG (Retrieval Augmented Generation)

- Limited context send to the LLM
  - Metadata of latest datasets
- LLM installed by TID: gpt-oss

Evaluation purposes only

- Gather feedback
- Inspiration about use cases

DRACBot: Example of “summarize my experiment”



Data-quality Summary (representative values)

Metric	Typical range across runs
Overall completeness	**7 % – 15 %** (most runs 12-15 %)
Overall ⟨I/σ(I)⟩	**0.8 – 2.5** (inner shells often much higher, up to ~40 for the best runs)
Overall CC½	**0.96 – 0.998** (inner shells often > 0.99)
Overall R_merge	**0 – 30 %** (many runs report 0, a few high values up to ~95 % for low-resolution shells)
Overall R_meas	**0 – 65 %** (some anomalously large negative values indicate statistical noise in very low-count data)
Overall R_pim	**0 – 27 %** (most ≤ 12 %)
Overall multiplicity	**1.0 – 1.3** (generally low, as expected for single-crystal mesh scans)
Resolution limits	**High-resolution shell** typically **1.63 – 1.66 Å**; **low-resolution cut-off** varies from **~18 Å** to **> 40 Å** depending on crystal quality.
Anomalous signal	Generally **absent** (overall anomalous multiplicity ≈ 0, anomalous completeness ≈ 0 – 0.3 %).

Note: Many parameters are reported as **0** for outer shells (e.g., R_merge, R_meas) because the corresponding reflections are either absent or have intensities below the detection threshold.

Key Observations

1. **Consistent crystal lattice** – All processed sub-datasets share nearly identical unit-cell dimensions, confirming that the same crystal (or identical crystals) were used throughout the mesh-and-collect experiment.
2. **High inner-shell data quality** – The inner resolution shells routinely show very high ⟨I/σ(I)⟩ (up to ~40) and CC½ > 0.99, indicating that the diffraction spots are strong and well measured.
3. **Low overall completeness** – Because each mesh-and-collect sub-dataset only samples a small portion of reciprocal space, overall completeness per sub-dataset is modest (7-15 %). The full experiment (combining all sub-datasets) would be expected to reach a much higher completeness.
4. **Resolution** – The majority of runs reach a high-resolution limit of **~1.64 Å**, sufficient for detailed atomic modelling of insulin and lysozyme.
5. **Anomalous signal** – No significant anomalous signal was detected, which is typical for native protein crystals without heavy-atom derivatization.
6. **Processing consistency** – All datasets were processed with the same pipeline (`grenades_fastproc` → `pyicat-plus_v0.9.1`) on ESRF HPC resources, ensuring comparable statistical treatment.

Conclusion

The mesh-and-collect experiment on the ID23-2 beamline generated a large set of high-resolution diffraction sub-datasets from insulin/lysozyme microcrystals. While each individual sub-dataset has limited completeness, the data quality within the measured resolution shells is excellent (high CC½, strong I/σ(I), and low R_merge/R_meas). When merged, these data should provide a robust structural model of the target proteins at ~1.6 Å resolution.

All information extracted from the dataset metadata supplied (27 re-processed integration datasets, IDs 2260453301 – 2260552388).

DRACBot: Utils

write a list of scp command line that will copy the data to demariaa@myhome:/data

Below are the **scp** commands you can run (one per dataset) to copy each processed-data folder to the remote host **demariaa@myhome** under **/data**.
Each command uses the `-r` flag to copy the directory recursively.

```
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_36_datacollection/
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_88_datacollection/
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_76_datacollection/
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_133_datacollection
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_54_datacollection/
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_56_datacollection/
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_113_datacollection
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_156_datacollection
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_59_datacollection/
```

Type your message...

Send

DRACbot can make mistakes. Please check the output with the beamline scientist(s) to be sure.

DRACBot: Utils

write a list of scp command lines that copies the data to demariaa@myhome:/data for the crystals that are not diffracting or resolution between 2 and 3 amstrongs

SCP commands for the datasets that show little or no diffraction (overall I/σ(I) < 1)

```
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_68_datacollection/a
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_74_datacollection/a
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_135_datacollection.
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_59_datacollection/a
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_56_datacollection/a
scp -r /data/visitor/ihls3744/id23eh2/20251001/PROCESSED_DATA/rco78_2/rco78_2-rco78_2/run_01_MeshAndCollect/run_01_88_datacollection/a
```

Type your message...

Send

DRACbot can make mistakes. Please check the output with the beamline scientist(s) to be sure.

DRACBot: Utils

do a python script that will plot the resolution of the crystal correlated with the exposure time

Below is a **stand-alone Python script** that

- reads the JSON you posted (or any similar list of dataset dictionaries),
- extracts the **overall high resolution limit** (`MXAutoprocIntegrationScaling_overall_resolution_limit_high`) and the **exposure time** (`__elapsedTime` – reported in seconds),
- skips any dataset that does not contain both values,
- creates a scatter plot **resolution (Å) vs. exposure time (s)**, adds a linear regression line for a quick trend, and
- saves the figure as `resolution_vs_exposure.png` and also shows it on screen.

You can copy-paste the script into a file (e.g. `plot_resolution_vs_exposure.py`) and run it with a standard Python 3 installation that has `matplotlib`, `numpy`, and `scipy` installed.

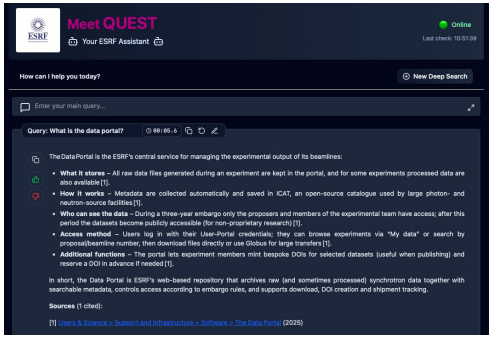
```python
#!/usr/bin/env python3
"""
Plot crystal resolution (high-resolution limit) versus exposure time
for a collection of ESRF “grenades_fastproc” datasets.

The script expects a JSON file that contains a list of dataset objects
exactly like the one you posted.  It extracts:
```

Type your message...

Send

DRACbot can make mistakes. Please check the output with the beamline scientist(s) to be sure.

Related developments

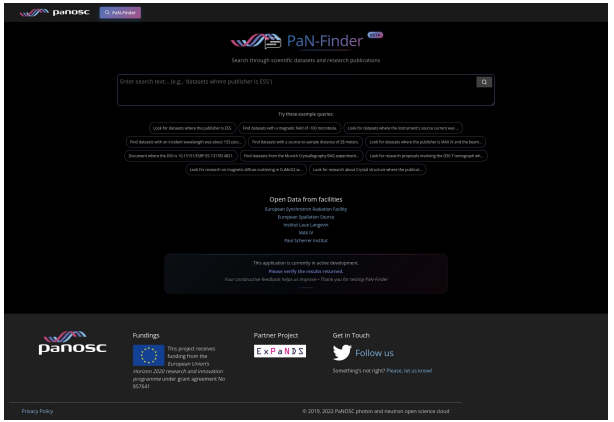


Quest

- QUEST is ESRF's AI assistant for scientific and administrative support.
- Uses Retrieval-Augmented Generation to deliver accurate, factual answers with inline references.
- Draws information from the ESRF extranet and 40,000+ publications in the PUMA corpus.
- Currently under evaluation BDO + UO

PaN-Finder

- AI-powered search tool from the OSCARS project that enhances the PaNOSC Data Portal
- Retrieval-Augmented Generation to provide intelligent, prompt-based searches with summaries and references.
- Developed by the European Spallation Source (ESS) team



Conclusions

What is the current status of AI tool usage in the daily operations of SWG?

- The use of AI, LLMs, and chatbots is modest.
- The perceived usefulness appears to be slightly higher than actual usage.
- Lack of time to experiment, along with ethical and ecological concerns, is the most important barrier.

How can AI be used to enhance user experience?

- A proof of concept (POC) has been implemented to gather feedback.
- The amount of metadata available at the ESRF, together with increasingly powerful LLMs, may provide a valuable tool for users.
- The implementation or adoption of other solutions will likely require several years of development if we aim to fully leverage them, but they could ultimately be a game changer

THANKS

Acceptancy

Article

AMERICAN MARKETING ASSOCIATION

Journal of Marketing
1-20

DOI: 10.1177/00222429251314491
journals.sagepub.com/home/jmx

Sage

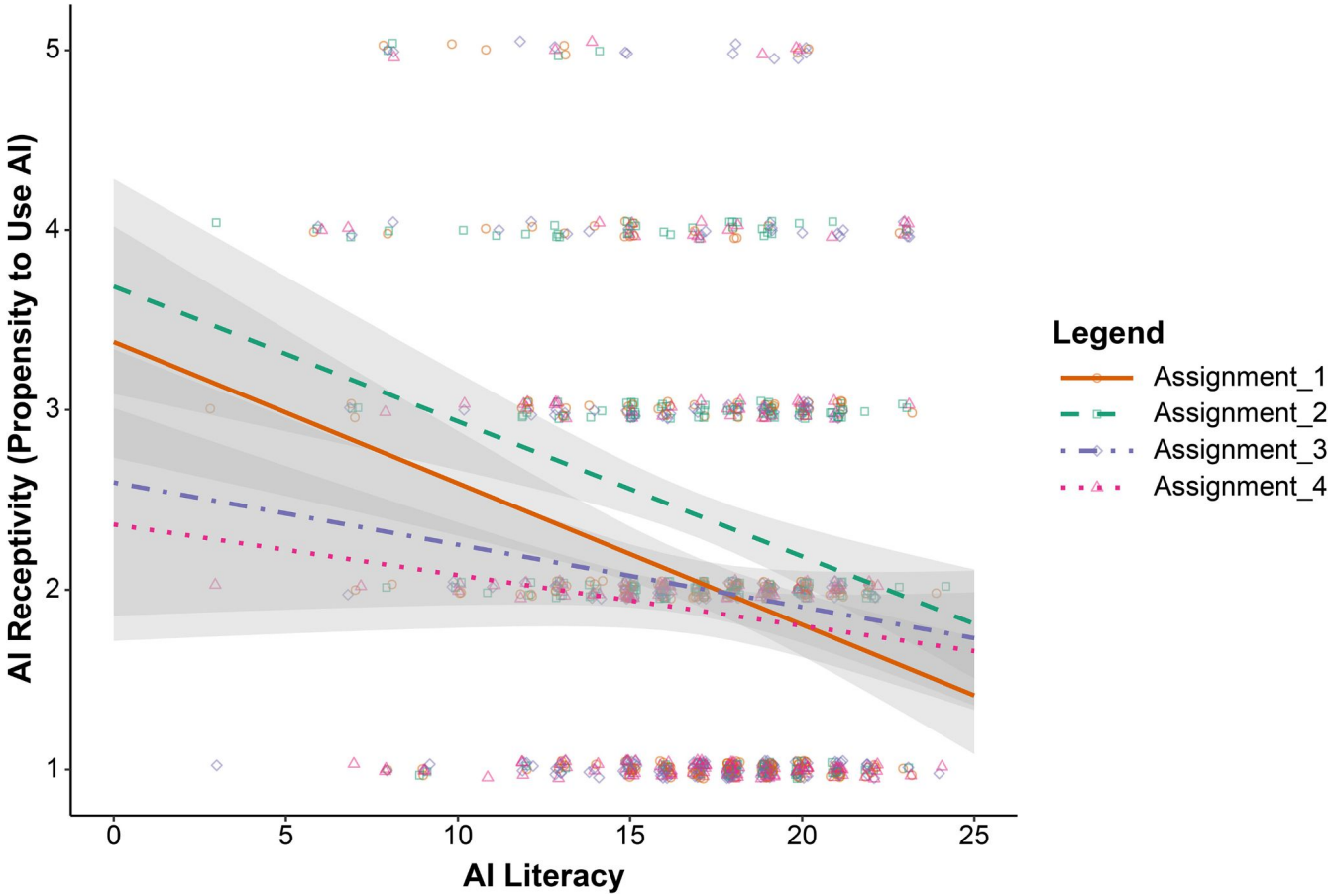
Lower Artificial Intelligence Literacy Predicts Greater AI Receptivity

Stephanie M. Tully, Chiara Longoni, and Gil Appel

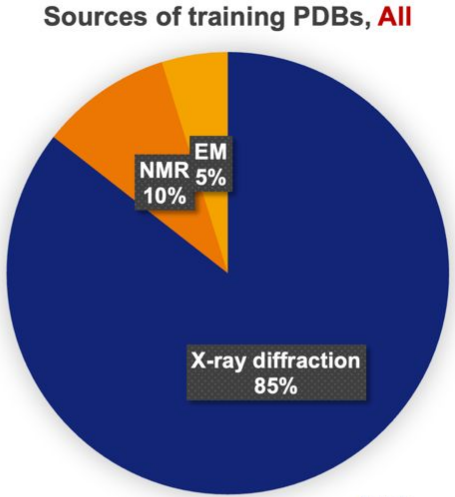
Abstract

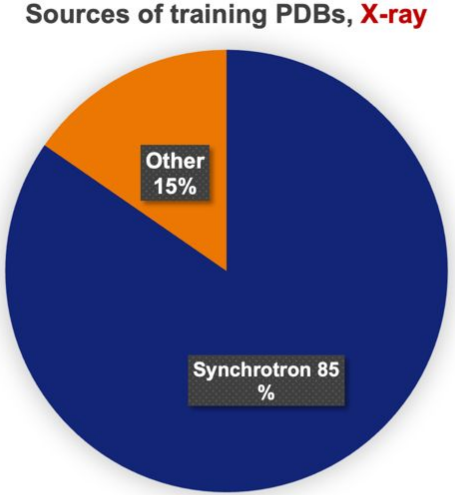
As artificial intelligence (AI) transforms society, understanding factors that influence AI receptivity is increasingly important. The current research investigates which types of consumers have greater AI receptivity. Contrary to expectations revealed in four surveys, cross-country data and six additional studies find that people with lower AI literacy are typically more receptive to AI. This lower literacy–greater receptivity link is not explained by differences in perceptions of AI's capability, ethicality, or feared impact on humanity. Instead, this link occurs because people with lower AI literacy are more likely to perceive AI as magical and experience feelings of awe in the face of AI's execution of tasks that seem to require uniquely human attributes. In line with this theorizing, the lower literacy–higher receptivity link is mediated by perceptions of AI as magical and is moderated among tasks not assumed to require distinctly human attributes. These findings suggest that companies may benefit from shifting their marketing efforts and product development toward consumers with lower AI literacy. In addition, efforts to demystify AI may inadvertently reduce its appeal.

Acceptancy


AI Receptivity (Propensity to Use AI)
5
4
3
2
1
0
5
10
15
20
25
AI Literacy
Legend
Assignment_1
Assignment_2
Assignment_3
Assignment_4

Alphafold





Fun Facts:

- 8.8 % of training set structures are from the ESRF (10% of X-ray structures)
- 7.8 %(62 million) of all training set atoms are from the ESRF (11% of X-ray atoms)
- 11.2 % (481 million) of all training set (total) reflections are from the ESRF